Psychiatry Research 293 (2020) 113410

Contents lists available at ScienceDirect

Psychiatry Research

journal homepage: www.elsevier.com/locate/psychres

Carving depression at its joints? Psychometric properties of the Sydney Melancholia Prototype Index

Lorenzo Lorenzo-Luaces[a,1,*], Lauren A. Rutter[a,1], Matthew D. Scalco[b]

[a] Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN, United States
[b] Department of Psychology, The University of New Orleans, New Orleans, LA, United States

ARTICLE INFO

Keywords:
Melancholia
Depression
Classification
Structural equation modeling

ABSTRACT

Parker and colleagues developed the Sydney Melancholia Prototype Index (SMPI), a 24-item measure to assess a potential subtype of depression: melancholia. While research supports the validity of the measure, no study has assessed its psychometric properties. We recruited 1633 participants online, of whom 487 reported a lifetime period of depressed mood or anhedonia and were administered the SMPI. We conducted confirmatory factor analyses (CFA) of the SMPI, to assess the proposed fit of the measure. We also conducted exploratory factor analyses (EFA) to explore the structure implied by the current data. CFA did not support the hypothesized factor structure of the SMPI, no matter what structure we assumed as primary (i.e., a one factor, two factor, or bifactor model). An EFA suggested a five-factor solution wherein several items did not appear to co-vary reliably and other factors captured the severity of melancholic symptoms, negative mood reactivity, positive mood reactivity, emotionality and family relationships, and early life adversity. The SMPI may not measure a single construct. Future research should explore the longitudinal association between depression severity, contaminant symptoms, positive and negative mood reactivity, and early life experiences.

1. Introduction

Experiences of depressed mood or low positive emotions can range from states of momentary sadness that anyone would consider non-disordered, to highly debilitating, chronic, and recurrent clinical conditions (Lorenzo-Luaces, 2015). The most common of the depressive disorders, major depressive disorder (MDD), is diagnosed if an individual endorses five symptoms, one of which must be depressed mood or anhedonia (hereafter both referred to as depressed mood) for at least two weeks (American Psychiatric Association APA, 2013). The optimal classification of MDD has been one of the major challenges in the history of psychiatry (Fried and Nesse, 2015; Lorenzo-Luaces, 2015; Parker, 2005). Questions include both the optimal differentiation of MDD from "normal" depressed mood (Wakefield and Schmitz, 2013, Wakefield and Schmitz, 2013; Wakefield and Schmitz, 2013; Wakefield et al., 2007) as well as identification of the various pathological states (e.g., subtypes, endophenotypes) likely subsumed under MDD. These challenges are made all the more pressing by the fact that depression is a leading cause of disability worldwide as well as arguably the most widely researched psychiatric diagnosis (Murray et al., 2012).

The longest-standing distinction in the history of thinking on depression is the differentiation between milder forms of depression, usually assumed to be psychogenic or triggered by a negative event, and depression without a cause, "out of the blue," or untriggered (Horwitz and Wakefield, 2007; Shorter, 2007). The specific terms used to reference these depressions have varied but include on one end non-melancholic, exogenous, or reactive depression and, on the other, melancholic, endogenous, or psychotic depression. Shorter (2007) put it succinctly when he said that "the concept of major depression popularized in DSM-III in 1980 is a historical anomaly," because it mixes different forms of pathology in a single diagnosis.

In addition to differing by how responsive depression is to external environmental influence, researchers and clinicians have also prioritized specific symptoms for subtyping melancholia, with psychomotor disturbances (PMDs), usually psychomotor retardation, emerging as a symptom consistently linked to melancholia (Parker, 2007). Parker and colleagues have gone so far as to call PMD the chief characteristic of melancholia, which they highlight as a specific disorder of mood and motor functioning (Malhi et al., 2005). They propose a tiered model of depression pathology wherein individuals with non-melancholic depression are a heterogeneous group that encompasses many individuals who have personality or cognitive vulnerabilities and are reacting

* Corresponding author at: 1101 E 10th St, Bloomington, IN 47405, United States.
E-mail addresses: lorenzl@sas.upenn.edu, lolorenz@iu.edu (L. Lorenzo-Luaces).
[1] Authors contributed equally to the manuscript

https://doi.org/10.1016/j.psychres.2020.113410
Received 5 March 2020; Received in revised form 5 August 2020; Accepted 17 August 2020
Available online 18 August 2020
0165-1781/

Table 1
Items in the Sydney Melancholia Prototype Index (SMPI).

Scale A
1. I have very low energy and find it extremely hard to get out of bed and get going.
2. My depressed mood completely prevents me from getting any real pleasure in life, and normally pleasing or humorous things won't lift my mood- or, at best, only superficially.
3. My mood and energy levels are worse in the mornings.
4. I completely lose interest in things, including hobbies and activities that I would usually enjoy when not depressed.
5. I find that I can't look forward to anything in life.
6. In walking and talking, I'm distinctly physically slowed, at times almost feeling 'paralyzed' or as if I'm walking through sand.
7. My concentration is distinctly affected and slowed.
8. I tend to lose weight when I'm depressed (and before any antidepressant or other drugs are commenced).
9. The severity of my depressive episodes appears far worse than would be expected given the circumstances that may precede them or appear to cause them.
10. I don't think that my early years were any more difficult - when compared to most people - in terms of having any major difficulties with parents or bullying.
11. When I'm not depressed my relationships and work performance are generally good.
12. My depressions can sometimes come 'out of the blue' without any particularly clear reason.

Scale B
1. Even when my depression is severe, I can generally look forward to something really nice coming up.
2. I find that I become distinctly more irritable and/or angry when I'm depressed.
3. Even when my depression is severe, I can generally be cheered up when people are really supportive.
4. My mood lifts (even if temporarily) and I can obtain some temporary relief when something nice happens.
5. If my concentration is affected during a depressive episode, it is usually because I am worrying too much and have lots of thoughts going through my head distracting me.
6. I often get (non-medication related) food cravings and/or increased appetite when I'm depressed.
7. I view myself as generally more inclined than most people to become emotional about things (regardless of whether I'm depressed or not).
8. Every time I get depressed, I can find some cause that explains the depression to me.
9. The severity of my depressions can be explained by the type of stressful events that precede them and the impact that these events have on me given my type of personality.
10. Even when I'm not depressed, I tend to have some difficulties in dealing with my partner, family and other relationships.
11. Even when I'm not depressed, I tend to worry more than most people, particularly when under stress.
12. In childhood and adolescence, I experienced more stressful events and major difficulties with my parents and others than most people experience.

Note. We allowed respondents to score items on a 5-point scale: 1 = "Completely disagree"; 2 = "Somewhat disagree"; 3 = "Neutral"; 4 = "Somewhat agree"; 5 = "Completely agree"

intensely enough to life stressors to meet MDD criteria (Malhi et al., 2005; Parker, 2005; Parker, 2007).

1.1. Sydney Melancholia Prototype Index (SMPI)

Parker and colleagues developed a self-reported and clinician-rated measure, first dubbed the Self-Report of Depressive Experiences (SERDEX), later renamed the Sydney Melancholia Prototypical Index (SMPI), to identify melancholia based on prototypical features (Parker et al., 2010; Parker et al., 2012; Parker et al., 2013; Parker et al., 2013; Parker et al., 2015; Parker et al., 2019). The SMPI includes 24 items that are clustered into two sets: the A items, worded to accord to the melancholic presentation, and the B items, worded to accord to the non-melancholic presentation. The questions on the SMPI query PMD, anhedonic symptoms, contextual triggers of depressive episodes, specific symptoms like food cravings, temperament, as well as pre-morbid personality and interepisode functioning (see Table 1).

The results of the published tests of the SMPI's validity suggest a number of differences in distinguishing melancholic and non-melancholic depression (Parker et al., 2010; Parker et al., 2012; Parker et al., 2013; Parker et al., 2013; Parker et al., 2015; Parker et al., 2019). For example, in one study, compared to participants with non-melancholic depression, participants with melancholic depression were more likely to report that their depressions came "out of the blue" (48 vs. 78%; Parker et al., 2012). Parker and colleagues have also compared response patterns between unipolar melancholic, unipolar non-melancholic, and bipolar depressions (Parker et al., 2013). They generally found that the pattern of responses for the melancholic patients differed markedly from the non-melancholic patients, who in turn provided responses similar to patients with bipolar disorder. SMPI-defined melancholia also appears to produce differences in clinically relevant variables including: personality, anxiety comorbidity, and substance use disorder comorbidity (Parker et al., 2010; Parker et al., 2012; Parker et al., 2013; Parker et al., 2013; Parker et al., 2015; Parker et al., 2019).

While the SMPI queries various features of melancholia, a recent data-driven analysis by Parker et al. (2015) suggests that items on the scale are differentially predictive of melancholic status. Specifically, a decision tree combining the anergia symptom and items querying the degree to which depression occurs as a response to the environment (e.g., "my depression occurs out of the blue"), can predict whether individuals are diagnosed as melancholic or non-melancholic (kappa = 0.74). Thus, responses to the SMPI, specifically the questions querying reactions to the environment, possibly identify a meaningfully different subset of individuals with depression.

1.2. Limitations of the SMPI

Despite support for the validity of the SMPI, there are a number of limitations to the measure and its research base. To our knowledge, no psychometric data on the structure and reliability of the SMPI have been reported. As a result, there is no way of knowing how the items are related to one another and whether they measure one construct or many. Although a clear structure may be implied from the nature of the scale and the way it is used, the only data-driven approach to using the scale suggests that the items capturing (negative) mood reactivity are the ones primarily distinguishing melancholic and non-melancholic depression. Moreover, with one recent exception (Parker et al., 2019) the SMPI has only been administered to clinical populations, usually those from tertiary clinics. It is well-known that clinical populations are only representative of a small portion of patients who meet for MDD (i.e., those with more severe illnesses and who have a perceived need for help; Lorenzo-Luaces, 2015). To our knowledge, the SMPI has mainly been used in Australian samples, and it has not been tested in non-Australian or representative samples. Additionally, the SMPI studies have accepted binary responses though some of the questions clearly are best suited for a continuum (e.g., "I have very low energy and find it extremely hard to get out of bed and going"). Indeed, dimensional models have repeatedly shown to be more representative of psychopathology compared to categorical models (i.e., Markon et al., 2011).

To address the limitations in the research base on the SMPI, we administered it to a large sample of individuals who had not been selected based on their clinical status. We also allowed respondents to answer questions based on a 5-point Likert Scale. Using the data from the present study, we tested the factor structure of the SMPI and alternative data-driven structures suggested by exploratory factor analysis.

2. Methods

2.1. Participants

This study was approved by the IRB at the University of New

Orleans. We recruited participants on Mechanical Turk as part of a survey on depression and substance use. Our inclusion criteria consisted of passing reCAPTCHA, passing an attention question that consisted of being able to read basic English, and completing more than just demographic questions. A total of 1891 participants responded to the survey after these initial exclusions. An additional 258 participants were excluded based on their responses to the validity items of the CAT-PD (Simms et al., 2011) for a total of 1633.

Because the SMPI queries episodes of depression, we asked participants whether they ever experienced a period of "sad mood" or if they "lost interest or pleasure …" for about two weeks. Respondents who answered affirmatively to that question were administered the SMPI. This approach captures individuals who would screen positive for a current or past major depressive episode but also includes those who have experienced depressed mood on a lower continuum of severity. A total of 487 participants completed the SMPI, for which we had complete data for 476.

Participants were predominantly male (53%; females = 44%, gender non-binary = 0.37%). The majority of the sample was White (76%; Black/African American = 8.9%, Asian = 6.5%, Biracial = 2.8%, Other = 1.8%, American Indian = 1%, Native Hawaiian = 0.1%) and non-Hispanic (89.8%, Hispanic = 8.5%). The demographics of our sample appeared representative of the U.S. population with participants from all 50 states and racial demographics that mapped on to census data. The average age was 35.74 (SD = 10.50, range = 18-92).

2.2. Measures

2.2.1. Sydney Melancholia Prototype Index (SMPI, Parker et al., 2012)

The current study used the self-report SMPI only (see Supplementary Materials). This 24-item measure includes 12 melancholic and 12 non-melancholic prototypic features. We allowed participants to respond to the prompts on a 1-5 scale with 1 indicating that they "completely disagree" with the characterization of their depressed mood and 5 indicating that they "completely agree."

2.2.2. Patient Health Questionnaire (PHQ-9, Kroenke et al., 2001)

The PHQ-9 is a widely used measure for detecting and screening depression based on the *DSM-5* criteria. Participants respond on the frequency with which they experience symptoms of depression on a 0-3 scale with 0 being "none of the days" and 3 being "all the days." The PHQ-9 has been well-established as a reliable and valid measure of depression severity (Kroenke et al., 2001). Internal consistency in our sample was .91. A score of 5 is considered mild and a 10 has shown to maximize sensitivity and specificity for screening for the diagnosis of depression (Levis et al., 2019).

2.2.3. Generalized Anxiety Disorder Questionnaire (GAD-7; Spitzer et al., 2006)

The GAD-7 is a 7-item self-report scale developed to assess symptoms of GAD. Scores range from 0 to 21, with higher scores indicating more severe GAD. Participants respond on the frequency with which they experience symptoms of anxiety on a 0-3 scale with 0 being "none of the days" and 3 being "all the days." Research has suggested that the GAD-7 is a valid screening tool for GAD in a primary care setting and in the general population (Lowe et al., 2008; Spitzer et al., 2006). Internal consistency in our sample was .92.

2.2.4. Snaith–Hamilton Pleasure Scale (SHAPS; Snaith et al., 1995)

The SHAPS is a 14-item self-report scale to assess hedonic capacity. Scores range from 0-14, with higher scores representing more severe anhedonia. We administered the measure with on a four-point Likert scale with 1 being "strongly agree" and 4 being "strongly disagree," which was dichotomized such that "strongly disagree" and "disagree" were rescored to "1" and "agree" and "strongly agree" were rescored as "0," as is typically done (Franken et al., 2007; Snaith et al., 1995). The SHAPS has been shown to have good psychometric properties in adult outpatients with depression (Nakonezny et al., 2010), and in a variety of clinical and non-clinical populations (Franken et al., 2007). Internal consistency in our sample was .90.

2.3. Analytic plan

All analyses were conducted using the R programming language. First, we conduct confirmatory factor analyses (CFAs) of the SMPI. Although no preliminary data on the factor structure of the SMPI have been provided, the measure is purported to measure melancholia and does so by providing two sets of illness features: melancholic and non-melancholic. Thus, we tested several models. The first simply considered all SMPI items to reflect a single construct (i.e., melancholia). The second model was a two-factor model in which the A items were all allowed to load on to a single construct ("melancholia") and all the B items were allowed to load on to another construct ("non-melancholia"). Additionally, a bifactor model where all the items load on to a general factor (i.e., melancholia) and there are two orthogonal factors capturing residual variance of the "A items" and of the "B" items, not captured by the general factor. Finally, we reran the bifactor model, estimating the correlations of the two group factors (Biderman et al., 2011).

As setting specific cut-offs for assessing "good" model fit cannot be generalized across all models (Hu and Bentler, 1999; Marsh et al., 2004), ranges were used to evaluate model fit (for Root Mean Square Error of Approximation (RMSEA), .08 is poor, .05 - .07 is acceptable, and < .05 is excellent; for Comparative Fit Index (CFI) and Tucker Lewis Index (TLI), < .9 is poor, .9 - .94 is acceptable, and > .95 is excellent; and for Standardized Root Mean Squared Residual (SRMR), .09 is poor, .06 - .09 is acceptable, and < .06 is excellent). When the CFAs did not yield acceptable fit, we conducted an exploratory factor analysis (EFA). We used parallel analysis, eigen values, and the scree plots to determine the optimum number of factors to extract. To establish whether the different factors identified by the EFA were meaningful, we correlated them to the PHQ-9, GAD-7, and SHAPS. Given that we did not select a sample of individuals with diagnosed MDD, we repeated the analyses above for patients who screened positive for MDD on the PHQ-9.

3. Results

The average GAD-7 score in the full sample who completed the GAD-7 in its entirety (n = 1523) was 4.41 (SD = 4.98, range 0 - 21), indicating generalized anxiety symptoms that are mild. Similarly, the average PHQ-9 score for those who completed that measure in its entirety (n = 1503) was 5.13 (SD = 5.78, range 0 -27), showing a mild level of depression. The average SHAPS score in our complete sample (n = 1523) was 2.19 (SD = 2.98, range 0-14), indicating an ability to experience pleasure that is considered normal. Participants completed the SMPI if they screened positive to the prompt ("Have you experienced a sad mood that lasts for longer than 2 weeks or an inability to enjoy the activities you once enjoyed?"), which reduced our sample to n = 487, with completed SMPI cases n = 476. The mean PHQ-9 score for the individuals who indicated a lifetime experience of depressed mood for two weeks was 9.80 (SD = 6.40), which was substantially higher (Cohen's d = 1.30) than individuals who denied ever being depressed for two weeks (M = 2.94, SD = 3.85). Of the participants who reported a lifetime depressed mood for two-weeks, and who completed the SMPI, roughly half screened positive for MDD on the PHQ-9 (i.e., PHQ-9 ≥10; n = 216).

3.1. Confirmatory factor analysis

We conducted confirmatory factor analyses in R using the lavaan package, standardizing latent factors to have a mean of 0 and a variance

of 1. The first model tested a one-factor solution (e.g., "melancholia). This model fit the data poorly (χ^2 (252) = 1922.29, p < .001, RMSEA = 0.12 (90% CI = 0.11, 0.12), SRMR = .11, CFI = .46, TLI = 0.41). Then, we examined a two-factor model whereby all the "A" items loaded on to one Factor 1 (scale "A") and all the "B" items loaded on to a second factor (scale "B"). This model provided slightly improved fit, but still was a poorly-fitted model (χ^2 (251) = 1659.19, p < .001, RMSEA = 0.11 (90% CI = 0.10, 0.11), SRMR = .11, CFI = .55, TLI = 0.50). We also explored a bifactor model in which all items from scale A and scale B loaded on a general factor and items specific to scale A and B loaded on specific factors for the two groups of item. This model did not fit the data well either (χ^2 (228) = 1231.58, p < .001, RMSEA = 0.10 (90% CI = 0.09, 0.10), SRMR = .08, CFI = .68, TLI = 0.61). Next, we estimated another bifactor model, but allowed the specific factors to covary. This was the best-fitting model but still did not have adequate fit (χ^2 (227) = 1229.22, p < .001, RMSEA = 0.10 (90% CI = 0.09, 0.10), SRMR = .08, CFI = .68, TLI = 0.61). Examining modification indices within this last model suggested that fit could be improved freely estimating the error covariances of several items which were mirror versions of each other (i.e., A10-B12, B8-B9, A12-B9, A6-B8). Of the models we tested, this bifactor model with corrections fit the data best, (χ^2 (223) = 715.59, p < .001, RMSEA = 0.07 (90% CI = 0.06, 0.07), SRMR = .06, CFI = .84, TLI = 0.80). Results of factor loadings in the corrected bifactor model are presented in Table 2. Importantly, even the corrected bifactor model did not have acceptable model fit by all indices, as evidenced by low CFI and TLI, though the RMSEA and SRMR values were improved. Taken together, these data suggest that the SMPI does not comport to a single-factor structure, at least in the current data.

3.2. Exploratory factor analysis

Given the lack of fit of any of the hypothesized models, we conducted an exploratory factory analysis in R using the GPArotation package to examine the factor structure and reliability of the SMPI. Parallel analyses revealed that a seven-factor model fit the data somewhat well (RMSEA = .05 [.04, .05], TLI = .91, mean item

Table 3
Exploratory factor analysis, seven-factor and five-factor solutions (n = 476).

Model	χ^2	Df	χ^2 *diff*	TLI	RMSEA	RMSR	BIC
Five-factor	410.72***	166		0.85	0.06 [0.05, 0.07]	0.04	-579.82
Seven-factor	167.43*	129	243.29	0.91	0.05 [0.05, 0.05]	0.02	-543.28

Factor loadings: five-factor model

	Factor 1 "Melancholia Severity"	Factor 2 "Reactive Depression"	Factor 3 "Positive Mood Reactivity"	Factor 4 "Family Relationships/ Emotionality"	Factor 5 "Early Years"
A1	0.72				
A2	0.63				
A3	0.35				
A4	0.70				
A5	0.53				
A6	0.52				
A7	0.54				
A8					
A9	0.48				
A10					0.91
A11			0.39		
A12		-0.52			
B1			0.63		
B2					
B3			0.60		
B4			0.58		
B5				0.32	
B6			0.33		
B7				0.34	
B8		0.85			
B9		0.79			
B10				0.56	
B11				0.83	
B12					-0.73
R^2	12%	8%	7%	6%	6%

*** = p < .001, * = p < .05, df = degrees of freedom, TLI = Tucker Lewis Index, RMSR = root mean square of residuals; RMSEA = root mean square error of approximation; BIC = Bayesian information criterion.

Table 2
Factor loadings for confirmatory factor analysis of the Sydney Melancholia Prototype Index, bifactor solution (n = 476).

General factor					SA~ and SB~			
	Estimate	Standard error	z-value	*P*	Estimate (SA)	Standard error	z-value	*P*
A1	0.30	0.06	4.98	0.00	0.74	0.06	13.16	0.00
A2	0.53	0.06	9.56	0.00	0.66	0.05	12.86	0.00
A3	0.22	0.07	3.16	0.00	0.41	0.07	5.83	0.00
A4	0.50	0.06	9.14	0.00	0.68	0.05	13.68	0.00
A5	0.67	0.06	11.76	0.00	0.56	0.05	10.44	0.00
A6	0.25	0.07	3.85	0.00	0.56	0.07	8.45	0.00
A7	0.32	0.06	5.02	00	0.56	0.06	8.77	0.00
A8	0.08	0.07	1.18	0.24	0.25	0.08	3.32	0.00
A9	0.52	0.06	8.73	0.00	0.45	0.06	7.58	0.00
A10	-0.23	0.07	-3.18	0.00	0.17	0.06	2.95	0.00
A11	-0.30	0.06	-5.40	0.00	0.21	0.06	3.64	0.00
A12	0.53	0.07	8.09	0.00	0.10	0.07	1.49	0.14
					Estimate (SB)			
B1	-0.86	0.06	-13.97	0.00	0.45	0.07	6.07	0.00
B2	0.18	0.06	2.91	0.00	0.25	0.06	3.95	0.00
B3	-0.80	0.06	-12.73	0.00	0.50	0.07	6.93	0.00
B4	-0.62	0.06	-10.39	0.00	0.46	0.06	7.29	0.00
B5	-0.06	0.07	-0.85	0.40	0.48	0.06	8.03	0.00
B6	0.17	0.08	2.13	0.03	0.33	0.08	4.20	0.00
B7	0.22	0.08	2.91	0.00	0.55	0.07	7.78	0.00
B8	-0.54	0.07	-8.28	0.00	0.17	0.07	2.33	0.02
B9	-0.39	0.07	-5.89	0.00	0.36	0.07	5.31	0.00
B10	0.60	0.07	8.78	0.00	0.43	0.07	5.96	0.00
B11	0.68	0.08	8.15	0.00	0.83	0.08	10.29	0.00
B12	0.25	0.08	3.33	0.00	0.30	0.06	5.19	0.00

complexity = 1.7), as did a five-factor model (RMSEA = .06 [.05, .07], TLI = .85, mean item complexity = 1.6). Only three factors had eigenvalues higher than one, but they did not produce an interpretable solution in that multiple items loaded on to more than one factor. Considering, interpretability of item loadings, the results of the parallel analysis, the scree plot, and the eigenvalues, we chose the more parsimonious but interpretable five-factor model.

The five-factor solution, using a .3 cutoff for each loading, is shown on Table 3. Factor 1 appears to be composed of items capturing "melancholic symptom severity," particularly symptoms of anhedonia and hopelessness. Factor 2 appeared to reflect items capturing "reactive depression," or the extent to which depression occurs in relation to stressors. We dubbed Factor 3 "positive mood reactivity," as the items contained in the factor related to being able to be cheered up, or mood lifting with positive experiences. Factor 4 captured general emotionality including anxiety and emotional instability and also family relations ("Family Relationships/Emotionality"). The final, Factor 5 ("Early years") was only comprised of two items asking about early life experiences.

3.3. Concurrent validity

We examined the correlations between the five factors from EFA model and PHQ-9 ("depression"), GAD-7 ("anxiety"), and SHAPS ("anhedonia"). In our sample, anxiety and anhedonia were moderately correlated (r = .32, 95% CI [.24, .40], $p < .001$), as were anhedonia and depression (r = .45, 95% CI [.37, .52], $p < .001$), and depression and anxiety were highly correlated (r = .73, 95% CI [.68, .77], $p < .001$). The melancholia severity factor (Factor 1) had a moderately strong relationship with depression (r = .42, 95% CI [.35, .49], $p < .001$), which was weaker with GAD (r = .30, 95% CI [.22, 38], $p < .001$) and anhedonia (r = .20, 95% CI [.11, .28], $p < .001$). The reactive depression factor (Factor 2) was similarly related to anhedonia (r = -.28, 95% CI [-.36, -.19], $p < .001$), GAD (r = -.14, 95% CI [-.23, -.05], $p < .01$), and depression (r = -.26, 95% CI [-.34, -.17], $p < .001$). The positive mood reactivity factor (Factor 3) showed the strongest relationship with anhedonia (r = -.34, 95% CI [-.42, -.26], $p < .001$), and was weakly correlated with anxiety (r = -.15, 95% CI [-.24, -.07], $p < .001$), and depression (r = -.24, 95% CI [-.32, -.15], $p < .001$). Our factor capturing emotionality, emotional instability, and family relationships (Factor 4), was most closely related to GAD (r = .58, 95% CI [.52, .64], $p < .001$), but also moderately correlated with depression (r = .46, 95% CI [.39, .53], $p < .001$), and weakly with anhedonia (r = .21, 95% CI [.12, .30], $p < .001$). Lastly, the early years factor (Factor 5) was weakly related to depression (r = -.10, 95% CI [-.19, -.01], $p < .05$) and GAD (r = -.13, 95% CI [-.22, -.04], $p < .01$), and not significantly related to anhedonia (r = -.02, 95% CI [-.11, .07], p = .62).

3.4. SMPI structure in depressed participants

Given that the current analysis was conducted on a sample that was not diagnosed with a depressive disorder, we re-ran all analyses using participants who screened positive for depression as per the PHQ-9 (i.e., PHQ-9 ≥10), n = 216. CFA results remained largely unchanged. A one factor solution fit the data poorly (χ^2 (252) = 996.39, $p < .001$, RMSEA = 0.12 (90% CI = 0.11, 0.13), SRMR = .12, CFI = .37, TLI = 0.31). A two-factor model also fit the data poorly χ^2 (251) = 888.65, $p < .001$, RMSEA = 0.11 (90% CI = 0.10, 0.12), SRMR = .11, CFI = .46, TLI = 0.41). We continued testing the same CFAs as above, and found no changes in our more depressed sample, with the corrected bifactor model remaining the best fitting model χ^2 (223) = 481.72, $p < .001$, RMSEA = 0.07 (90% CI = 0.06, 0.08), SRMR = .07, CFI = .78, TLI = 0.73, see Table 4.

The EFA results suggest that a nine-factor solution fit the data best, (RMSEA = .05 [.02, .06], TLI = .91, mean item complexity = 2), with factor loadings ranging from .31 to .89. The nine-factor solution did not aid in making sense of the SMPI conceptually. A five-factor solution showed a similar pattern as in the unrestricted, full sample (RMSEA = .07 [.05, .08], TLI = .77, mean item complexity = 1.7), with factor loadings ranging from .32 to .86, and items 3 and 11 from scale "A" and items 5 and 10 from scale "B" loading onto none of the factors (see Table 5). Overall, the results did not change dramatically, indicating that even in a more severely depressed sample, the items do not appear to load on to each other as a single factor.

Table 4
Sensitivity analyses, confirmatory factor analysis factor loadings for bifactor corrected solution (n = 216).

General factor					SA~ and SB~			
	Estimate	SE	z-value	p	Estimate (SA)	SE	z-value	P
A1	-0.09	0.10	-0.86	0.39	0.66	0.07	8.93	0.00
A2	-0.44	0.09	-4.84	0.00	0.58	0.08	7.39	0.00
A3	-0.22	0.11	-2.09	0.04	0.36	0.10	3.52	0.00
A4	-0.31	0.09	-3.33	0.00	0.60	0.07	8.28	0.00
A5	-0.50	0.08	-6.31	0.00	0.39	0.08	4.62	0.00
A6	-0.09	0.12	-0.70	0.48	0.67	0.10	6.64	0.00
A7	-0.07	0.11	-0.62	0.54	0.60	0.09	7.03	0.00
A8	-0.01	0.11	-0.05	0.96	0.16	0.11	1.48	0.14
A9	-0.06	0.10	-0.59	0.56	0.56	0.09	6.43	0.00
A10	0.23	0.11	2.09	0.04	0.05	0.09	0.53	0.59
A11	0.25	0.08	3.05	0.00	0.18	0.09	2.03	0.04
A12	-0.07	0.10	-0.67	0.50	0.28	0.10	2.94	0.00
					Estimate (SB)			
B1	0.88	0.08	11.23	0.00	0.12	0.16	0.73	0.47
B2	0.06	0.13	0.48	0.63	-0.57	0.10	-5.46	0.00
B3	0.88	0.08	10.51	0.00	0.23	0.16	1.42	0.16
B4	0.65	0.08	8.49	0.00	-0.01	0.13	-0.07	0.95
B5	0.21	0.10	2.11	0.04	-0.31	0.10	-3.03	0.00
B6	0.15	0.13	1.12	0.26	-0.41	0.13	-3.18	0.00
B7	0.30	0.15	2.04	0.04	0.67	0.13	-5.31	0.00
B8	0.38	0.10	3.92	0.00	-0.07	0.12	-0.58	0.56
B9	0.40	0.10	4.12	0.00	-0.15	0.12	-1.22	0.22
B10	-0.22	0.10	-2.24	0.03	-0.29	0.10	-2.80	0.01
B11	-0.00	0.12	-0.02	0.99	-0.58	0.09	-6.26	0.00
B12	-01	0.11	0.13	0.90	-0.20	0.10	-2.00	0.05

4. Discussion

We undertook what is, to our knowledge, the first published analysis of the psychometric properties of the SMPI, which was designed as a measure of the melancholic presentation of depression (Parker et al., 2010; Parker et al., 2012; Parker et al., 2013; Parker et al., 2013; Parker et al., 2015; Parker et al., 2019). Rather than measuring a single dimension of melancholia, or the constructs of reactive vs. melancholic depression, the SMPI appears to assess overall severity of depressive symptoms, especially anhedonia and hopeless, environmental triggers to depression, positive mood reactivity, general mood instability, and early life experiences. The measure also includes various items that do not appear to co-vary with each other or the rest of the scale. These items measure the presence of depression-associated irritability and weight-loss during times of depression. It is important to note, and underscore, that lack of unidimensionality or items that do not co-vary do not necessarily diminish the validity of a scale (Fried, 2020). Thus, our results should not be taken to undermine the data which support the validity of the SMPI (Parker et al., 2010; Parker et al., 2012; Parker et al., 2013; Parker et al., 2013; Parker et al., 2015; Parker et al., 2019). Instead, the lack of unidimensional factor structure of the measure raises questions about the exact nature of melancholic afflictions, at least as measured by the SMPI.

Before interpreting the current results, it is important to consider several limitations of the current sample. First, we surveyed

Table 5
Sensitivity analyses, exploratory factor analysis, nine-factor and five-factor solutions (n = 216).

Model	χ^2	Df	χ^2 *diff*	TLI	RMSEA	RMSR	BIC
Five-factor	281.11***	166		0.77	0.07 [0.05, 0.08]	0.05	-576.17
Nine-factor	71.74, *ns*	96	209.37	0.91	0.05 [0.02, 0.06]	0.02	-384.92

Factor loadings: five-factor model

	Factor 1 "Melancholia Severity"	Factor 2 "Reactive Depression"	Factor 3 "Positive Mood Reactivity"	Factor 4 "Early Years"	Factor 5 "Appetite"
A1	0.59				
A2	0.52		-0.43		
A3					
A4	0.59				
A5	0.34		-0.48		
A6	0.48				
A7	0.51				
A8					-0.48
A9	0.51				
A10				-0.78	
A11					
A12		-0.68			
B1			0.67		
B2	0.32				
B3			0.65		
B4			0.56		
B5					
B6					0.79
B7	0.36				
B8		0.73			
B9		0.66			
B10					
B11	0.34				
B12				0.86	

*** = $p < .001$, *ns* = not significant, df = degrees of freedom, TLI = Tucker Lewis Index, RMSR = root mean square of residuals; RMSEA = root mean square error of approximation; BIC = Bayesian information criterion.

participants on Amazon's Mechanical Turk. There are concerns about the quality of the data from these samples related to inattention and the presence of automated respondents ("bots"; Kennedy et al., 2018). Prior studies of the SMPI include a clinician-administered version of the questions, which was not possible in our sample. Some items assessed by the SMPI, may be better measured by methods other than self-report. In addition, we opted to measure the responses to the SMPI on a dimensional rather than categorical scale. The original SMPI was categorical indicating the presence or absence of specific features. It is possible that the items are best thought of as binary indicators, and our results would have been different if we had presented participants with a different response format. Although this is an empirical question, to date most psychological and psychometric suggests psychological constructs are best represented on a continuum (e.g, Brown and Barlow, 2005; Kessler et al., 2003; Kraemer et al., 2004; Markon et al., 2011) . Finally, it is possible that our results differ because our sample is not drawn from a tertiary clinic and not Australian. Additional replication in a variety of clinical and subclinical samples is recommended.

While the nature of the current sample may be a source of concern, MTurk workers are at least as attentive or more attentive than other community samples (e.g., college students; Hauser and Schwarz, 2015). Additionally, we employed various checks for attention. Levels of depression severity are relatively well-represented on MTurk (Ophir et al., 2019), and were so in our sample. Moreover, the factor structure we observed from our EFA yielded a sensible result and we obtained reasonable correlations between measures. Several strengths of the study also bear highlighting. First, we performed the first analysis of the factor structure of the SMPI. Additionally, we undertook this analysis in a relatively large sample. We went beyond prior reports to study responses to the SMPI across the range of depression severity. Moreover, we allowed participants to rate their responses on a dimensional scale, which probably does a better job of capturing the specific constructs (e.g., being more emotional than the average person).

One implication of the current findings may be that if there is validity to the concept of melancholia, its associated features may need to be reconceptualized. Some items, like those capturing depression-associated irritability and weight loss, worry-related concentration difficulties, and good relationship/work functioning are included on the measure under the assumption that individuals with melancholia generally have good interepisode functioning but when depressed their depressions are so severe and pervasive that they lose interest in all things (Parker, 2005). The stereotypical view of reactive/non-melancholic depression is that individuals have more general vulnerability to negative affectivity (i.e., including to worry and irritability) that may interact with negative life events to produce depression. It is likely, however, that the tendency to become depressed, even severely depressed, co-varies with other negative emotions like worry or irritability, as analyses of the structure of psychopathology have suggested (Kotov et al., 2017). If anything, anxious depression and irritable depression both seem to be related to more severe presentations than non-anxious or non-irritable depression (Benazzi and Akiskal, 2005; Fava et al., 2004). Accordingly, it does not seem promising to conjecture that there is a subtype of depression that is very severe but less likely to be associated with irritability or worry and another subtype that is milder yet happens to co-vary with worry and irritability. In other words, endogenous, psychotic, or melancholic depressions likely have high rates of worry and irritability.

In a prior study, Parker et al. (2015) had reported that items capturing the extent to which depression occurs as a response to the environment, along with an item capturing anergia, could be used to diagnose melancholia without referencing other SMPI items. Here, we found that the anergia item loaded on to a factor capturing overall symptom severity while the items representing the extent to which depression was "reactive" to the environment loaded on to a single

factor. Horwitz and Wakefield (2007) conceptualized the endogenous activation of depressed mood can be conceptualized as a maladaptive activation of naturally evolved loss responses. However, most (93%) individuals who meet the criteria for MDD report the occurrence of a stressor before the start of their episode (Wakefield et al., 2007). Thus, "true" endogenous depression (i.e., occurring completely "out of the blue") is extremely uncommon. Severely depressed mood in response to minor or no stress may represent either a sequelae of recurrent depression ("stress sensitization") or may happen in a small subgroup of people vulnerable to highly recurrent and endogenous depression (Monroe et al., 2019). An interesting pattern we observed was a distinction between positive mood reactivity (e.g., "even when I am depressed I can generally be cheered up") and negative mood reactivity (i.e., perceiving one's depression has external causal triggers). This is consistent with current conceptualizations of the Research Domain Criteria (RDoC), which places negative affect and positive affect as belonging to two different brain systems (Sanislow et al., 2010).

Although the concept of melancholia is rather old (Parker, 2000, 2005; Parker et al., 2010), it is by no means the *ne plus ultra* of depression subtypes. For starters, adding melancholic symptoms to the symptoms of depression may increase heterogeneity. Fried et al. (2020) estimated that there 10,377 ways to qualify for a diagnosis of major depression, but up to 341,737 to qualify for depression and the melancholic subtype. Additionally, there is limited evidence (Weitz et al., 2015) that melancholia predicts overall depression outcomes or that it predicts differential response to interventions (but, see Parker et al., 2013a). More progress may be made by exploring the different features that make up the construct of melancholia. For these purposes, we recommend further study of the SMPI. Indeed, melacholia as a construct is likely more complicated than a single underlying factor. Among the symptoms and features that are found in the SMPI, those capturing psychomotor disturbances and anhedonia appear to be well-supported as potential endophenotypes of depression (Webb et al., 2016). Given our findings with regard to the fit of the reactive depression factor, future research should explore whether the extent the degree of negative mood reactivity in response to environmental triggers are a meaningful marker of psychopathology. Future work should also examine longitudinal associations between depression severity, contaminant symptoms, positive and negative mood reactivity, and early life experiences.

Authors' contributions

LL-L and MDS designed the study. LAR analyzed the data. LL-L and LAR wrote the initial draft of the manuscript. All authors edited subsequent drafts.

Funding source

This study was partially funded by the Eugenie and Joseph Jones Family Foundation, New Orleans, who had no role in the writing of the current manuscript.

Declaration of Competing Interest

None.

Acknowledgements

We thank the subjects in the study for their participation.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.psychres.2020.113410.

References

American Psychiatric Association (APA), 2013. Diagnostic and Statistical Manual of Mental Disorders: DSM-5. American Psychiatric Publications.

Benazzi, F., Akiskal, H., 2005. Irritable-hostile depression: further validation as a bipolar depressive mixed state. J. Affect. Disord. 84 (2–3), 197–207. https://doi.org/10.1016/j.jad.2004.07.006.

Biderman, M.D., Nguyen, N.T., Cunningham, C.J., Ghorbani, N., 2011. The ubiquity of common method variance: the case of the big five. J. Res. Personal. 45 (5), 417–429. https://doi.org/10.1016/j.jrp.2011.05.001.

Brown, T.A., Barlow, D.H., 2005. Dimensional versus categorical classification of mental disorders in the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders and beyond: Comment on the special section. J. Abnorm. Psychol. 114 (4), 551–556. https://doi.org/10.1037/0021-843X.114.4.551.

Fava, M., Alpert, J.E., Carmin, C.N., Wisniewski, S.R., T1rivedi, M.H., Biggs, M.M., Balasubramani, G., 2004. Clinical correlates and symptom patterns of anxious depression among patients with major depressive disorder in STAR* D. Psychol. Med. 34 (7), 1299–1308. https://doi.org/10.1017/S0033291704002612.

Franken, I.H.A., Rassin, E., Muris, P., 2007. The assessment of anhedonia in clinical and non-clinical populations: further validation of the Snaith–Hamilton Pleasure Scale (SHAPS). J. Affect. Disord. 99 (1–3), 83–89. https://doi.org/10.1016/j.jad.2006.08.020.

Fried, E.I., 2020. Lack of theory building and testing impedes progress in the factor and network literature. PsyArxiv. https://osf.io/6cts9/.

Fried, E.I., Coomans, F., Lorenzo-Luaces, L., 2020. The 341 737 ways of qualifying for the melancholic specifier. Lancet Psychiatry 7 (6), 479–480. https://doi.org/10.1016/S2215-0366(20)30169-3.

Fried, E.I., Nesse, R.M., 2015. Depression is not a consistent syndrome: an investigation of unique symptom patterns in the STAR* D study. J. Affect. Disord. 172, 96–102. https://doi.org/10.1016/j.jad.2014.10.010.

Hauser, D.J., Schwarz, N., 2015. Attentive Turkers: MTurk participants perform better on online attention checks than do subject pool participants. Behav. Res. Methods 48 (1), 400–407. https://doi.org/10.3758/s13428-015-0578-z.

Horwitz, A.V., Wakefield, J.C., 2007. The Loss of Sadness: How Psychiatry Transformed Normal Sorrow into Depressive Disorder. Oxford University Press, New York.

Hu, L.t., Bentler, P.M., 1999. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct. Equ. Model. 6 (1), 1–55. https://doi.org/10.1080/10705519909540118.

Kennedy, R., Clifford, S., Burleigh, T., Jewell, R., Waggoner, P., Jewell, R., Winter, N.J.G., 2018. The shape of and solutions to the MTurk quality crisis. Polit. Sci. Res. Methods 1–56. https://doi.org/10.1017/psrm.2020.6.

Kessler, R.C., Merikangas, K.R., Berglund, P., Eaton, W.W., Koretz, D.S., Walters, E.E., 2003. Mild disorders should not be eliminated from the DSM-V. Arch. Gen. Psychiatry 60 (11), 1117–1122. https://doi.org/10.1001/archpsyc.60.11.1117.

Kotov, R., Krueger, R.F., Watson, D., Achenbach, T.M., Althoff, R.R., Bagby, R.M., Clark, L.A., et al., 2017. The hierarchical taxonomy of psychopathology (HiTOP): a dimensional alternative to traditional nosologies. J. Abnorm. Psychol. 126 (4), 454. https://doi.org/10.1037/abn0000258.

Kraemer, H.C., Noda, A., O'Hara, R., 2004. Categorical versus dimensional approaches to diagnosis: Methodological challenges. J. Psychiatric Res. 38 (1), 17–25. https://doi.org/10.1016/S0022-3956(03)00097-9.

Kroenke, K., Spitzer, R.L., Williams, J.B., 2001. The PHQ-9: validity of a brief depression severity measure. J. Gen. Intern. Med. 16 (9), 606–613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x.

Levis, B., Benedetti, A., Thombs, B.D., 2019. Accuracy of patient health questionnaire-9 (PHQ-9) for screening to detect major depression: individual participant data meta-analysis. Br. Med. J. 365, l1476.

Lorenzo-Luaces, L., 2015. Heterogeneity in the prognosis of major depression: from the common cold to a highly debilitating and recurrent illness. Epidemiol. Psychiatr. Sci. 24 (6), 466–472. https://doi.org/10.1017/S2045796015000542.

Lowe, B., Decker, O., Muller, S., Brahler, E., Schellberg, D., Herzog, W., Herzberg, P.Y., 2008. Validation and standardization of the generalized anxiety disorder screener (GAD-7) in the general population. Med. Care 46 (3), 266–274. https://doi.org/10.1097/MLR.0b013e318160d093.

Malhi, G.S., Parker, G., Greenwood, J., 2005. Structural and functional models of depression: from sub-types to substrates. Acta Psychiatr. Scand. 111 (2), 94–105. https://doi.org/10.1111/j.1600-0447.2004.00475.x.

Markon, K.E., Chmielewski, M., Miller, C.J., 2011. The reliability and validity of discrete and continuous measures of psychopathology: A quantitative review. Psychological Bulletin 137, 856–879. https://doi.org/10.1037/a0023678.

Marsh, H.W., Hau, K.-T., Wen, Z., 2004. In search of golden rules: comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. Struct. Equ. Model. 11 (3), 320–341. https://doi.org/10.1207/s15328007sem1103_2.

Monroe, S.M., Anderson, S.F., Harkness, K.L., 2019. Life stress and major depression: the mysteries of recurrences. Psychol. Rev. 791–816. https://doi.org/10.1037/rev0000157.

Murray, C.J., Vos, T., Lozano, R., Naghavi, M., Flaxman, A.D., Michaud, C., Abdalla, S., 2012. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 380 (9859), 2197–2223. https://doi.org/10.1016/S0140-6736(12)61689-4.

Nakonezny, P.A., Carmody, T.J., Morris, D.W., Kurian, B.T., Trivedi, M.H., 2010. Psychometric evaluation of the Snaith-Hamilton Pleasure Scale in adult outpatients with major depressive disorder. Int. Clin. Psychopharmacol. 25 (6), 328–333. https://doi.org/10.1097/YIC.0b013e32833eb5ee.

Ophir, Y., Sisso, I., Asterhan, C.S.C., Tikochinski, R., Reichart, R., 2019. The Turker Blues: hidden factors behind increased depression rates among Amazon's Mechanical Turkers. Clin. Psychol. Sci. 8 (1), 65–83. https://doi.org/10.1177/2167702619865973.

Parker, G., 2000. Classifying depression: should paradigms lost be regained? Am. J. Psychiatry 157 (8), 1195–1203. https://doi.org/10.1176/appi.ajp.157.8.1195.

Parker, G., 2005. Beyond major depression. Psychol. Med. 35 (4), 467–474. https://doi.org/10.1017/S0033291704004210.

Parker, G., 2007. Defining melancholia: the primacy of psychomotor disturbance. Acta Psychiatr. Scand. 115, 21–30. https://doi.org/10.1111/j.1600-0447.2007.00959.x.

Parker, G., Blanch, B., Paterson, A., Hadzi-Pavlovic, D., Sheppard, E., Manicavasagar, V., Gilfillan, D., 2013. The superiority of antidepressant medication to cognitive behavior therapy in melancholic depressed patients: a 12-week single-blind randomized study. Acta Psychiatr. Scand. 128 (4), 271–281. https://doi.org/10.1111/acps.12049.

Parker, G., Fletcher, K., Barrett, M., Synnott, H., Breakspear, M., Rees, A.-M., Hadzi-Pavlovic, D., 2010. Inching toward Bethlehem: mapping melancholia. J. Affect. Disord. 123 (1), 291–298. https://doi.org/10.1016/j.jad.2009.10.001.

Parker, G., Fletcher, K., Hadzi-Pavlovic, D., 2012. Is context everything to the definition of clinical depression? A test of the Horwitz and Wakefield postulate. J. Affect. Disord. 136 (3), 1034–1038. https://doi.org/10.1016/j.jad.2010.11.021.

Parker, G., McCraw, S., Blanch, B., Hadzi-Pavlovic, D., Synnott, H., Rees, A.-M., 2013. Discriminating melancholic and non-melancholic depression by prototypic clinical features. J. Affect. Disord. 144 (3), 199–207. https://doi.org/10.1016/j.jad.2012.06.042.

Parker, G., McCraw, S., Fletcher, K., Friend, P., Futeran, S., 2013. Validation of a new prototypic measure of melancholia. Compr. Psychiatry 54 (7), 835–841. https://doi.org/10.1016/j.comppsych.2013.02.010.

Parker, G., McCraw, S., Hadzi-Pavlovic, D., 2015. The utility of a classificatory decision tree approach to assist clinical differentiation of melancholic and non-melancholic depression. J. Affect. Disord. 180, 148–153. https://doi.org/10.1016/j.jad.2015.03.057.

Parker, G., McCraw, S., Hadzi-Pavlovic, D., Hong, M., Barrett, M., 2013. Bipolar depression: prototypically melancholic in its clinical features. J. Affect. Disord. 147 (1), 331–337. https://doi.org/10.1016/j.jad.2012.11.035.

Parker, G., Tavella, G., Hadzi-Pavlovic, D., 2019. Identifying and differentiating melancholic depression in a non-clinical sample. J. Affect. Disord. 243, 194–200. https://doi.org/10.1016/j.jad.2018.09.024.

Sanislow, C.A., Pine, D.S., Quinn, K.J., Kozak, M.J., Garvey, M.A., Heinssen, R.K., Cuthbert, B.N., 2010. Developing constructs for psychopathology research: Research Domain Criteria. J. Abnorm. Psychol. 119 (4), 631–639. https://doi.org/10.1037/a0020909.

Shorter, E., 2007. The doctrine of the two depressions in historical perspective. Acta Psychiatr. Scand. 115 (s433), 5–13. https://doi.org/10.1111/j.1600-0447.2007.00957.x.

Simms, L.J., Goldberg, L.R., Roberts, J.E., Watson, D., Welte, J., Rotterman, J.H., 2011. Computerized adaptive assessment of personality disorder: introducing the CAT-PD project. J. Personal. Assess. 93 (4), 380–389. https://doi.org/10.1080/00223891.2011.577475.

Snaith, R.P., Hamilton, M., Morley, S., Humayan, A., Hargreaves, D., Trigwell, P., 1995. A scale for the assessment of hedonic tone the Snaith–Hamilton Pleasure Scale. The British Journal of Psychiatry 167 (1), 99–103. https://doi.org/10.1192/bjp.167.1.99.

Spitzer, R.L., Kroenke, K., Williams, J.B., Lowe, B., 2006. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch. Intern. Med. 166 (10), 1092–1097. https://doi.org/10.1001/archinte.166.10.1092.

Wakefield, J.C., Schmitz, M.F., 2013. When does depression become a disorder? Using recurrence rates to evaluate the validity of proposed changes in major depression diagnostic thresholds. World Psychiatry 12 (1), 44–52. https://doi.org/10.1002/wps.20015.

Wakefield, J.C., Schmitz, M., 2013. Normal vs. disordered bereavement-related depression: are the differences real or tautological? Acta Psychiatr. Scand. 127 (2), 159–168. https://doi.org/10.1111/j.1600-0447.2012.01898.x.

Wakefield, J.C., Schmitz, M., 2013. Can the DSM's major depression bereavement exclusion be validly extended to other stressors?: Evidence from the NCS. Acta Psychiatr. Scand. 128 (4), 294–305. https://doi.org/10.1111/acps.12064.

Wakefield, J.C., Schmitz, M.F., First, M.B., Horwitz, A.V., 2007. Extending the bereavement exclusion for major depression to other losses: evidence from the National Comorbidity Survey. Arch. Intern. Med. 64 (4), 433–440. https://doi.org/10.1001/archpsyc.64.4.433.

Webb, C.A., Dillon, D.G., Pechtel, P., Goer, F.K., Murray, L., Huys, Q.J., Parsey, R., 2016. Neural correlates of three promising endophenotypes of depression: evidence from the EMBARC study. Neuropsychopharmacology 41 (2), 454–463. https://doi.org/10.1038/npp.2015.165.

Weitz, E.S., Hollon, S.D., Twisk, J., Van Straten, A., Huibers, M.J., David, D., Cristea, I.A., 2015. Baseline depression severity as moderator of depression outcomes between cognitive behavioral therapy vs pharmacotherapy: an individual patient data meta-analysis. JAMA Psychiatry 72 (11), 1102–1109. https://doi.org/10.1001/jamapsychiatry.2015.1516.