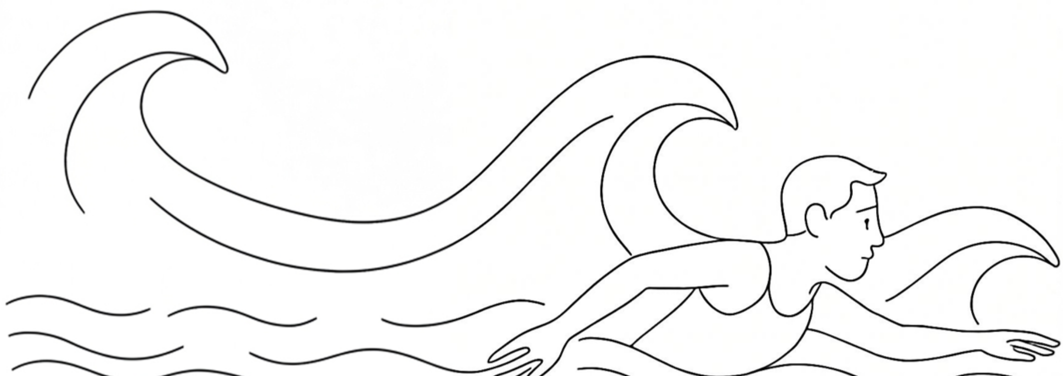
Learning to swim the endless waves: examining self-management and risk factors for (chronic) depression

Solis, E.C.



CHAPTER 7

Psychometric Properties of the Leiden Index of Depression Sensitivity (LEIDS)

Ericka Solis, Niki Antypa, Judith M. Conijn, Henk Kelderman,
Willem Van der Does

Institute of Psychology, Leiden University, Leiden, The Netherlands

Psychological Assessment. 2017; 29(2): 158-17

Abstract

The Leiden Index of Depression Sensitivity (LEIDS) is a self-report measure of cognitive reactivity (CR) to sad mood. The LEIDS and its revised version (LEIDS-R) reliably distinguish between depression-vulnerable and healthy populations. They also correlate with other markers of depression vulnerability, but little is known about the other psychometric properties. Our aim was to examine the factor structure and validity of the LEIDS-R. We used data from the Netherlands Study of Depression and Anxiety (NESDA; *N* = 1,696) and a student sample (*N* = 811) for exploratory (EFA) and confirmatory factor analysis (CFA). CFA showed that model fit of the six-factor structure was satisfactory in the NESDA sample, but some factors were highly correlated. After removing four poor items, EFA yielded an alternative five-factor structure and could not replicate the original six-factor model. Testing for measurement invariance across recruitment groups of NESDA showed support for strong invariance. Due to high interfactor correlations, a bifactor model with one general factor and five specific factors was fitted in two samples. This model supported use of a general factor, but high factor loadings in specific factors supported retaining a five subscale structure. Higher scores on the general factor were associated with a history of depression, especially in participants with a history of comorbid anxiety. We conclude that the LEIDS-R has good psychometric properties. A modified version, LEIDS-RR, comprised of five subscales and a total cognitive reactivity score, is recommended for future research. One of the subscales is suitable as a short form.

Keywords: Cognitive reactivity; LEIDS; depression; psychometric properties; factor structure; validity; relapse

Major depressive disorder (MDD) can be a chronic, lifelong disease (Richards, 2011). Several risk factors for chronicity have been identified, such as number of depressive episodes, last episode severity, high neuroticism, and ongoing psychosocial difficulties (Hardeveld, Spijker, De Graaf, Nolen, & Beekman, 2013; Mueller & Leon, 1999). Cognitive models of depression have stressed the importance of negative cognitions in the development and maintenance of MDD (Beck, 1967; Clark & Beck, 1999; Jarrett et al., 2012).

Cognitive vulnerability factors of depression include rumination, dysfunctional attitudes, and negative cognitive styles (Hankin et al., 2009). These factors are often highly evident during a depressive episode, but the sad mood and other depressive symptoms are often interwoven and inseparable from such negative thinking. Rumination, for instance, relates to the focusing on depressive symptoms and its causes and consequences (Nolen-Hoeksema, 1991). Because depression is a recurrent disorder, vulnerability factors that can be detected outside the depressed state and can predict relapse are important. Teasdale (1988), in his *differential activation* hypothesis, amended cognitive models to state that the negative cognitions that remain latently present in some individuals even when depressive symptoms are in remission, may be reactivated not only by life events or stress, but also by negative mood states. Research has indeed shown that negative cognitions become activated during naturally-occurring or experimentally-induced sad mood states, in remitted depressed but not in never-depressed individuals (Hollon, DeRubeis, & Evans, 1987; Ingram & Ritter, 2000; Miranda, Gross, Persons, & Hahns, 1998; Taylor & Ingram, 1999). The ease with which these negative cognitions can be (re-

)activated by mild dysphoria is called cognitive reactivity (CR) (Teasdale, 1988; Segal, Gemar, & Williams; 1999).

CR is typically measured by assessing negative cognitions before and after inducing a sad mood, using parallel forms of the Dysfunctional Attitudes scale (DAS; Weissman, 1979). The difference in scores, or DAS change score, reflects CR (Miranda et al., 1998; Segal et al., 1999). This research has shown that recovered-depressed individuals tend to have higher CR (i.e., higher DAS change scores) than never-depressed individuals (Hedlund & Rude, 1995; Miranda, Persons, & Byers, 1990; Miranda & Persons, 1988; Miranda et al., 1998). High DAS change scores also predicted a shorter time to relapse in remitted depressed individuals, independent of prior treatment (Segal et al., 1999; Segal, Kennedy, Gemar, Hood, Pedersen, & Buis, 2006).

Although DAS change scores are the current standard for the measurement of CR, this index is far from ideal. Basic psychometric properties, such as test-retest reliability, are unknown. Secondly, the DAS is not very sensitive to change, so mean CR scores are typically quite low. Thirdly, up to 20% of participants in mood induction experiments fail to react with increased sadness (Martin, 1990), which suggests that CR may not be measurable in this specific group. Finally, the parallel versions of the DAS are not interchangeable and provide different mean scores. In experiments, this can be statistically corrected (e.g., by including order of administration in the analyses), but these change scores have no meaning for an individual case in clinical practice. Using the same version of the DAS before and after mood induction requires repeated administration of the same items within seven minutes, likely leading to demand effects. Given these limitations, it is not surprising that research has produced some inconsistent findings. Several studies did not find any significant differences in DAS change scores between recovered depressed and never-depressed groups (e.g., Brosse, Craighead, & Craighead, 1999; Van der Does, 2005; Van der Does, Manthey, & Hermans, 2012), despite the fact that the former group is more vulnerable to depression.

The Leiden Index of Depression Sensitivity (LEIDS; Van der Does, 2002a) was developed to measure CR without the use of mood induction. Whereas the DAS uses 'absolute' wordings (e.g., "I feel like a failure every time I make a mistake"), the LEIDS items are phrased conditionally (e.g., "*When I feel down,* I am more bothered by perfectionism"), similar to the Anxiety Sensitivity Index (ASI; Peterson & Reiss, 1992). All items are also worded in the comparative form. For instance, rather than asking whether participants see themselves as irritable, the LEIDS asks whether they become *more* irritable when in a state of dysphoria. Prior to answering the items, participants are instructed to take a moment to imagine or recall a mildly-sad mood state. An initial study using a 26-item version of the LEIDS showed that four subscales could be distinguished. Moreover, the total score was moderately associated with the DAS change score before and after a mood induction and differentiated between never-depressed and recovered-depressed samples (Van der Does, 2002a; 2002b). The LEIDS was later revised and expanded (LEIDS-R; Van der Does & Williams, 2003). Since then, the 34-item LEIDS-R has been in use with six subscales, based on an unpublished factor analysis using a sample of N = 499.

Studies from independent research groups have consistently shown that recovered-depressed patients have higher LEIDS-R scores than never-depressed individuals (e.g., Van der Does, 2005; Merens, Booij, & Van der Does, 2008;

Booij & Van der Does, 2007; Moulds, Kandris, Williams, Lang, Yap, & Hoffmeister, 2008; Antypa, Van der Does, & Penninx, 2010; Raes, Dewulf, Van Heeringen & Williams, 2009; Kruijt et al, 2013). Two attempts to validate the revised scale scores against DAS change scores were unsuccessful. This was however due to the failure of the DAS change scores to differentiate between vulnerable and healthy samples (Van der Does, 2005; Van der Does et al., 2012). As far as we are aware, in more than ten studies, the LEIDS(-R) has not failed to replicate this basic distinction, and as such, we may conclude that the LEIDS-R scores have higher criterion validity than the DAS scores. We cannot be certain that the LEIDS-R measures CR, or rather another aspect of depression vulnerability. We do know, however, that the scale measures a clinically relevant construct. High scores on the LEIDS-R predicted the first onset of depression over the next two years (Kruijt et al., 2013). Furthermore, mindfulness-based cognitive therapy reduced LEIDS-R scores (Raes, Dewulf, Van Heeringen, & Williams, 2009). Finally, LEIDS-R scores were associated with a genetic vulnerability marker of depression (the serotonin transporter gene polymorphism; Antypa & Van der Does, 2010). LEIDS scores also predicted the depressive response to a serotonergic manipulation (tryptophan depletion) in remitted depressed patients (Booij & Van der Does, 2007). In summary, many studies have shown that the scale is a convenient measure of depression vulnerability.

The subscales of the LEIDS-R have also been informative. The LEIDS-R subscale profile during remission was shown to be associated with the symptom profile during the prior depressive episode. For instance, remitted depressed patients who had been irritable during a prior episode had higher aggression reactivity scores than remitted depressed patients who had not been irritable

(Verhoeven et al., 2011). Individuals who had been suicidal during their latest depressive episode had specifically higher scores on the hopelessness subscale during remission than remitted depressed patients who had not been suicidal (Williams, Van der Does, Barnhofer, Crane, & Segal, 2008). This latter finding was replicated in a larger sample (Antypa et al., 2010), which also showed that remitted patients with a past suicide attempt had higher scores not only on hopelessness reactivity but also on aggression reactivity. In women, aggression reactivity was associated with variations in the monoamine-oxidase-A gene (Verhoeven et al., 2012).

Some of these findings await independent replication, but they support the validity of the LEIDS-R as an index of depression vulnerability. The other psychometric properties of the scale scores also remain to be investigated. Therefore, the aims of the current study were (1) to compare the commonly-used six-factor structure with alternative factor structures using exploratory and confirmatory factor analyses, (2) to evaluate measurement invariance of the factor structure across recruitment groups using multigroup analysis, and (3) to assess criterion validity, discriminant validity, and internal consistency of the LEIDS-R scale scores.

Method

Participants

The present study used two samples: a mixed clinical/healthy sample and a college student sample. Participants in the mixed sample were selected from the Netherlands Study of Depression and Anxiety (NESDA; Penninx et al., 2008) and were predominantly of North-European descent (96.3%). NESDA includes 2,981 individuals (18 – 65 years of age) recruited from three research locations (i.e., Leiden, Groningen and Amsterdam) and settings (community, primary care and specialized outpatient mental health care clinics). Participants completed a four-hour medical and psychological test battery at one of the research sites and additional questionnaires at home (Penninx et al., 2008). The protocol was approved by the Ethical Review Board of each participating location, and each participant signed a consent form. Detailed information regarding the rationale, sample characteristics and methodology is described elsewhere (Penninx et al., 2008).

The college student sample consisted of 899 individuals (between 17 and 45 years of age), predominantly of North-European descent (90.6%), who had been recruited for a number of studies at various sites at Leiden University by means of advertisements. The research was approved by the Ethics Committee of the Leiden University Medical Center in the Netherlands, and each participant signed a consent form.

Exclusion criteria. In both samples, participants with current depression were excluded because CR is the cognitive response to mild sadness, and we expect that depressed individuals would have difficulty imagining a mild sad mood. Although a mild mood would be rated as a 3 or 4 on a scale of 1 to 10, a currently-depressed individual may rate their mood at 9 or 10 on the scale. While the structure is not expected to greatly alter by including currently-depressed individuals, it would not make sense theoretically to include them for the aforementioned reason.

Of the 2,986 NESDA participants, 1,115 had a current diagnosis of MDD (i.e., fulfilling diagnostic criteria in the past six months), as diagnosed using the lifetime Composite International Diagnostic Interview (CIDI-WHO, version 2.1; Robins et al., 1988), and were excluded. Another 164 did not return the questionnaires and six completed no or only a few items of the LEIDS-R. The final sample included 1,696 participants (n = 431 from the community, n = 1108 from primary care, and n = 157 from specialized care). Sixty-eight participants had missing scores, missing at random, on one to four items, or less than .5% missing data per item.

In the college sample, psychiatric disorders were not clinically diagnosed. Instead, we screened the participants in the college sample for current depression and anxiety using the Hospital Anxiety and Depression Scale (HADS; Zigmond & Snaith, 1983). Those who were more likely to be experiencing a current depressive episode, as indicated by a cut-off score of ≥ 8 on the depression subscale (specificity of .79 and sensitivity of .83; Bjelland, Dahl, Haug, & Neckelmann, 2002), were excluded. Of the 899 participants, 74 who had screened positive for a current depressive disorder and 14 who had not completed the LEIDS-R were excluded. No data were missing in the college sample, and history of MDD or anxiety was not assessed. Possible current anxiety disorders were identified using a cut-off score and ≥ 8 on the anxiety subscale (specificity of .88 and sensitivity of .80; Olsson, Mykeltun, & Dahl, 2005). The demographic and clinical characteristics of both samples are shown in **Table 1**.

Table 1. *Sample Characteristics of the Total NESDA Sample and Student Sample*

	NESDA		Student
	No MDD	Remitted MDD	Nondepressed
Characteristic	(*n* = 966)	(*n* = 730)	(*n* = 811)
Age, mean (SD), years	42.0 (14.3)	44.1 (12.6)	21.0 (3.6)
Gender			
Female	616 (63.8%)	523 (71.6%)	626 (77.2%)
Education, mean (SD), years	12.6 (3.3)	12.4 (3.2)	n/a
Ethnicity			
North-European ancestry / Other	925 / 41	708 / 22	729 / 82
Number of previous episodes			
Single / Recurrent	0 / 0	371 / 359	n/a
Current comorbid diagnoses	248 (25.7%)	267 (36.6%)	248 (30.6%)[a]
Social phobia	47	67	n/a
Panic disorder	69	61	n/a
Agoraphobia	22	29	n/a
Generalized anxiety disorder	21	19	n/a
Dysthymia	6	4	n/a
Multiple comorbid diagnoses	83	87	n/a
Any comorbid diagnoses (lifetime)	358 (37.1%)	511 (70.0%)	n/a

Note. MDD = Major Depressive Disorder. Remitted MDD = no episode in 6 months. Data regarding lifetime diagnoses were not available for the Student Sample. [a]Screened positive on anxiety subscale of the Hospital Anxiety and Depression Scale; cut-off score ≥ 8. n/a = not assessed.

Instruments

LEIDS-R. The Leiden Index of Depression Sensitivity Revised (LEIDS-R) is a 34-item self-report measure of CR that has six subscales: Hopelessness / Suicidality (HOP; "When I feel sad, I feel as if I care less if I lived or died."), Acceptance / Coping (ACC; "When I am in a sad mood, I feel more like myself"), Aggression (AGG; "When I feel bad, I feel more like breaking things."), Perfectionism / Control (CTR; "When I feel somewhat depressed, I think I can permit myself fewer mistakes."), Risk Aversion (RAV; "When I feel sad, I feel less inclined to express disagreement with someone else."), and Rumination

(RUM; "When I feel sad, I spend more time thinking about the possible causes of my moods"). Participants rate the extent to which each statement applies to them on a five-point Likert scale (0 = *not at all*; 1 = *a bit*; 2 = *moderately*; 3 = *strongly*; and 4 = *very strongly*). Items are all positively worded for cognitive reactivity and the total score is obtained by summing item responses. A higher total score indicates stronger cognitive reactivity.

ASI. The Anxiety Sensitivity Index (ASI; Reiss, Peterson, Gursky, & McNally, 1986) was used in the NESDA study and is a 16-item scale (total score ranging from 0 to 64) measuring the fear of possible negative consequences stemming from anxiety symptoms (Taylor, Koch, McNally, & Crockett, 1992). Research has repeatedly shown a higher-order unifactorial structure (Zinbarg & Barlow, 1996; Zinbarg, Barlow, & Taylor, 1997; Mohlman & Zinbarg, 2000; Vujanovic, Zvolensky, Bernstein, Feldner, & McLeish, 2007) and a lower-order multifactorial structure. However, studies have also proposed two-factor (Schmidt & Joiner, 2002), three-factor (Mohlman & Zinbarg, 2000; Zinbarg & Barlow, 1996; Zinbarg et al., 1997) or four-factor structures (Cox, Parker, & Swinson, 1996; Vujanovic et al., 2007). ASI total scores have shown adequate test-retest reliability (r = .75 for 2 weeks), criterion validity (Reiss et al., 1986), and high internal consistency (especially with regard to the total score). The ASI scores have also shown good convergent and discriminant validity with other mood scales in line with the anxiety sensitivity theory (Vujanovic et al., 2007). In the NESDA sample, internal consistency of the total score was high (α =.89).

NEO-FFI. The NEO Five Factor Inventory (NEO-FFI; Costa & McCrae, 1992) is a 60-item self-report questionnaire used to measure the Big Five personality domains, each comprised of 12 items. In the Dutch translation, the five personality factors were clearly demonstrated in factor analysis and internal consistency was considered adequate (Hoekstra, Ormel, & de Fruyt, 1996). The current study used the Neuroticism scale. Neuroticism scores have shown good reliability (Robins, Franley, Roberts & Trzesniewski, 2001) and convergent validity (e.g., Anisi, Majdian Joshanloo, & Gohari-Kamel, 2012). In the NESDA sample, internal consistency for Neuroticism was high, $\alpha = .93$. Both the ASI and the NEO-FFI were administered during the NESDA study but not in the student sample.

Statistical analyses

We conducted five sets of analyses. The first set of analyses focused on the current six-factor LEIDS-R structure. Using the total NESDA sample, we examined basic psychometric properties, such as item response frequencies, item-total correlations, and internal consistency of the current six-factor LEIDS-R. Confirmatory factor analysis (CFA) was then used to assess the appropriateness of the theory-driven six-factor structure of the LEIDS-R (Van der Does & Williams, 2003).

In the second set of analyses, we investigated alternative factor structures for the LEIDS-R by means of exploratory factor analyses (EFA) and thereafter confirming the models in CFA, using two split samples of the NESDA sample. NESDA sample 1 was used to identify models in EFA and NESDA sample 2 was used to confirm the models. The two subsamples were statistically equivalent in sex distribution, age, education level, birth country, MDD type, lifetime diagnoses, total LEIDS-R and subscale scores (all *p*s > .10, two-tailed). As a second validation of the selected models, the student sample was used in CFA.

The third set of analyses focused on measurement invariance. Demonstrating measurement invariance, or identical dimensions across groups, allows one to make meaningful comparisons across such populations (Raykov, 2004), such as recruitment groups. To test for measurement invariance in the new structure across recruitment groups of the NESDA sample (community, n = 431, primary care, n = 1108, and specialized care, n = 157), three models were estimated (Berrios, Kellett, Fiorani, & Poggioli, 2015; Millsap & Yun-Tein, 2004; Spinhoven, Penninx, Hickendorff, van Hemert, Bernstein, & Elzinga, 2014) and compared using model-fit indices. These are described in further detail below.

Because we expected high correlations among the LEIDS-R factors, which would suggest a higher-order or hierarchical factor model (Thompson, 2004), we also tested bifactor models and the reliability thereof in the fourth set of analyses. The bifactor models included a general CR factor, on which all items were allowed to load, and specific factors on which only the corresponding items were allowed to load. All factors were specified to be orthogonal to one another and to the general CR factor (Reise, Bonifay, & Haviland, 2013). To test the scale score reliability, we calculated Cronbach’s alpha, α; omega, ω (McDonald, 1999; Revelle & Zinbarg, 2009; Zinbarg, Revelle, Yovel, & Li, 2005); and the hierarchical omega coefficient, ω_h (McDonald, 1970) for the confirmatory bifactor model. However, the Cronbach’s alpha reliability estimate can be biased when including the variance of group factors (Widhiarso & Ravand, 2014), and while omega assumes that items are congeneric rather than tau equivalent (Graham, 2006), the interpretability of omega, similar to alpha, may depend on unidimensionality (Graham, 2006; McDonald, 1999; Rodriguez, Reise, & Haviland, 2015). Therefore, omega hierarchical may be a better reliability estimate as it can account for measurement error due to the variance of the specific factor scores (McDonald, 1999; Zinbarg, Barlow, & Brown, 1997; Zinbarg et al., 2005). $\omega_{h\ total}$ was calculated for the total score, which is the percent of total score variance associated with variation of one general factor, and for each subscale, $\omega_{h\ subscale}$, which is the percentage that the subscale scores provide reliable variance after accounting for general factor (Olatunji, Ebesutani, & Reise, 2015; Rodriguez et al., 2015). We also computed the explained common variance (EVC) of the total scale which represents the percentage of common variance attributable to the general factor in a bifactor model (Bentler, 2009; Reise, Moore, & Haviland, 2010; Rodriguez et al., 2015; Ten Berge & Sočan, 2004). Hierarchical omega coefficient does not clarify dimensionality and cannot indicate general factor strength; therefore ECV is a useful tool to evaluate the relative strength of the general factor (Rodriguez et al., 2015).

Moreover, in the fifth set of analyses, we assessed criterion and discriminant validity of the scores of the revised LEIDS-R and its subscales. All validity estimates were compared to those of the original LEIDS-R.

Factor analyses. Considering the five-point Likert scale of the LEIDS-R, we conducted EFA and CFA for ordered categorical (ordinal) data in MPLUS version 7.2 using the robust weighted least squares means and variance adjusted (WLSMV) estimator which uses the diagonal weight matrix and polychoric correlations (p. 603, Muthén & Muthén, 1998-2012). Thresholds were computed from the univariate marginals, and oblique Geomin rotation was selected to allow correlation between factors. Missing values were dealt with using pairwise deletion, which provides unbiased parameter estimates when data are missing at random. To determine the number of factors to be retained in the EFA, scree plots of eigenvalues, parallel analysis (PA; Horn, 1965; Glorfeld, 1995) and lowest minimum average partial correlations by means of Velicer's minimum average partial test (MAP; Velicer, 1976) were examined. PA and MAP for categorical data were conducted using the *random.polychor.pa* package in *R* version 3.1.1 (The R Foundation for Statistical Computing Platform, 2014).

The estimated EFA models were evaluated by examining the number of large residual correlations (i.e., > |.1|), factor loading pattern, interpretability of the factors, and inter-factor correlations. Low item communalities (<.40; Costello & Osbourne, 2005) were also inspected. We also used the root-mean-square error of approximation (RMSEA; Steiger, 1990) and the standardized root-mean-square residual (SRMR; Bentler, 1995) for EFA model selection (Muthén & Muthén, 2009). CFA models were evaluated by the number of weak standardized factor loadings (< .50), cross-loadings (as identified by modification indices, which reflect improvement in model fit by freeing a particular model parameter), inter-factor correlations, and several model-fit indices: RMSEA, comparative fit index (CFI; Bentler, 1990), and the Tucker-Lewis index (TLI; Tucker & Lewis, 1973). Values ≤ .08 and ≤ .05 on the RMSEA indicate an acceptable and good fit, respectively (Browne & Cudeck, 1993; McCallum, Browne, & Sugawara, 1996). For the SRMR, values < .06 indicate good model fit (Hu & Bentler, 1999). For the CFI and TLI, values ≥ .90 and ≥ .95 indicate acceptable and good model fit, respectively (Hu & Bentler, 1999). In the models, inter-factor correlations >.70 were considered high. Factors correlated > .85 brought the independence of the factors into question and suggested combining the factors.

Measurement invariance. Three models were used in multigroup analysis to test for measurement invariance across recruitment groups (community, n = 431, primary care, n = 1108, and specialized care, n = 157). In the baseline model to test configural invariance, or whether the same pattern of item loadings and factors were found across groups, all parameters were allowed to vary across recruitment groups. To test metric invariance, or the relative strength of the relations between subscale items and their respective underlying conceptual factors, we tested a second model in which the factor loadings were held equal across group and thresholds were allowed to vary. In the third model which tested scalar invariance, or whether the groups used the response scale similarly, both factor loadings and item thresholds were held equal across groups. The commands for the various invariance models in MPlus version 7.2 (i.e, configural, metric, and scalar) were utilized. After each parameter change, chi-square difference tests (χ^2 difftests) were used to test each measurement invariance model with the next progressively constrained model. The χ^2 difftests adjust the chi-square square values and degrees of freedom to obtain interpretable p values (Muthén & Muthén, 1998-2012). We evaluated model fit using the TLI, CFI and RMSEA, using the same aforementioned criteria of the factor analyses, and the difference between models using ΔCFI $\leq$ .01 (Cheung & Rensvold, 2002) and ΔRMSEA $\leq$ .015 (Chen, 2007) to indicate acceptable measurement invariance.

Criterion and discriminant validity. Criterion validity was assessed in the NESDA sample by correlating the LEIDS-R total and subscale scores with MDD (lifetime) history (0 = no MDD, 1 = past MDD) as based on the CIDI-WHO. A positive relationship between LEIDS-R scores and MDD history was regarded as support for criterion validity as it reflects ability of the LEIDS-R to identify individuals with and without a previous depressive episode.

To assess discriminant validity, the LEIDS-R was correlated with the ASI and the NEO-FFI Neuroticism. Moderate correlations were taken as support for discriminant validity. Spearman's rho (r_s) was used for all correlations due to non-normal distribution of the LEIDS-R subscales. Furthermore, we compared LEIDS-R scores between participants with no history of depression or anxiety, a history of either depression or anxiety, or a history of comorbid depression and anxiety by means of a one–way ANOVA. Considering that an anxiety disorder is a risk factor for future depressive episodes (Pine, Cohen, Gurley, Brook, & Ma, 1998; Wittchen, Kessler, Pfister & Lieb, 2000), we expected high LEIDS-R scores for both remitted depressed and (remitted) anxious patients. Since comorbidity is associated with a poor prognosis of illness (Sherbourne & Wells, 1997), we expected particularly high LEIDS-R scores for participants with a history of both depression and anxiety disorder. We repeated these analyses for the ASI and the NEO-FFI Neuroticism scores to assess the relative discriminative ability of the LEIDS-R. For the ANOVA tests, we excluded participants with a current anxiety disorder in the NESDA sample.

Results

Psychometric Properties of the Original LEIDS-R

Item properties and internal consistency. Inspection of frequency tables showed that for 12 items (items 7, 18, 19, 21, 26, 30, 34 and all five ACC items), at least 87% of the sample responded *Not at all* (0) or *A bit* (1). In particular, the ACC items were skewed as each had less than 2% responses in the categories *Strongly* (3) and *Very Strongly* (4). We inspected the items for severe skewness and kurtosis using the criteria that skewness is > 2 in absolute value or kurtosis is > 7 in absolute value for samples larger than 300 (Kim, 2013), which identified items 4, 10, 18, 19, 24, 26, 28, 30, and 34. For each item, except item 8, the corrected item-total correlation within their respective subscale was acceptable and ranged from .34 to .67. Item 8 (CTR subscale) correlated weakly with the corrected total CTR score (r_s = .09), suggesting that this item may not be appropriate for measuring the concept of perfectionism/control. Cronbach's alpha for the LEIDS-R was α = .92. The subscales ACC (α = .64), CTR (α = .66), AGG (α = .80), RAV (α = .81), RUM (α = .82), and HOP (α = .83) showed moderate to high internal consistency. Most subscale scores were at least moderately inter-correlated (r_s = .41 to .77), apart from the ACC subscale which correlated moderately with CTR (r_s = .46) but below .32 with the other subscales.

Confirmatory factor analysis. CFA was used to test the model fit of the original six-factor structure of the LEIDS-R in the total NESDA sample. Examination of the residual correlations indicated that 13.90% and 2.14% were > |.1| and > |.2|, respectively. The model-fit indices suggested an acceptable model fit (RMSEA = .063, TLI = .923, and CFI = .930). Factor loadings ranged from .51 to .87, except for items 8 ("*go out and do more pleasurable activities*") and 19 ("*work harder*"), which loaded weakly, .02 and .34, respectively, on the CTR factor. The RAV and RUM factors (r_s = .93), and the CTR factor with both RAV (r_s = .83) and RUM factors (r_s = .83) were strongly correlated. The RAV and RUM correlation was > .85, which suggested combining the factors. The remaining inter-factor correlations ranged from .36 (ACC with AGG) to .79 (HOP with RUM), which were acceptable. The standardized parameter estimates and inter-factor correlations of the six-factor structure can be found in the supplemental material. In summary, the six-factor structure of the LEIDS-R seems inappropriate considering the high correlations between several factors. In subsequent analysis, we examined alternative structures for the LEIDS-R.

Examination of Alternative Factor Structures of the LEIDS-R

Identification of poor items. EFA was conducted in NESDA subsample 1 to investigate alternative first-order models for the LEIDS-R. Eigenvalues (with percent variance explained) of the third, fourth, fifth, and sixth factors were 2.30 (51.31%), 1.70 (56.32%), 1.25 (60.00%), and 1.04 (63.05%), respectively, with all remaining eigenvalues less than 1. The examined criteria indicated an unequal number of factors: the Guttman criterion (eigenvalues > 1) supported a six-factor model; the scree plot suggested either a four- or five-factor model; PA revealed an 11-factor model; and MAP supported a four-factor model. Because PA may overestimate the number of factors in large sample sizes (Warne & Larsen, 2014), this result was disregarded.

EFA models including four-, and five--factors were estimated and further inspected. The four- and five-factor models had between 5-13 low item communalities (<.40), respectively. Further analyses revealed that item 6 ("*more busy keeping images and thoughts at bay*"), item 8 ("*go out and do more pleasurable activities*"), and item 33 ("*think about how my life could have been different*") had substantial cross-loadings in two or more EFA models, including the five-factor model, and could be considered for deletion. In English, items 6 and 8 have a formal, rather than colloquial formulation (but that is less the case in the original Dutch scale, which was used in the present studies). Item 8 had a highly skewed frequency distribution with only 5.7% responding *Strongly* or *Very Strongly*. Item 33 was deleted due to its conceptual ambiguity; it pertained to both hopelessness and rumination, but it could also be construed as a *positive* coping strategy. Although item 1 ("*I can only think positive when I am in a good mood*") did not have particularly poor statistical properties, it was deleted because the wording of the item differed somewhat from the usual conditional formulation. In summary, items 1, 6, 8, and 33 were deleted from the questionnaire, which was renamed LEIDS-RR.

Exploratory factor analysis. We re-estimated the four- and five-factor models using the remaining 30 items of the LEIDS-RR in NESDA sample 1. Eigenvalues (with percent variance explained) of the third, fourth, fifth, and sixth factors were 2.28 (53.72%), 1.65 (59.23%), 1.21 (63.25%), and 1.03 (66.67%), respectively, with all remaining eigenvalues less than 1. Results of the eigenvalue and MAP analyses using 30-items were similar compared to the 34-item EFA results.

The four-factor model showed good model fit (RMSEA = .053; SRMR = .045) but only three items (from the HOP subscale) loaded distinctly on factor three. As in the three-factor model, factor one contained 15 items. We concluded that this model was not optimal mainly due to the low number of items loading highly on factor three.

The five-factor model showed good model fit (RMSEA = .045; SRMR = .034) and an interpretable factor loading pattern. **Table 2** shows the EFA parameter estimates of the five-factor model. Inter-factor correlations of the five-factor model can be found in the supplementary material. Primary factor loadings ranged from .37 to .94. Four factors corresponded to the original subscales of the LEIDS-R and are named accordingly: Acceptance (5 ACC items, 1 CTR item), Control (4 CTR items, 1 RUM item), Aggression (6 AGG items), and Hopelessness (5 HOP items). Item content supported the empirical subdivision of items into factors. For instance, item 19 ("*I work harder when I feel down*"), originally a CTR item, fit conceptually in the ACC factor as it reflects a positive coping strategy. Item 32 ("*When I feel sad, I spend more time thinking about the possible causes of my moods*"), originally a RUM item, also fit conceptually in the CTR factor. Items 3 and 32 (CTR) and items 5 and 17 (HOP) had substantial cross-loadings, however, obscuring the CTR and HOP factors. The fifth factor of the five-factor model combined most original RAV and RUM items and related conceptually to *avoidant coping* (AVC).

Table 2. *Standardized Factor Loadings and Communalities of the 30-Item Five-Factor Model Derived from EFA in NESDA Sample 1*

LEIDS-RR Item Statements (Item no.)	Original Subscales	EFA[a] ACC	AVC	CON	AGG	HOP	Com[a]
More creative than usual (4)	ACC	.63	*.35*				.46
More helpful (10)	ACC	.66					.48
Better intuitive feeling for meaning (15)	ACC	.56					.59
Work harder (19)	CTR	.48		*.36*			.41
Feel more like myself (24)	ACC	.53					.40
Nicer than usual (28)	ACC	.72					.52
Take fewer risk (2)	RAV		.50	*.29*			.51
Less inclined to disagree (11)	RAV		.57				.48
Feel easily overwhelmed (13)	RUM		.43	*.29*			.63
Avoid difficulties (14)	RAV		.71				.65
Less able to cope with everyday tasks (20)	RUM		.63		*.31*		.59
Want to escape everything (23)	RAV		.63				.69
Neglect tasks (25)	RUM		.76	*.28*	*.44*		.71
Problems concentrating (27)	RUM		.70				.61
Think about what my mood says about me (3)	CTR		*.29*	.37			.44
Permit myself fewer mistakes (12)	CTR		*.31*	.56			.62
Bothered by my perfectionism (16)	CTR			.69			.59
Higher need for control (31)	CTR			.69			.40
Think about the causes of my mood (32)	RUM		*.29*	.37			.38
Do more things I will regret (7)	AGG				.51		.46
Feel like breaking things (18)	AGG				.64		.63
Bothered by aggressive thoughts (21)	AGG				.61		.65
Feel cynical or sarcastic (22)	AGG				.58		.49
Take more risks (26)	AGG				.65		.69
Lose my temper more easily (29)	AGG				.64		.60
Feel hopeless about everything (5)	HOP		*.33*			.38	.62
Care less if I lived or died (9)	HOP					.80	.72
Think I can make nobody happy (17)	HOP			*.33*		.39	.60
Feel I would be better off dead (30)	HOP					.94	.88
Think about harming myself (34)	HOP					.82	.72

Note. NESDA Sample 1, *N* = 856. Items 1, 6, 8, and 33 were removed from analyses. Cross-loadings ≥ .25 are displayed and provided in italics. [a]Comm = item communalities.

To conclude, considering model fit and conceptual interpretation of the factors, the five-factor model was preferred over the alternative models. The appropriateness of the 30-item, five-factor model (LEIDS-RR) was then assessed by means of CFA in two different cross-validation samples.

Cross-Validation of the Five-Factor Model for the 30-item LEIDS-RR

Using CFA, we examined the EFA-derived five-factor model in NESDA sample 2 and the student sample, along with three competing first-order factor models: a one-factor model, the original six-factor model, and the four-factor EFA-derived model. Next to that, we estimated two bifactor models, which posited a general CR factor and domain-specific factors that corresponded with the EFA-derived five-factor model and the original six-factor model, respectively. **Table 3** shows the model fit indices of the models tested in CFA.

Table 3. *Model Fit Indices of Competing Models for the 30-item LEIDS-R fitted in NESDA Sample 2 and Student Sample in Confirmatory Factor Analysis*

Model	χ^2	*df*	RMSEA [90% CI]	CFI	TLI
NESDA Sample B					
One-factor	3243.21*	405	.091 [.088 .094]	.879	.870
Four-factor EFA-derived	1575.58*	399	.059 [.056, .062]	.950	.945
Five-factor EFA-derived	1436.38*	395	.056 [.053, .059]	.956	.951
Original six-factor	1578.63*	390	.060 [.057, .063]	.949	.943
Bifactor five-factor	1234.09*	375	.052 [.049, .055]	.963	.957
Bifactor six-factor	3234.23*	493	.057 [.055, .059]	.945	.937
Student Sample					
One-factor	4038.92*	405	.106 [.102, .108]	.704	.682
Four-factor EFA-derived	2156.59*	399	.074 [.071, .077]	.857	.844
Five-factor EFA-derived	2044.88*	395	.072 [.069, .075]	.865	.852
Original six-factor	2209.50*	390	.076 [.073, .079]	.852	.834
Bifactor five-factor	1646.24*	375	.065 [.062, .068]	.896	.880
Bifactor six-factor	1968.68*	493	.061 [.058, .064]	.895	.881

Note. RMSEA = root-mean-square error of approximation; CFI = comparative fit index; TLI =Tucker-Lewis Index. Items 1, 6, 8, and 33 were deleted. *$p < .001$.

Among the first-order factor models, in both the NESDA and student sample, model-fit indices suggested best fit for the five-factor model. Model fit was good in NESDA sample 2 (RMSEA = .056, CFI = .956, TLI = .951), although in the student sample, model fit was poor according to the TLI and CFI (≤ .865) but acceptable according to the RMSEA (.072). All standardized factor loadings were acceptable, ranging from .53 to .89, in the NESDA sample 2. Inter-factor correlations ranged from .40 to .84. The CTR and AVC factors (r_s = .84) and HOP and AVC factors (r_s = .80) were strongly correlated. In the student sample, factor loadings were somewhat lower than the corresponding estimates in the NESDA sample, ranging from .35 to .85. Factor correlations in the student sample were substantially lower, ranging from .16 (AVC with ACC) to .68 (AVC with HOP), suggesting that although highly correlated in the NESDA sample, the AVC, HOP and CTR factors may be more distinct in other population samples.

Further examining the bifactor models indicated that the bifactor model with five specific factors showed improved model-fit indices (RMSEA = .052, CFI = .963, TLI = .957) compared to the first-order five-factor model. **Table 4** shows the standardized factor loadings of the models tested in CFA using both NESDA sample 2 and the student sample.

In the bifactor model with five specific factors, item 2 did not significantly load on its specific factor, AVC, p = .07. Several items, including ACC items 4, 10, 15, and 19, loaded weakly (<.30) on the general factor, but all ACC items loaded well in its respective factor. In general, the item loadings on the ACC factor were substantially higher for the specific factor than for the general factor, while for AVC and CTR, the factor loadings showed the opposite pattern of relative strength. For the HOP subscale, most but not all general-factor loadings were stronger, and for the AGG subscale, the relative strength of the specific and general loadings was mixed. These results suggest that for the AVC, CTR, and HOP subscales, the general CR factor appeared to account for most variance, while for the other subscales either the specific factor was more important (ACC) or the factor loadings show that the relative strength of the general factor is item dependent. In the student sample, the RMSEA fit index showed that an improvement of model fit in the six-factor bifactor model compared to the five-factor bifactor model, .061 versus .065, respectively, but not according to the TLI and CFI. Also, the six-factor bifactor model included eight items with negative or near zero loadings, ranging from -.25 to .10, indicating that the six-factor bifactor model was a poor model. However, inspection of the factor loadings in the five-factor bifactor model showed that the ACC items had low loadings on the general factor (range: .09 - .34) but high loadings on the specific factor, similar to that in the NESDA sample. Other scales showed that the relative strength of the general and specific factors was mixed.

Table 4. *Standardized Factor Loadings of the 30-Item First-Order Five-Factor Model and Bifactor Model with Five Specific Factors from Confirmatory Factor Analysis using NESDA Sample 2 and Student Sample*

	First-Order		Bifactor			
LEIDS-RR Item Statements (Item no.)	1[a]	2[b]	SF1	General1	SF2	General2
More creative than usual (4)	.59	.50	.60	.27	.48	.15
More helpful (10)	.66	.60	.68	.31	.65	.12
Better intuitive feeling for meaning (15)	.70	.66	.58	.37	.64	.20
Work harder (19)	.56	.47	.50	.28	.46	.09
Feel more like myself (24)	.77	.73	.41	.47	.56	.34
Nicer than usual (28)	.77	.74	.70	.39	.74	.19
Take fewer risk (2)	.72	.42	**.09**	.70	.23	.36
Less inclined to disagree (11)	.63	.35	.22	.59	.41	.24
Feel easily overwhelmed (13)	.85	.77	.12	.83	.09	.74
Avoid difficulties (14)	.68	.38	.30	.62	.47	.26
Less able to cope with everyday tasks (20)	.81	.71	.45	.72	.41	.59
Want to escape everything (23)	.82	.74	.41	.74	.39	.64
Neglect tasks (25)	.74	.65	.39	.66	.45	.52
Problems concentrating (27)	.78	.61	.26	.73	.38	.50
Think about what my mood says about me (3)	.72	.67	.25	.64	.13	.55
Permit myself fewer mistakes (12)	.79	.66	.19	.71	.33	.51
Bothered by my perfectionism (16)	.66	.53	.49	.55	.61	.33
Higher need for control (31)	.68	.54	.48	.58	.69	.33
Think about the causes of my mood (32)	.53	.36	.35	.45	.23	.24
Do more things I will regret (7)	.76	.61	.35	.59	.34	.42
Feel like breaking things (18)	.76	.76	.59	.54	.64	.42
Bothered by aggressive thoughts (21)	.78	.76	.62	.55	.75	.38
Feel cynical or sarcastic (22)	.74	.58	.48	.55	.39	.38
Take more risks (26)	.74	.58	.31	.59	.50	.31
Lose my temper more easily (29)	.69	.70	.57	.48	.59	.39
Feel hopeless about everything (5)	.88	.83	.16	.79	.09	.79
Care less if I lived or died (9)	.83	.81	.68	.64	.69	.56
Think I can make nobody happy (17)	.79	.74	.20	.72	.22	.68
Feel I would be better off dead (30)	.89	.85	.57	.72	.68	.59
Think about harming myself (34)	.79	.75	.60	.61	.67	.50

Note. Items 1, 6, 8 and 33 were removed from the analyses. The bifactor model includes five specific factors and one general factor. SF = specific factors. Factor loadings displayed in bold are *NS, p* > .05.
[a]NESDA Sample 2 (1), N = 840.
[b]Student Sample (2), N = 811.

The reliability statistics of the five-factor bifactor model were evaluated using the total NESDA sample. Cronbach's alpha estimates are shown in the supplementary material. Compared to the LEIDS-R with 34-items, Cronbach's alpha estimates remained the same for the total scale and HOP and AGG subscales (α = .92, .83, and .80, respectively), but improved slightly for the ACC subscale (α = .66) and substantially for the CTR subscale (α = .75). Cronbach's alpha of the new AVC subscale was α = .88. The ω estimates for the general CR factor and ACC, AVC, CON, AGG, and HOP subscales were .96, .84, .92, .82, .88, and .93, respectively. The ω_h estimates for the same factors were .87, .62, .13, .22, .39, and .27, respectively. Therefore, the total score and all subscale scores were associated with high alpha and omega reliability estimates. The ω_h for the general factor and the lower ω_h of the subscales after accounting for the general factor supported a strong general CR dimension. However, the ECV was .moderate (.63), suggesting that the LEIDS-R may be multidimensional. The ω_h of the ACC subscale showed a score beyond the general CR factor.

Overall, the bifactor analyses show relative support for keeping the LEIDS-R subscales. The generally substantial factor loadings on the general factor support interpretation of the general CR dimension, but the non-trivial and strong factor loadings on the specific dimensions and the moderate ECV also highlight the importance of accounting for additional multidimensionality.

Measurement Invariance

We tested the measurement invariance between the three recruitment groups (community, primary care, and specialized care) of the NESDA sample.

Unfortunately, due to the low sample size of the specialized care sample (specialized care, n = 157) the analysis could not be run using all three groups. Further testing was conducted using the remaining groups (community, n = 431, primary care, n = 1108). First, we tested a configural model (model 0) to evaluate whether the LEIDS-R invariance by recruitment group demonstrated the same pattern of factors and loadings and satisfactory goodness of fit indices (see **Table 5**). For both recruitment groups, all factor loadings were statistically significant, $p < .001$. The pattern of factor loadings was generally the same between community and primary care groups. The fit statistics for both the configural and metric models were acceptable. When the metric invariance model was tested against the configural model, the $\chi 2$ difftest was statistically different, $p < .05$, which did not support measurement invariance. We therefore systematically examined factor loadings to identify sources of misfit.

Table 5. *Measurement Invariance Models Across Recruitment Groups (Primary Healthcare, N =1108; Community, N = 431)*

Model	$\chi2$ (*df*)	Reference model no.	$\Delta\chi2$(Δ*df*)	CFI	TLI	RMSEA [90% CI]	ΔCFI	ΔTLI	ΔRMSEA
0. Configural	3389.31 (790)**			.933	.926	.065 [.063, .068]			
1. Metric	3368.47 (815)**	0	51.20 (25)*	.934	.930	.064 [.062, .066]	.001	.004	-.001
1A. Metric	3347.41 (806)**	0	14.86 (16)	.935	.930	.064 [.062, .066]	.001	0	0
2. Scalar	3246.06 (886)**	1A	109.47 (80)*	.939	.940	.059 [.057, .061]	.004	.010	-.005

Note. CFI = comparative factor index; RMSEA = root-mean-square error of approximation; TLI = TLI = Tucker-Lewis Index. Categories 3 and 4 were collapsed for items 15 and 28. ΔCFI scores ≤ .01 indicate that that null hypothesis of invariance should be accepted (Cheung & Rensvold, 2002). ΔRMSEA scores ≤ .015 support measurement invariance (Chen, 2007).
In Model 1A, items 3, 5, 13, 18, 20, 25 and 32 were not constrained. * $p < .05$, ** $p < .001$

The items 3CTR, 5HOP, 13RUM, 18AGG, 20RUM, 25RUM, AND 32CTR were freed in a revised metric model (model 1A). Model 1A was compared against the configural model, no statistical (p = .54) or descriptive differences were detected. The scalar invariance model (model 2) showed similar fit indices compared to revised metric model (model 1A) although the $\chi 2$ difftest was statistically different, p < .05, which partially supports the assertion that the observed differences between community and primary care recruitment groups was attributed to CR scores.

Properties of the LEIDS-RR

The correlations among the total and subscale scores for the revised 30-item LEIDS-RR tested in the total NESDA sample are shown in the supplementary material. The inter-correlations between subscales scores ranged from .25 to .71. The AVC score was correlated .91 with the total score whereas the ACC score was correlated .47 with the total score.

The correlations between LEIDS-RR scores, depression history, anxiety sensitivity and neuroticism were also examined for convergent and discriminant validity (see supplementary material). The total score of the LEIDS-RR correlated moderately with depression history (r_s = .33; p < .001). Among the subscales, the HOP and the new AVC subscale had the highest correlations with depression history (r_s = .30 and .32, respectively). The associations between the 30-item LEIDS-RR general factor scores and depression history were essentially equal to those for the LEIDS-R scores. The only notable differences were small increases for the CTR, ACC, and the combined AVC subscale (as compared to the original RUM and RAV scales). The correlation between depression history and Neuroticism was r_s = .31 (p < .001) and between depression history and ASI was r_s = .17 (p < .001).

The LEIDS-RR was weakly to moderately related to the ASI (r_s range, .24 - .53), suggesting that these measures assess sufficiently distinct concepts, especially for the ACC and AGG subscales. Moreover, the LEIDS-RR was weakly- to moderately-related to the Neuroticism scale (r_s range, .25 - .62). While the LEIDS-RR total and most subscales were more related to Neuroticism than to anxiety sensitivity, the ACC subscale remained weakly correlated with Neuroticism.

To assess the relative performance of the LEIDS-RR to distinguish between four lifetime disorders groups (i.e., no disorder history, past anxiety disorder, past depression, past comorbid anxiety and depression), we conducted ANOVA tests for the LEIDS-RR, ASI, and NEO-FFI Neuroticism scale. **Table 6** shows the results including the means and standard deviations of these groups.

All ANOVA tests were statistically significant, p < .001. For the total score and AVC and CTR subscales of the LEIDS-RR, post hoc tests showed that participants with no psychiatric history scored lower than the lifetime anxiety group, p < .05. For the total score, AVC, CTR, and HOP subscales, the depression group scored lower than the comorbid lifetime depression and anxiety group, p < .001. The separate remitted depression and anxiety groups scored similarly, p range .11 - .87, except on the AVC subscale.

Table 6. *One-Way Analysis of Variance of the 30-Item LEIDS-RR Total and Subscales, ASI, and Neuroticism Personality Trait by Lifetime Depression and Anxiety Disorders in the NESDA sample.*

		Groups $M(SD)$			ANOVA			
Test	No History	Anxiety	Depression	Cormorbid	df (1)	df (2)	Welch F	η^2
LEIDS-RR^{x}	15.3 (12.1)^{a}	20.0 (12.2)^{b}	23.3 (13.8)^{b}	29.4 (15.1)^{c}	3	357.5	61	.15
AVC^{x}	6.5 (5.3)^{a}	8.3 (5.5)^{b}	10.1 (5.8)^{c}	12.2 (6.1)^{d}	3	358.1	59.2	.14
ACC^{x}	1.3 (2.0)^{a}	1.4 (2.0)^{a}	1.7 (2.5)^{ab}	2.1 (2.7)^{b}	3	357.2	7.6	.02
CTR^{x}	3.2 (3.1)^{a}	4.5 (3.4)^{b}	4.7 (3.5)^{b}	6.3 (4.4)^{c}	3	348.9	37.2	.10
AGG^{x}	2.7 (2.7)^{a}	3.4 (3.0)^{ab}	3.8 (3.6)^{bc}	4.5 (3.6)^{c}	3	349.3	19.7	.05
HOP^{x}	1.7 (2.3)^{a}	2.3 (2.5)^{ab}	3.0 (2.9)^{b}	4.3 (3.8)^{c}	3	345.5	39.4	.12
ASI^{x}	23.4 (5.8)^{a}	26.2 (6.5)^{bc}	24.9 (6.6)^{b}	27.6 (7.7)^{c}	3	350.6	21.5	.06
Neuroticism^{ə}	27.1 (7.4)^{a}	3.3 (7.2)^{b}	31.7 (7.6)^{b}	34.8 (7.2)^{c}	3	1184	67.5	.15

Note. Participants with current anxiety disorders were excluded from the analyses (N = 1191). Due to a violation of the homogeneity of variance, the Welch F-ratio is reported with adjusted degrees of freedom for the LEIDS-RR total and subscales and the ASI.
df (1) = degrees of freedom between groups. df (2) = degrees of freedom within groups. ASI = Anxiety Sensitivity Index.
Cormorbid = lifetime MDD and anxiety. ^{x}Tamhane's T2 post-hoc tests were used due to different group variances.
^{ə}Neuroticism was measured with the Revised Five-Factor
Inventory (NEO-FFI), df were not adjusted, and Scheffé post-hoc tests for equal group variances were used.
All ANOVA's were statistically significant, $p < .001$. Means with different suffixes (a,b,c,d) differ significantly,
$p < .05$.

The same pattern was observed for the Neuroticism scale. For the ASI, post hoc tests showed a slightly different pattern: the mean score for the depression group was lower than that for anxiety, and the lifetime anxiety group did not differ from the remitted depression group, p = .45, and from the comorbid remitted depression and anxiety groups, p = .37. Only the no psychiatric history group was clearly differentiated from the other groups, p < .01.

Moreover, for both the total LEIDS-RR and Neuroticism scale, post hoc tests showed that the no psychiatric history group scored significantly lower than both the depression group and the comorbid depression and anxiety groups, p < .01. For both measures, Cohen's d showed a moderate effect size for the test between the no psychiatric history group and the separate depression group in the total LEIDS-RR, d = .64 (90% CI [.51 - .76]), and the Neuroticism scale, d = .62 (90% CI [.49 - .74]). Similarly, there was a large effect (>.80;) for the analyses between the no psychiatric history group and the comorbid lifetime depression and anxiety group on the total LEIDS-RR, d = 1.10 (90% CI [.96 - .1.23]), and the Neuroticism scale, d = 1.06 (90% CI [.92 – 1.20]).

Discussion

This paper provides psychometric results on the LEIDS-R and presents an improved, abbreviated version of the questionnaire: the LEIDS-RR. Rearrangement of the items into five subscales, rather than six subscales, the removal of four items and the addition of a general factor improved the model fit of the scale. In the LEIDS-RR, the reactivity subscales are as follows: Hopelessness/Suicidality (HOP), Acceptance Coping (ACC), Aggression (AGG), Control/Perfectionism (CTR), and Avoidant Coping (AVC). Multigroup group analysis indicated that the internal structure of the new model was invariant across two recruitment groups, suggesting that LEIDS-RR scores represented genuine differences in cognitive reactivity between primary care and community recruitment groups, rather than measurement artifacts.

One important difference between the LEIDS-R and LEIDS-RR is the combination of the RAV and RUM subscales into one AVC subscale. Combing these subscales was warranted considering the high correlation between these subscales in the LEIDS-R. The theoretical background of rumination and risk aversion also supports the combining of these subscales. For instance, previous studies using non-clinical samples have shown that depression, rumination, and avoidance (behavioral, cognitive, and experiential subtypes) were associated with one another even after controlling for anxiety (Cribb, Moulds, & Carter, 2006; Moulds, Kandris, Starr, & Wong, 2007). Also, individuals who are more likely to engage in behavioral avoidance are more likely to ruminate, regardless of anxiety symptoms (Moulds et al., 2007). Rumination may also be construed as a method of avoiding active problem solving (Rijsbergen, Kok, Elgersma, Hollon, & Bockting, 2015). While we also explored the possibility of a bifactor model using the original six-factor subscales, the five-factor bifactor model still demonstrated optimal fit. These findings support the combining of the RAV and RUM subscales into an *Avoidance Coping* (AVC) subscale in the LEIDS-RR. Avoidance coping is the maladaptive coping mechanism characterized by cognitive and behavioral efforts to deny, minimize, or avoid dealing directly with stressful demands. It is also closely linked to depression (Cronkite & Moos, 1995; Penley, Tomaka, & Wiebe, 2002; Holahan, Moos, Holahan, Brennan, & Schutte, 2005), which fits in well with the LEIDS-RR.

Moreover, the LEIDS-RR includes a general cognitive reactivity scale and the reliability coefficients indicate that the total scale score could be utilized. The ACC subscale of the LEIDS-RR showed higher factor loadings in its respective factor and a relatively substantial omega hierarchical coefficient, suggesting that the ACC factor measures a distinct concept, differentiated from the general factor. Items from the other subscales also showed relatively higher loadings in their respective factors, and the ECV indicated that the general factor only accounted for approximately 60% of the common variance, thereby suggesting a possible multidimensional model. The results of our CFAs suggest that both overall LEIDS-R total and subscale scores have research and clinical utility. In some cases, a bifactor model that measures cognitive reactivity as a general scale, but allows for multidimensionality by maintaining the subscales due to the variety of item content, provides a better model conceptualizing dimensionality (Reise et al., 2010). Although the concepts of the subscales could be measured by a broader construct of cognitive reactivity, there are instances where multidimensional models should be considered for the testing of relevant theory-driven hypotheses (Reise et al., 2013; Spinhoven et al., 2014). Given the specific findings with some of the subscales in previous research (e.g., Antypa et al., 2010; Verhoeven et al., 2011; Verhoeven et al., 2012) we think that this is the case with the LEIDS-RR. However, results of analyses of subscale scores should be interpreted with caution, so as not to misinterpret or overinterpret scores, given the smaller unique variance provided by most of the subscale scores compared to the total score.

Moreover, the original first-order six-factor structure of the LEIDS-R demonstrated an acceptable model fit and use of the scale should still provide satisfactory results. While the LEIDS-RR is recommended for future research, the subscales of the five-factor model are maintained, which are similar to those of the LEIDS-R, suggesting that the interpretation of previous studies using the LEIDS-R are not altered by these new results. The scores of the LEIDS-R and the LEIDS-RR both showed good criterion validity, in accordance with similar results previously shown for the original LEIDS (Van der Does, 2002a). This suggests that the LEIDS-RR can be useful in identifying individuals with an increased risk for depression.

LEIDS-RR scores did not discriminate between the previously anxious and depressed groups, but did distinguish the comorbid anxious and depressed group from the remaining groups. This pattern may be expected from a scale measuring vulnerability to depression, since an anxiety disorder is a risk factor for depression (Pine et al., 1998; Wittchen et al., 2000) and comorbidity leads to an increasingly higher risk (Sherbourne & Wells, 1997). The Neuroticism scale had similar discriminative ability and both the LEIDS-RR and Neuroticism correlated higher with depression history than with the ASI. The ASI showed the expected pattern of scores among groups, with the comorbid remitted anxiety/depression group scoring highest.

The HOP and AVC subscales had the greatest ability to discriminate between individuals with and without a past depressive episode and also loaded the strongest on the general cognitive reactivity factor. The CTR subscale also performed well. In general, the AVC subscale correlated highly with the total LEIDS-RR score (r = .91), which suggests that the AVC subscale may also be used as a short version to estimate CR. Future studies should investigate whether higher scores on either the AVC subscale, HOP subscale, or CTR subscale alone, can predict depression incidence and relapse using a longitudinal design. Future research investigating a short version of the LEIDS-RR should also further examine three items (5 *"feel hopeless about everything"*, 19 "*work harder*" and 32 "*think about the causes of my moods*") since they had the tendency to cross-load when examining the five-factor model in both samples. Although these items were inspected during the EFA process and appeared to fit conceptually into the questionnaire and their corresponding subscale, further research is necessary to confirm that these items should remain in the questionnaire.

A limitation of the current study is that the screening procedure for the student sample may have overestimated the number of participants with a depressive disorder and may have thus excluded too many participants. Considering the young age (M = 20.95, SD = 3.62) of the participants and the ability to study at a university level, we would not expect too many participants in the student sample to have psychiatric disorders, making this sample less of a clinical sample than the NESDA sample. This conservative participant selection and resulting limited variability in cognitive reactivity may explain the relatively poorer model fit in the student sample compared to the NESDA sample. Further cross-validation is necessary to confirm the five-factor bifactor model.

Although participants did not frequently endorse the ACC items, the items did result in a homogeneous factor in the analyses. These items relate to increased interpersonal sensitivity, creativity and acceptance when one experiences sad mood. Although not commonly endorsed, the subscale might be related to cognitive reactivity in individuals with bipolar (disorder) tendencies, who have been previously shown for example to have enhanced creativity compared to controls (Santosa, Strong, Nowakowska, Wang, Rennicke & Ketter, 2007). Bipolar and MDD patients, as well as creative controls, have been shown to share common temperamental traits (Nowakowska, Strong, Santosa, Wang, & Ketter, 2005). On the other hand, statements from the ACC subscale could characterize individuals with chronic sad mood, such as highly recurrent depression or dysthymia; in such cases, a depressive identity may become internalized leading to a sense of acceptance. Future research may investigate the pertinence and utility of the ACC subscale. Future research should also investigate scoring norms to establish specificity and sensitivity cut-off scores for the LEIDS-RR to determine at which scores individuals are at a higher risk for depression.

Conclusions

In conclusion, the LEIDS-R showed satisfactory psychometric properties, but the revised 30-item LEIDS-RR was superior to the LEIDS-R structure and is recommended for future research. A general cognitive reactivity factor is recommended, although the subscales may provide additional multidimensional information. The HOP and AVC subscales have the strongest association with depression history. The AVC subscale has the highest correlation with the total score and may be a suitable short form.

References

Andrews, G., & Peters, L. (1998). The psychometric properties of the Composite International Diagnostic Interview. *Social Psychiatry and Psychiatric Epidemiology.*

Anisi, J., Majdian, M., Joshanloo, M., & Gohari-kamel, Z. (2012). Validity and reliability of the NEO Five-Factor Inventory (NEO-FFI) on university students. *Journal of Behavioral Sciences, 5,* 351-355.

Antypa, N., Van der Does, A.J.W., & Penninx, B.W.J.H. (2010). Cognitive reactivity: Investigation of a potentially treatable marker of suicide risk in depression. *Journal of Affective Disorders, 122,* 46-52.

Antypa, N., & Van der Does, A.J.W. (2010). Serotonin transporter gene, childhood emotional abuse and cognitive vulnerability to depression. *Genes, Brain and Behavior, 9,* 615-620.

Beck, A.T. (1967). *Depression: Clinical, experimental, and theoretical aspects.* New York: Harper & Row.

Bentler, P. (1990). Comparative fit indices in structural models. *Psychological Bulletin, 107,* 238-246.

Bentler, P. M. (1995). *EQS structural equations program manual.* Encino, CA: Multivariate Software.

Bentler, P. M. (2009). Alpha, dimension-free, and model-based internal consistency reliability. *Psychometrika, 74*, 137–143.

Berrios, R., Kellett, S., Fiorani, C., & Poggioli, M. (2015). Assessment of Identity Disturbance: Factor Structure and Validation of the Personality Structure Questionnaire in an Italian Sample. *Psychological Assessment.* Advance online publication. http://dx.doi.org/10.1037/pas0000143

Bjelland, I., Dahl, A.A., Haug, T.T., & Neckelmann, D. (2002). The validity of the Hospital Anxiety and Depression Scale: an updated literature review. *Journal of Psychosomatic Research, 52,* 69-77.

Blazer, D.G., Kessler, R.C., McGonagle, K.A., Swartz, M.S. (1994). The prevalence and distribution of major depression in a national community sample: The National Comorbidity Survey. *American Journal of Psychiatry, 151*, 979-986.

Booij, L., & Van der Does, A.J.W. (2007). Cognitive and serotonergic vulnerability to depression: Convergent findings. *Journal of Abnormal Psychology, 116,* 86-94.

Brosse, A.L., Craighead, L.W., & Craighead, W.E. (1999). Testing the mood-state hypothesis among previously depressed and never-depressed individuals. *Behavior Therapy, 30,* 97-115.

Browne, M.W, & Cudeck, R. (1993). Alternative ways of assessing model fit. In K.A. Bollen & J.S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Newbury Park, CA: Sage.

Brunner, M., Nagy, G., & Wilhelm, O. (2012). A tutorial on hierarchically structured constructs. Journal of Personality, 80, 796–846.

Chen, F.F. (2007). Sensitivity of goodness of fit indices to lack of measurement invariance. *Structural Equation Modeling, 14*, 464–504.

Cheung, G.W., & Rensvold, R.B. (2002). Evaluating goodness-of-fit indices for testing measurement invariance. *Structural Equation Modeling, 9,* 233–255.

Clark, D.A., & Beck, A.T. (1999). *Scientific Foundations of Cognitive Theory and Therapy of Depression.* New York: John Wiley.

Clark, L.A., & Watson, D. (1991). Tripartite model of anxiety and depression: psychometric evidence and taxonomic implications. *Journal of Abnormal Psychology, 100,* 316-336.

Cortina, J.M. (1993). What is coefficient alpha? An examination of theory and application*.* *Journal of Applied Psychology*, *78*(1), 98-104.

Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual.* Odessa, FL: Psychological Assessment Resources.

Costello, A.B., & Osbourne, J.W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research, and Evaluation, 10*(7). Retrieved on October 26, 2015, from http://pareonline.net/getvn.asp?v=10&n=7
Cox, B.J., Parker, J.D.A., & Swinson, R.P. (1996). Anxiety sensitivity: Confirmatory evidence for a multidimensional construct. *Behaviour Research and Therapy, 34*(7), 591-598.
Cribb, G., Moulds, M.L., & Carter, S. (2006). Rumination and experiential avoidance in depression. *Behaviour Change, 23,* 165-176.
Cronkite, R. C., & Moos, R. H. (1995). Life context, coping processes, and depression. In E. E. Beckham, & W. R. Leber (Eds.), *Handbook of depression* (pp. 569–587). New York: Guilford.
Glorfeld, L.W. (1995). An improvement on Horn's parallel analysis methodology for selecting the correct number of factors to retain. *Educational and Psychological Measurement, 55,* 377-393.
Gomez, R., & McLaren, S. (2015). The Center for Epidemiological Studies Depression Scale: Measurement and structural invariance across ratings of older adult men and women. *Personality and Individual Differences, 75,* 130-134.
Graham, J.M. (2006). Congeneric and (essentially) tau-equivalent estimates of score reliability what they are and how to use them. *Educational and Psychological Measurement*, *66*(6), 930-944.
Hankin, B.L., Oppenheimer, C., Jenness, J., Barrocas, A., Shapero, B.G., & Goldband, J. (2009). Developmental origins of cognitive vulnerabilities to depression: Review of processes contributing to stability and change across time. *Journal of Clinical Psychology, 65*(12), 1327-1338.
Hardeveld, F., Spijker, J., De Graaf, R., Nolen, W.A., & Beekman, A.T.F. (2013). Recurrence of major depressive disorder and its predictors in the general population: results from The Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Psychological Medicine, 43,* 39-48.
Hoekstra, H.A., Ormel J., & de Fruyt, F. (1996). *Handleiding NEO persoonlijkheids-vragenlijsten NEO-PI-R en NEO-FFI.* Lisse: Swets Test Services.
Holahan, C.J., Moos, R.H., Holahan, C.K., Brennan, P.L., & Schutte, K.K. (2005). Stress generation, avoidance coping, and depressive symptoms: A 10-year model. *Journal of Consulsting and Clinical Psychology, 73,* 658-666. doi: 10.1037/0022-0006X.73.4.658
Hollon, S. D., DeRubeis, R. J., & Evans, M. D. (1987). Causal mediation of change in treatment for depression: Discriminating between nonspecificity and noncausality. *Psychological Bulletin*, *102*, 139–149.
Hedlund, S., & Rude, S. (1995). Evidence of latent depressive schemas in formerly depressed individuals. *Journal of Abnormal Psychology, 95,* 53-59.
Horn, J.L. (1965). A rationale and test for the number of factor in factor analysis. *Psychometrika, 30,* 179-185.
Hu, L., & Bentler, P. (1999). Cutoff criteria for fit indices in covariance structure analysis. Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1-55.
Ingram, R.E., & Ritter, J. (2000). Vulnerability to depression: Cognitive reactivity and parental bonding in high-risk individuals. *Journal of Abnormal Psychology, 109,* 588-596.
Jarrett, R.B., Minhajuddin, A., Borman, P.D., Dunlap, L., Segal, Z.V., Kidner, C.L., Friedman, E.S., & Thase, M.E. (2012). Cognitive reactivity, dysfunctional attitudes, and depressive relapse and recurrence in cognitive therapy responders. *Behaviour Research and Therapy, 50,* 280-286.
Kim, H.Y. (2013). Statistical notes for clinical researchers: assessing normal distributions (2) using skewness and kurtosis. *Restorative Dentistry & Endodontics, 38*; 52-54.
Kruijt, A.W., Antypa, N., Booij, L., de Jong, P.J., Glashouwer, K., Penninx, B.W.J.H., & Van der Does, W. (2013). Cognitive reactivity, implicit associations, and the incidence

of depression: a two-year prospective study. *PloS ONE*, *8*, e70245. doi: 10.1371/journal.pone.0070245

Lin, A., Yung, A.R., Wigman, J.T.W., Killacky, E., Baksheev, G., & Wardenaar, K.J. (2014). Validation of a short adaptation of the Mood and Anxiety Symptoms Questionnaire (MASQ) in adolescents and young adults. *Psychiatry Research, 215,* 778-783.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1,* 130-149.

Martin, M. (1990). On the induction of mood. *Clinical Psychology Review, 10,* 669-697.

McDonald, R.P. (1970). The theoretical foundation of common factor analysis, principal factor analysis, and alpha factor analysis. *British Journal of Mathematical and Statistical Psychology, 23,* 1-21.

McDonald, R.P. (1999). *Test theory: A unified treatment.* Mahwah, NJ: Lawrence Erlbaum.

Merens, W., Booij, L., & Van der Does, A.J.W. (2008). Residual Cognitive Impairments in Remitted Depressed Patients. *Depression and Anxiety, 25,* E27- E36.

Millsap, R.E., & Yun-Tein, J. (2004). Assessing factorial invariance in ordered-categorical measures. *Multivariate Behavioral Research, 39*(3), 479-515.

Miranda, J., Gross, J.J., Persons, J.B., & Hahn, J. (1998). Mood matters: Negative mood induction activates dysfunctional attitudes in women vulnerable to depression. *Cognitive Therapy and Research, 22,* 363-376.

Miranda, J., & Persons, J.B. (1988). Dysfunctional attitudes are mood-state dependent. *Journal of Abnormal Psychology, 97,* 76-79.

Miranda, J., Persons, J.B., & Byers, C. (1990). Endorsement of dysfunctional beliefs depends on current mood state. *Journal of Abnormal Psychology, 99,* 237-241.

Mohlman, J., & Zinbarg, R. (2000). The structure and correlates of anxiety sensitivity in older adults. *Psychological Assessment, 12,* 440-446.

Moulds, M.L., Kandris, E., Starr, S., & Wong, A.C.M. (2007). The relationship between rumination, avoidance, and depression in a non-clinical sample. *Behaviour Research and Therapy, 45,* 251-261.

Moulds, M.L., Kandris, E., Williams, A.D., Lang. T., Yap, C., & Hoffmeister, K. (2008). An investigation of the relationship between cognitive reactivity and rumination. *Behavior Therapy, 39,* 65-71.

Mueller, T.I., & Leon, A.C. (1999). Recurrence after recovery from major depressive disorder during 15 years of observational follow-up. *The American Journal of Psychiatry, 156,* 1000-1006.

Muthén, L. K., & Muthén, B. O. (2009). Mplus Short Course Videos and Handouts. Muthén & Muthén – Home Page. Retrieved from http:/www.statmodel.com/download/Topic 1-v11.pdf.

Muthén, L.K., & Muthén, B.O. (1998-2012). Mplus User's Guide. Seventh Edition. Los Angeles, CA: Muthén & Muthén.

Nolan-Hoeksema, S. (1991). Responses to depression and their effects on the duration of depressive episodes. *Journal of Abnormal Psychology, 100*(4), 569-582. http://dx.doi.org/10.1037/0021-843X.100.4.569

Nowakowska, C., Strong, C.M., Santosa, C.M., Wang, P.W., & Ketter, T.A. (2005). Temperamental commonalities and differences in euthymic mood disorder patients, creative controls, and healthy controls. *Journal of Affective Disorders, 85*, 207-215.

Olatunji1, B.O., Ebesutani, C., & Reise, S.P. (2015). A bifactor model of disgust proneness: Examination of the disgust emotion scale. *Assessment, 22*(2), 248-262.

Olsson, I., Mykletun, A., & Dahl, A.A. (2005). The hospital anxiety and depression rating scale: A cross-sectional study of psychometrics and case finding abilities in general practice. *BMC Psychiatry, 5(46)*. doi: 10.1186/1471-244X-5-46.

Penley, J.A., Tomaka, J., & Wiebe, J.S. (2002). The association of coping to physical and physiological health outcomes: A meta-analytic review. *Journal of Behavioural Medicine, 6,* 551-603.

Penninx, B.W.J.H., Beekman, A.T.F., Smit, J.H., Zitman, F.G., Nolen, W.A., Spinhoven, P., et al. (2008). The Netherlands Study of Depression and Anxiety (NESDA): Rationale, objectives, and methods. *International Journal of Methods in Psychiatric Research, 17,* 121-140.
Peterson, R. A., & Reiss, S. (1992). *Anxiety sensitivity index manual.* (2nd ed.). Worthington, OH: International Diagnostic Systems.
Pine, D.S., Cohen, P., Gurley, D., Brook, J. & Ma, Y. (1998). The risk for early-adulthood anxiety and depressive disorders in adolescents with anxiety and depressive disorders. *Archives of General Psychiatry, 55*, 56- 64.
Raes, F., Dewulff, D., Van Heeringen, C., & Williams, J.M.G. (2009). Mindfulness and reduced cognitive reactivity to sad mood: Evidence from a correlational study and a non-randomized waiting list controlled study. *Behaviour Research and Therapy, 47*, 623-627.
Raykov, T. (2004). Behavioural scale reliability and measurement invariance evaluation using latent variable modeling. *Behavior Therapy, 35,* 299-331.
Reise, S.P. (2012). The rediscovery of bi-factor measurement models. Multivariate Behavioral Research, 47, 667-696
Reise, S.P., Bonifay, W.E., & Haviland, M.G. (2013). Scoring and modeling psychological measures in the presence of multidimensionality. *Journal of Personality Assessment, 95*, 129–140.
Reise, S.P., Moore, T.M., & Haviland, M.G. (2010). Bifactor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. *Journal of Personality Assessment, 92*, 544-559. doi: 10.1080/00223891.2010.496477
Reiss, S., Peterson, R.A., Gursky, D.M., & McNally, R.J. (1986). Anxiety sensitivity, anxiety frequency, and the prediction of fearfulness. *Behaviour Research and Therapy, 24,* 1-8.
Revelle, W., & Zinbarg, R. E. (2009). Coefficients alpha, beta, omega, and the GLB: Comments on Sijtsma. *Psychometrika, 74*, 145-154
Richards, D. (2011). Prevalence and clinical course of depression: A review. *Clinical Psychology Review, 31,* 1117-1125.
Rijsbergen, G.D., Kok, G.D., Elgersma, H.J., Hollon, S.D., & Bockting, C.L.H. (2015). Personality and cognitive vulnerability in remitted recurrently depressed patients. *Journal of Affective Disorders, 173,* 97-104.
Robins, R.W., Fraley, R.C., Roberts, B.W, & Trzesniewski, K.H. (2001). A longitudinal study of personality change in young adulthood. *Journal of Personality, 69,* 617-640.
Robins, L.N., Wing. J., Wittchen, H.U., Helzer, J.E., Babor, F.F., Burke, J., …Towle, L.H. (1988). The Composite International Diagnostic Interview: An epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. *Archives of General Psychiatry, 45,* 1069-1077.
Rodriguez, A., Reise, S. P., & Haviland, M. G. (2015). Evaluating Bifactor Models: Calculating and Interpreting Statistical Indices. Psychological Methods. Advance online publication. http://dx.doi.org/10.1037/met0000045
Romera, B., Vilagut, G., Puigvert, J.M., Palacín, C., Ferrer, M., Haro, J.M., & Alsonso, J. (2002). Reliability and validity of the Composite International Diagnostic Interview (CIDI). *Quality of Life Research, 11,* Abstracts: 9th Annual Conference of the International Society for Quality of Life Research (ISOQOL) (Nov., 2002), p. 641.
Rush, A.J., Weissenburger, J., & Eaves, G. (1986). Do thinking patterns predict depressive symptoms? *Cognitive Therapy and Research, 10,* 225-235.
Santosa, C.M., Strong, C.M., Nowakowska, C., Wang, P.W., Rennicke C.M., Ketter, T.A. (2007). Enhanced creativity in bipolar disorder patients: a controlled study. *Journal of Affective Disorders, 100,* 31-39.
Schmitt, T.A. (2011). Current methodological considerations in exploratory and confirmatory factor analysis. *Journal of Psychoeducational Assessment, 29,* 304. doi: 10.1177/0734282911406653

Schmidt, N. B., & Joiner, T. E. (2002). Structure of the Anxiety Sensitivity Index psychometrics and factor structure in a community sample. *Journal of Anxiety Disorders, 16*(1), 33-49.
Segal, Z.V., Gemar, Z., & Williams, S. (1999). Differential cognitive response to a mood challenge following successful cognitive therapy or pharmacotherapy for unipolar depression. *Journal of Abnormal Psychology, 108,* 3-10.
Segal, Z.V., Kennedy, S., Gemar, M., Hood, L., Pedersen, R., & Buis, T. (2006). Cognitive reactivity to sad mood provocation and the prediction of depressive relapse. *Archives of General Psychiatry, 63,* 749-755.
Sherbourne, C.D. & Wells, K.B. (1997). Course of depression in patients with comorbid anxiety disorders. *Journal of Affective Disorders, 43,* 245-250.
Spinhoven, P., Penninx, B.W., Hickendorff, M., van Hemert, A.M., Bernstein, D.P., & Elzinga, B.M. (2014). Childhood Trauma Questionnaire: Factor structure, measurement invariance, and validity across emotional disorders. *Psychological Assessment, 26*(3), 717-729.
Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research, 25,* 173–180.
Taylor, S., Koch, W.J., McNally, R.J., & Crockett, D.J. (1992). Conceptualizations of anxiety sensitivity. *Psychological Assessment, 4,* 245-250.
Taylor, L., & Ingram, R.E. (1999). Cognitive reactivity and depressotypic information processing in children of depressed mothers. *Journal of Abnormal Psychology, 108,* 202-210.
Teasdale, J.D. (1988). Cognitive vulnerability to persistent depression. *Cognition and Emotion, 2,* 247-274.
Ten Berge, J. M., & Sočan, G. (2004). The greatest lower bound to the reliability of a test and the hypothesis of unidimensionality. *Psychometrika, 69*, 613–625. http://dx.doi.org/10.1007/BF02289858
Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications.* Washington, DC: American Psychological Association.
Van der Does, W. (2002a). Cognitive reactivity to sad mood: Structure and validity of a new measure. *Behaviour Research and Therapy, 40,* 105-120.
Van der Does, W. (2002b). Different types of experimentally induced sad mood? *Behavior Therapy, 33,* 551-561.
Van der Does, W. (2005). Thought suppression and cognitive vulnerability to depression. *British Journal of Clinical Psychology, 44,* 1-14.
Van der Does, A.J.W., & Williams, J.M.G. (2003). Leiden Index of Depression Sensitivity-Revised (LEIDS-R). Retrieved June 29, 2015, from http://www.dousa.nl/publications.htm#LEIDS.
Van der Does, A.J.W., Manthey, L., & Hermans, K. (2012). Dysphoric mood, thought suppression, and cognitive control in recovered depressed and never-depressed individuals. *Journal of Experimental Psychopathology, 3,* 69-83.
Van der Sluis, S., Vinkhuyzen, A.A.E., Boomsma, D.I., and Posthuma, D. (2010). Sex differences in adults' motivation to achieve. *Intelligence, 38,* 433-446.
Verhoeven, F.E., Booij, L., Van der Wee, N.J.A., Penninx, B.W.H.J., & Van der Does, A.J.W. (2011). Clinical and physiological correlates of irritability in depression: Results from the Netherlands Study of Depression and Anxiety. *Depression Research and Treatment, 2011.* doi:10.1155/2011/126895
Verhoeven, F.E., Booij, L., Kruijt, A.W., Antypa, N., & van der Does, W. (2012). The effects of MAOA genotype, childhood trauma, and sex on trait and state-dependent aggression. *Brain and Behavior, 2,* 806-813.
Vujanovic, A.A., Zvolensky, M.J., Bernstein, A., Feldner, M.T., & McLeish, A.C. (2007). A test of the interactive effects of anxiety sensitivity and mindfulness in the prediction of anxious arousal, agoraphobic cognitions, and body vigilance. *Behaviour Research and Therapy, 45*(6), 1393-1400.
Wardenaar, K.J., van Veen, T., Giltay, E.J., de Beurs, E., Penninx, B.W., & Zitman, F.G. (2010). Development and validation of a 30-item short adapation of the Mood

and Anxiety Symptoms Questionnaire (MASQ). *Psychiatric Research, 179,* 101-106.
Wardenaar, K.J., van Veen, T., Giltay, E.J., den Hollander-Gijsman, M., Penninx, B.W.J.H., & Zitman, F.G. (2010). The structure and dimensionality of the Inventory of Depressive Symptomatology Self Report (IDS-SR) in patients with depressive disorders and healthy controls. *Journal of Affective Disorders 125,* 146-154.
Warne , R.T, & Larsen, R. (2014). Evaluating a proposed modification of the Guttman rule for determining the number of factors in an exploratory factor analysis. *Psychological Test and Assessment Modeling 56,* 104-123.
Weissman, A. (1979). *Dysfunctional Attitude Scale: A validation study.* Unpublished doctoral dissertation, University of Pennsylvania, Philadelphia.
Widhiarso, W., & Ravand, H. (2014). Reliability coefficient for multidimensional measures. *Review of Psychology, 21*(2), 111-121.
Williams, J.M.G., Van der Does, A.J.W., Barnhofer, T., Crane, C., & Segal, Z.S. (2008). Cognitive reactivity, suicidal ideation, and future fluency: Preliminary investigation of a differential activation theory of hopelessness/suicidality. *Cognitive Therapy Research, 32,* 83-104.
Wittchen, H.U. (1994). Reliability and validity studies of the WHO-composite international diagnostic interview (CIDI): a critical review. *Journal of Psychiatric Research, 281,* 57-84.
Wittchen, H.U., Robins, L.N., Cottier, L.B., Sattorius, N., Burke, J.D., & Regier, D. (1991). Cross-cultural feasibility, reliability, and sources of variance of the composite international diagnostic interview (CIDI). *The multicentre WHO/ADAMHA field trials. British Journal of Psychiatry, 159,* 645-653, 658.
Wittchen, H.U., Kessler, R.C., Pfister, H. & Lieb, R. (2000). Why do people with anxiety disorders become depressed? A prospective-longitudinal community study. *Acta Psychiatrica Scandinavica, 102* (suppl 406), 14-23.
World Health Organization [WHO] (1993). *The ICD-10 classification of mental and behavioural disorders. Diagnostic criteria for research.* Geneva: World Health Organization.
Zigmond, A. S., & Snaith, R.P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica, 67,* 361-370.
Zinbarg, R.E, & Barlow, D.H. (1996). Structure of anxiety and the anxiety disorders: a hierarchical model. *Journal of Abnormal Psychology, 105,* 181-193.
Zinbarg, R.E., Barlow, D.H., & Brown, T.A. (1997). Hierarchical structure and general factor saturation of the anxiety sensitivity index: Evidence and implications. *Psychological Assessment, 9*(3), 277-284.
Zinbarg, R. E., Revelle, W., Yovel, I., & Li, W. (2005). Cronbach's, Revelle's, and McDonald's: Their relations with each other and two alternative conceptualizations of reliability. *Psychometrika, 70,* 123-133.

Supplementary Material

Table A. *Standardized Factor Loadings and Standardized Cronbach's Alpha Scores (α) of the Six-Factor LEIDS-R Model from Confirmatory Factor Analysis using the Total NESDA Sample.*

LEIDS-R Item Statement (Item No.)	Factor					
	ACC	AGG	HOP	CTR	RAV	RUM
More creative than usual (4)	.51					
More helpful (10)	.66					
Better intuitive feeling for meaning (15)	.70					
Feel more like myself (24)	.76					
Nicer than usual (28)	.68					
Do more things I will regret (7)		.72				
Feel like breaking things (18)		.75				
Bothered by aggressive thoughts (21)		.81				
Feel cynical or sarcastic (22)		.73				
Take more risks (26)		.73				
Lose my temper easily (29)		.70				
Feel hopeless about everything (5)			.87			
Care less if I lived or died (9)			.79			
Think I can make nobody happy (17)			.82			
Feel I would be better off dead (30)			.84			
Think about harming myself (34)			.76			
Think about what mood says about me (3)				.73		
Go out and do pleasurable activities (8)				.02		
Permit myself fewer mistakes (12)				.79		
Bothered by perfectionism (16)				.68		
Work harder (19)				.34		
Higher need for control (31)				.67		
Think positively only in good mood (1)					.67	
Take fewer risks (2)					.72	
Busy keeping thoughts or images at bay (6)					.61	
Less inclined to express disagreement (11)					.64	
Avoid difficulties (14)					.71	
Want to escape everything (23)					.84	
Feel overwhelmed (13)						.83
Less able to cope with everyday tasks (20)						.77
Neglect tasks (25)						.73
Problems concentrating (27)						.76
Think about the causes of mood (32)						.52
Think about different life choices (33)						.67
α	.63	.79	.83	.65	.81	.82

Note. All 34-items were included in the analysis.

Table B. *Inter-Factor Correlation of the Original LEIDS-R Six-Factor Structure and EFA-derived Five-Factor Structure*

		Factor					
CFA Model	Factor	CTR	HOP	ACC	AGG	RUM	RAV
Six-factor	CTR		0.68	.67	.59	.83	.83
34-items[a]	HOP			.38	.67	.79	.78
	ACC				.36	.40	.50
	AGG					.71	.59
	RUM						.93
Five-factor		CTR	HOP	ACC	AGG	AVC[c]	
30-items[b]	CTR		.69	.66	.57	.84	
	HOP	.59		.40	.67	.80	
	ACC	.47	.20		.40	.40	
	AGG	.30	.52	.19		.67	
	AVC	.65	.68	.16	.50		

[a]NESDA sample, N = 1696.
[b]Items 1, 6, 8, and 33 were removed from the analyses. NESDA Sample 2 (1), N = 840, above diagonal; Student Sample (2), N = 811, below diagonal.
[c]Factors RUM and RAV of Six-Factor Model combined to form AVC factor.

All correlations were statistically significant, $p < .001$.

Table C. *Cronbach's Alpha Estimates (α) and Spearman Correlations between the 30-Item LEIDS-RR Scores, CIDI-WHO Lifetime Depression Status, ASI Scores, and NEO-FFI Neuroticism Scores.*

LEIDS-RR Scores	α	Correlations: LEIDS-RR Scores ACC	AVC	CON	AGG	HOP	Total	MDD Lifetime	ASI	NEO-FFI Neuroticism
ACC	.67							.13	.24	.25
AVC	.88	.29						.32	.50	.58
CON	.76	.42	.68					.25	.43	.46
AGG	.80	.25	.55	.48				.20	.34	.40
HOP	.83	.27	.71	.59	.54			.30	.47	.62
Total	.91	.47	.91	.83	.70	.82		.33	.53	.62

Note. All correlations were conducted using Spearman's rho and significant, $p < .05$. Results were obtained in the total NESDA sample ($N = 1{,}696$). ASI = Anxiety Sensitivity Inventory. NEO-FFM = Revised NEO Five Factor Inventory.