Mood Instability and Irritability as Core Symptoms of Major Depression: An Exploration Using Rasch Analysis

Lloyd Balbuena[1]*, Rudy Bowen[1]*, Marilyn Baetz[1] and Steven Marwaha[2]

[1] Psychiatry, University of Saskatchewan, Saskatoon, SK, Canada, [2] Mental Health and Wellbeing, Warwick Medical School, Coventry, UK

Background: Mood instability (MI) and irritability are related to depression but are not considered core symptoms. Instruments typically code clusters of symptoms that are used to define syndromic depression, but the place of MI and irritability has been under-investigated. Whether they are core symptoms can be examined using Rasch analysis.

Method: We used the UK Psychiatric Morbidity Survey 2000 data ($n = 8,338$) to determine whether the nine ICD/DSM symptoms, plus MI and irritability, constitute a valid depression scale. Rasch analysis was used, a method concerned with ensuring that items constitute a robust scale and tests whether the count of symptoms reflects an underlying interval-level measure. Two random samples of 500 were drawn, serving as calibration and validation samples. As part of the analysis, we examined whether the candidate symptoms were unidimensional, followed a Guttman pattern, were locally independent, invariant with respect to age and sex, and reliably distinguished different levels of depression severity.

Results: A subset of five symptoms (*sad, no interest, sleep, cognition, suicidal ideas*) together with *mood instability* and *irritability* satisfactorily fits the Rasch model. However, these seven symptoms do not separate clinically depressed persons from the rest of the population with adequate reliability (Cronbach α = 0.58; Person Separation Index = 0.35), but could serve as a basis for scale development. Likewise, the original nine DSM depression symptoms failed to achieve satisfactory reliability (Cronbach α = 0.67; Person Separation Index = 0.51).

Limitations: The time frame over which symptoms were experienced varied, and some required recall over the last year. Symptoms other than those examined here might also be core depression symptoms.

Conclusion: Mood instability and irritability are candidate core symptoms of the depressive syndrome and should be part of its clinical assessment.

Keywords: mood instability, depression, psychometrics, diagnosis, health surveys

INTRODUCTION

Depression is a common condition with an estimated lifetime prevalence in the USA of about 16% (1). It is an important cause of workdays lost to disability (2) and is as impairing as arthritis, diabetes, and cardiovascular disease (3). The cost of sub-syndromal symptoms probably exceeds that of formally diagnosed major depression (3–6). It is a concern that the incidence of suicide – the most tragic consequence of depression – has not decreased over decades (7). Clearly, we need to better understand the depressive syndrome and the symptoms used in its assessment.

The conceptualization, assessment, and measurement of major depression are tricky, and this shows in the poor reliability and validity of its instruments (8–10). Non-cohesive symptoms might partly explain why specific genetic, biological, or psychological underpinnings are poorly understood (11–13). While depressive symptoms diverge in their association with external variables – as with cognitive and neuro-vegetative symptoms (14) – a particular symptom can be shared by different disorders. For example, it is unclear whether agitation is an indicator of anxiety or depression and whether it is because agitation is related to the higher construct of distress (14, 15). Two individuals can share the same major depression diagnosis without sharing a single symptom (16). Calculating the prevalence and burden of depression is made challenging by the heterogeneity of studies, partly a result of differences in measurement (17).

In clinics worldwide, diagnosing major depression is fairly straightforward. Primary care and specialist physicians follow the DSM (which requires 5 of 9 symptoms) or the ICD (which requires 4 of 10 symptoms) (18). Interestingly, prevalence estimates are similar between systems, although somewhat different populations are identified (19). Having equivalent diagnostic systems has simplified the work of health systems with regard to billing for services and clinical communication (18), but has left important conceptual work unattended. Two problems with the DSM criteria, and by extension, the ICD were raised by Kendler (20). First, the criteria are narrower than the symptoms known to the Western tradition of psychiatry, resulting in an impoverished concept (20). This is perhaps understandable because a list of diagnostic symptoms needs to be brief. Second, the DSM criteria are reified, in the sense that they are thought to constitute depression itself, instead of selected signs of depression (20). In health systems where time is a premium, relying solely on the checklist of symptoms, and ultimately on symptom counts, is a common practice.

Two potential candidate symptoms of the depressive syndrome are MI and irritability. By mood, we mean a valenced emotional state (i.e., positive or negative) (21) in a patient. MI can be defined as "rapid oscillations of intense affect, with a difficulty in regulating these oscillations or their behavioral consequences" (22). MI is closely associated with depression in both cross-sectional and longitudinal studies (23, 24). The prevalence of MI is about 14% in the UK general population and about 61% in participants with depression (25), suggesting that MI could be important in diagnosing depression. MI is central to neuroticism (26) that, in turn, is the personality trait most consistent predictor of depression (27).

DSM-V and ICD-10 accounts of major depression mention irritability in their narrative descriptions, but do not include it in the list of diagnostic symptoms (4, 28). Hence, it could be ignored by clinicians who follow the nine standard symptoms, as if it were an exhaustive list. Yet, it is reported that irritability occurs in one-third to one-half of child and adult patients with major depression (29–31), and is part of a strong principal factor of major depression (29). Irritable depression is also associated with greater severity, lower quality of life, and a history of suicidal attempts, which is itself a criterion for depression (29). These findings, as well as Kendler's critique suggest that the ICD or DSM symptom lists are incomplete.

Our research questions are:

(i) Do the DSM/ICD symptoms for major depression constitute a valid measure?
(ii) Are MI and irritability symptoms of depression?

We addressed these questions using Rasch analysis, which tests a crucial assumption in scales: the total score (or count of symptoms) is an adequate, equidistant representation of depression levels. In brief, the objective of Rasch modeling is to verify that questionnaires have the properties of physical measures (e.g., a ruler). The units are equally spaced, measure a single attribute (i.e., length), and are additive. Moreover, the reading does not depend on the properties of the entity being measured or the person making the measurement.

MATERIALS AND METHODS

Data

We used data from the 2000 Psychiatric Morbidity Survey (PMS) of Great Britain. The main purpose of the survey was to estimate the prevalence of psychiatric disorders and their correlates using a stratified random multistage design. Participants were 8,580 adults, aged 16–74 years, living in private households in Great Britain. Of these, the 8,338 people (97%) who had complete records of symptoms of interest were the population from which our calibration and verification samples were drawn. Full details of the PMS methods are available in the main survey report (32).

Measures
Depressive Symptoms

Participants were assessed for depression and anxiety disorders by trained lay interviewers who used the Clinical Interview Schedule-Revised (CIS-R). This is a reliable and valid instrument that can be used to algorithmically assign an ICD-10 diagnosis (33). We selected CIS-R questions that were similar in meaning and wording to the nine symptoms of major depression specified by DSM-V. The DSM-V depression symptoms only has "subtle changes" over DSM-IV (34), while both ICD-10 and the upcoming ICD-11 are designed to harmonize with DSM (35, 36). Where the DSM-V symptoms had multiple parts, we combined the participant's answers to multiple CIS-R questions. The CIS-R questions we included for analysis are the following: "sad, miserable, or depressed," "unable to enjoy or take an interest

in things," "loss of appetite/weight except on a diet," "problems getting to sleep/sleeping more than usual," "restless, walking more slowly, less talkative," "tired except from doing exercise/ lacking in energy," "felt guilty/blamed self/felt not as good as other people," "problems concentrating/forgetting things," "life not worth living/wished for death/thought of suicide." The time frame over which symptoms were experienced differed for different symptoms (weeks to years), so the duration and timing of occurrence were disregarded.

MI and Irritability

These were assessed within the participant-completed Structured Clinical Interview for DSM-IV Axis II Personality Disorders (SCID-II): borderline personality disorder section (37). The question that assessed MI was "Do you have a lot of sudden mood changes?" There were two questions on irritability: (a) "Many people become irritable or short tempered at times, though they may not show it. Have you felt irritable or short tempered with those around you in the past month?" and (b) "During the past month did you get short tempered or angry over things which now seem trivial when you look back on them?"

Analytical Strategy

Although Rasch analysis is increasingly used in other medical specialties, it is still largely underutilized in psychiatry (38, 39). As mentioned in the Section "Introduction," Rasch analysis determines whether mental or psychological scales have the characteristics of physical measures. For this to be the case, five conditions must be met (**Figure 1**). First, the instrument is designed to measure a single attribute (unidimensionality). Just as a ruler measures length only, depression scales must measure depression alone. Second, the responses follow a Guttman pattern: persons are ranked from lowest to highest levels of the trait, while items are ranked by highest to lowest levels of endorsement. The

appearance of a Guttman pattern is like a staircase. The ranking of persons and items in this manner produces units that are interval-level measures called logits ("log odds unit"). Third, endorsing a particular question should be independent of the endorsement of another question except with respect to the attribute being measured (local independence). This requirement guards against spurious correlations – those that are due to external factors, such as wording or position in the scale. Fourth, the items and the overall instrument must *not* have differential item or test function. This means that at the question and instrument levels, there must be invariance with respect to person characteristics like age and gender. Fifth, the overall scale must be internally consistent (as measured by Cronbach's alpha) and able to distinguish different strata of respondents along the latent trait (40, 41).

We performed Rasch analysis in two samples of 500 subjects, one serving as calibration and the other as validation sample. The requirements, tests performed, and the criteria in each test are summarized in **Table 1**. For the complete description of analysis steps, please refer to the Appendix in Data Sheet S1 in Supplementary Material.

RESULTS

We refer the interested reader elsewhere (32) for a description of the demographic characteristics of all 8,580 PMS participants. Our calibration and verification samples were similar in age distribution (mean = 45 years), mean frequency of depression symptoms endorsed (about four symptoms), sex, and living arrangements. Please refer to **Table 2** for details.

Unidimensionality

Parallel analysis of the 11 candidate items in the calibration sample showed two dimensions (adjusted eigenvalues: 2.82 and 1.02). In the validation sample, a similar result was reached with

FIGURE 1 | **Schematic diagram of analysis steps**.
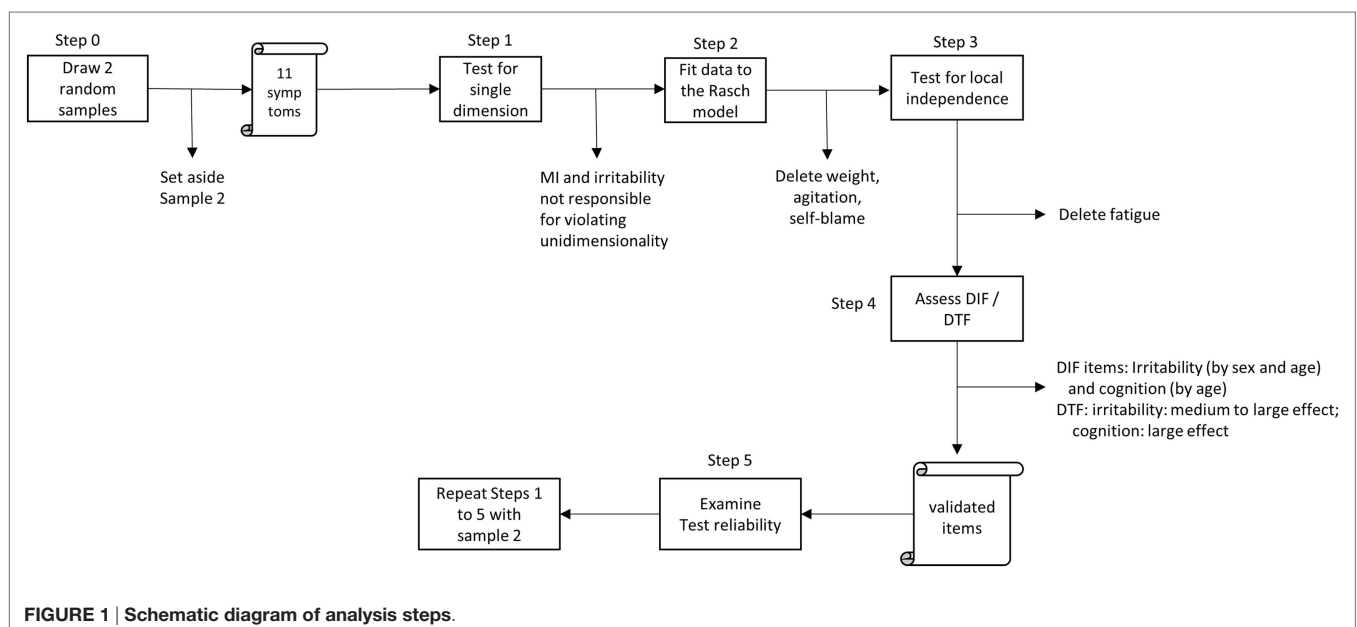
TABLE 1 | Assessment of fit with Rasch model assumptions and the criteria used.

Requirement	Test(s)	Criterion
1. Unidimensional	Parallel analysis comparing actual data with artificial data (scope: the entire set of questions) (42)	Retain the nth component for which the eigenvalue in the actual data is greater than the 95th percentile in 5,000 simulated data (43)
2. Guttman pattern	Infit mean square value (scope: per question)	Must be between 0.91 and 1.09 (44)
3. Local independence	Residual correlations (scope: by pairs of questions)	Not greater than 0.2 and p values (with Holm's adjustment) not greater than 0.05 (45, 46)

4. Invariance	(a) Uniform DIF: Mantel-Haenszel test (scope: per question)
	(b) Non-uniform DIF: Breslow-Day (scope: per question)
	(c) Differential Test Functioning: tau-squared: the variance of DIF across all items (scope: the entire set of questions)
		(a)-(b) Neither of the two tests has a significant value (47)
		(c) tau-squared < 0.07: small DTF
		    0.07 < tau-squared < 0.14: medium DTF
		    tau-squared > 0.14: large DTF (48)

5. Reliability	(a) Cronbach alpha (b) Person Separation Index (scope: the entire set of questions)	(a) alpha > 0.8 (b) PSI > 0.7 (41)

TABLE 2 | Demographic variables of the calibration and validation samples taken from the UK Psychiatric Morbidity Survey 2000.

	Sample 1 (n = 500)	Sample 2 (n = 500)	p
Mean number of symptoms endorsed (of 11 symptoms)	3.83 (2.52)	3.93 (2.46)	0.52
Mean age (SD)	44.48 (15.78)	45.23 (15.64)	0.46
Sex			
Male (%)	213 (42.60)	193 (38.60)	0.20
Female (%)	287 (57.40)	297 (61.40)	
Living arrangements			
Married/cohabiting (%)	273 (54.60)	286 (57.20)	0.70
Single (%)	115 (23.00)	107 (21.40)	
Widowed/divorced/separated (%)	112 (22.40)	107 (21.40)	

eigenvalues: 2.84 and 1.04. To determine whether MI and/or irritability were responsible for violating unidimensionality, we omitted them both and performed parallel analysis using only the typical nine depression symptoms. Once again, two dimensions were detected. The eigenvalues were 2.60 and 1.02 for the calibration sample and 2.57 and 1.04 for the validation sample. After misfitting items (i.e., weight/appetite change, agitation/retardation, fatigue, self-blame) were dropped, we tested dimensionality with the retained items (i.e., sad, lack of interest, sleep problems, cognition problems, and suicidal ideas) plus MI and irritability. Only one adjusted eigenvalue was greater than one in both the calibration and validation samples (1.87 and 1.81, respectively). We took these results as evidence that MI and irritability are part of the core indicators of depression.

Assessment of Fit with a Probabilistic Guttman Pattern

Of the 11 symptoms we examined, the most common symptom was irritability, while MI was the least common. Please refer to **Table 3** for the complete list of item locations in logits. **Figure 2** is a visual representation of the item locations and the corresponding fraction of the population that they demarcate.

The symptoms that misfit the Rasch model were weight/appetite change, agitation/retardation, fatigue, and self-blame. With consistent findings in both samples, these symptoms were

eliminated. See **Table 4** for the initial and final symptom lists and fit statistics.

Test of Local Independence

In the calibration sample, large residual correlations were observed between the item pairs fatigue and cognition, $\chi^2 = 14.36$, df = 1, p with Holm's adjustment = 0.004. The validation sample showed a similar large residual correlation for these two items, $\chi^2 = 11.26$, df = 1, p with Holm's adjustment = 0.02. After fatigue was eliminated from the symptom list, no large residuals remained.

Differential Item/Test Function

With the assessment of differential response by gender, no items were flagged for DIF in the calibration sample. In the validation sample, irritability showed both uniform and non-uniform effects. Female respondents more frequently endorsed irritability, and the disparity with male respondents also differed by depression level. The impact of DIF by gender with respect to irritability had a medium sized effect on the test for the validation sample. See **Table 5** for details.

With the DIF assessment by age group, cognition showed both uniform DIF. In the calibration sample, respondents above 45 years of age endorsed cognitive problems more frequently. In the validation sample, respondents above age 45 endorsed irritability more frequently. Both cognition and irritability had large test effects in both samples. See **Table 6** for details.

Test Reliability

The PSI for the initial 11 symptoms was 0.60 for the calibration sample and 0.58 for the validation sample. For the seven retained symptoms (i.e., sad, no interest, sleep, cognition, suicidal ideas, MI, irritability) the PSIs were 0.38 and 0.35 for the calibration and validation samples, respectively. Cronbach alphas for the 11 symptoms (nine original, plus MI and irritability) were 0.71 and 0.70 for the calibration and validation samples, respectively, and 0.58 for both samples for the seven retained symptoms.

In *post hoc* analysis, we examined the PSI and alpha for the standard nine DSM/ICD symptoms. PSIs were 0.52 and 0.51, and Cronbach alphas were 0.70 and 0.67 for calibration and validation, respectively.

TABLE 3 | Item difficulties (in logits) in the calibration and validation datasets.

Item	Calibration		Validation	
	Item difficulties in the initial list	Item difficulty after eliminating misfitting items	Item difficulties in the initial list	Item difficulty (after eliminating misfitting items)
1. Sad	0.05	0.06	−0.04	−0.02
2. No interest	1.60	1.57	1.53	1.50
3. Weight/appetite change	0.72	N/A	0.74	N/A
4. Sleep	0.17	0.18	0.14	0.14
5. Agitation, retardation	2.23	N/A	2.00	N/A
6. Fatigue	−0.25	N/A	−0.19	N/A
7. Self-blame	2.23	N/A	2.18	N/A
8. Cognition	0.39	0.39	0.40	0.40
9. Suicidal ideas	1.14	1.13	1.05	1.03
10. MI	2.33	2.29	2.10	2.05
11. Irritability	−0.77	−0.75	−0.96	−0.92

FIGURE 2 | Distribution of persons and items along the depression latent trait.

TABLE 4 | Item fit statistics in the calibration and validation datasets.

Item	Calibration		Validation	
	Infit t in the initial list	Infit t after eliminating misfitting items	Infit t in the initial list	Infit t after eliminating misfitting items
1. Sad	0.91	0.93	0.94	0.93
2. No interest	0.97	1.00	1.00	1.04
3. Weight/appetite change	1.19	N/A	1.14	N/A
4. Sleep	1.08	1.07	1.04	1.02
5. Agitation, retardation	0.80	N/A	0.83	N/A
6. Fatigue	0.97	N/A	0.99	N/A
7. Blame no good	0.81	N/A	0.82	N/A
8. Cognition	0.96	0.98	0.97	0.97
9. Suicidal ideas	1.11	1.04	1.10	1.04
10. MI	1.01	0.96	1.00	0.97
11. Irritability	1.05	1.00	1.09	1.03

TABLE 5 | DIF by gender.

Item	Sample 1 (n = 500)					Sample 2 (n = 500)				
	Mantel–Haenszel χ^2[♣]	p-Value	Breslow-Day[†]	p-Value	Combined decision rule*	Mantel–Haenszel χ^2[♣]	p-Value	Breslow-Day[†]	p-Value	Combined decision rule*
Sad	0.85	0.36	5.34	0.25	Ok	1.17	0.28	4.00	0.55	Ok
No interest	0.03	0.85	0.65	0.99	Ok	0.00	0.97	2.39	0.79	Ok
Sleep	1.82	0.18	7.54	0.18	Ok	5.00	0.03	2.70	0.75	Ok
Cognition	2.94	0.09	4.15	0.53	Ok	1.82	0.18	2.47	0.78	Ok
Suicidal ideas	0.5	0.82	12.04	0.03	Ok	0.41	0.52	2.36	0.80	Ok
Mood instability	0.07	0.79	3.36	0.50	Ok	0.19	0.66	4.32	0.50	Ok
Irritability	1.00	0.32	4.60	0.33	Ok	9.87	0.00*	4.10	0.54	Flag
Differential test function (DTF)	Tau²: −0.001 (small effect)					Tau²: 0.11 (medium effect)				

[♣]*Tests uniform DIF which means that one group systematically scores higher.*

[†]*Tests non-uniform DIF which means that one group scores higher at lower levels of depression but the pattern reverses at higher levels of depression.*

An item is flagged for differential function if either the Mantel-Haenszel or the Breslow-Day test is significant.

TABLE 6 | DIF by age group.

Item	Sample 1 (n = 500)					Sample 2 (n = 500)				
	Mantel–Haenszel χ^2[♣]	p-Value	Breslow-Day[†]	p-Value	Combined decision rule*	Mantel–Haenszel χ^2[♣]	p-Value	Breslow-Day[†]	p-Value	Combined decision rule*
Sad	1.29	0.26	1.82	0.77	Ok	2.27	0.13	1.65	0.90	Ok
No interest	1.68	0.19	1.98	0.85	Ok	0.24	0.62	2.81	0.73	Ok
Sleep	0.20	0.66	1.68	0.89	Ok	6.36	0.01	2.41	0.79	Ok
Cognition	12.26	<0.001*	3.19	0.67	Flag	3.86	0.05	6.74	0.24	Ok
Suicidal ideas	0.27	0.61	4.63	0.46	Ok	3.82	0.05	5.96	0.31	Ok
Mood instability	6.62	0.01	2.56	0.63	Ok	3.11	0.07	4.23	0.52	Ok
Irritability	8.50	0.004	6.35	0.17	Ok	10.95	<0.001*	2.80	0.73	Flag
Differential test function (DTF)	Tau²: 0.29 (large effect)					Tau²: 0.21 (large effect)				

[♣]*Tests uniform DIF which means that one group systematically scores higher.*

[†]*Tests non-uniform DIF which means that one group scores higher at lower levels of depression but the pattern reverses at higher levels of depression.*

An item is flagged for differential function if either the Mantel-Haenszel or the Breslow-Day test is significant.

DISCUSSION

In this work, we sought to clarify what symptoms form the most statistically cohesive set to measure depression as a construct. We now discuss the practical and theoretical implications of our findings.

Measuring Depression

We found that the nine depression symptoms in the DSM/ICD systems are not unidimensional. In practice, this means that the "5 of 9" rule (or "4 of 10" for ICD) is probably not warranted because the symptoms do not all tap the same attribute. A more homogeneous set of indicators is achieved by removing weight and appetite change, agitation and retardation, and feelings of worthlessness and inappropriate guilt and fatigue from the core of the major depression syndrome. The findings that MI and irritability fit the Rasch model indicates that they belong to the core network/cluster of symptoms that includes sadness and anhedonia (49).

While irritability appears to be endorsed differentially by sex and by age group, MI is invariant with respect to both characteristics. That irritability is identified as a DIF item should not

automatically exclude it from the list. One possibility is to adjust for the DIF effect of gender and/or age in assessing the severity of depression (50). This would be difficult to implement in a paper and pencil test, but could be solved by computer adaptive testing that takes covariates into account.

Seven items – sad, lack of interest, sleep, cognition, suicidal ideas, MI, and irritability – could serve as the kernel for a depression measure, but on their own do not reliably separate depressed persons from the rest. Likewise, the list of DSM/ICD symptoms fails the typical criterion for internal consistency (alpha < 0.80) and also falls short of distinguishing the depressed from the rest (PSI < 0.70). Clinical judgment and contextual information may need to be taken into account apart from the canonical list of symptoms. For use outside of the clinic, scales such as the PHQ-9, Beck Depression Inventory, HADS, CES-D, and the like are only recently being examined using item response theory (51–53).

Reconceptualizing Depression

Mood instability is common in depression (23, 25) and has been shown to be a precursor of depression (24). Our current results provide evidence that MI is a symptom of depression. According to DSM, relatively short durations of MI phases would not meet

episode criteria for major depression (2 weeks) or for hypomania (4 days) (4). If the patient reports rapid mood fluctuations, it is typical to either dismiss MI as clinically unimportant or to concurrently diagnose a personality disorder, particularly borderline personality disorder where both MI and irritability are DSM-V criteria (4). Unfortunately, people who do not fulfill duration criteria may also be considered "well" or at least not depressed and receive no treatment (54), but these people are at higher risk of developing future depression (55). The evidence indicates that intense and frequent mood swings are associated with severe distress (56, 57), and there is merely a quantitative difference between the mood fluctuations of normal individuals and those of patients (58). MI is linked to other indicators of distress and impairment such as health care utilization, medication use, and suicidal thoughts (25, 59) and has recently been proposed to fit the characteristics of the Research Domain Criteria (60).

Irritability is associated with emotional lability in patients with unipolar depression (30) and in university students (30, 61) and is common in depression (29, 30). Mixed depression, which may be defined as "an overlapping of manic and depressive symptoms" includes irritability and emotional lability among its symptoms (62). This presentation is characterized by psychic and motor agitation, accompanied by intense suffering, which put the patient at increased risk of suicide (54). Irritability could be a core symptom of depression (29), an indicator of a more severe and chronic course (30, 63) or a feature of bipolar disorders (64). DSM-V has included irritability as a core symptom of mania, generalized anxiety disorder, and borderline personality disorder. but excluded it as a symptom of major depression. It has previously been rejected as a symptom of major depression as it does not appreciably increase the prevalence above that of sadness and loss of interest (29, 65). Increased prevalence is not necessarily a good basis for defining a syndrome. Conversely, irritability, along with MI, could lead to longstanding interpersonal and adjustment difficulties that could lead to depression (30, 49). In summary, both MI and irritability are observed in a range of psychiatric disorders.

It is uncertain whether agitation and retardation are specific distinguishing features of major depression, melancholia, mixed mood states, atypical depression, bipolar II depression, or anxiety comorbid with mood disorders (66, 67). Agitation is a defining characteristic of a proposed mixed depressive state that has both melancholic and excitatory features, but which does not have the levity in mood of hypomanic patients (62). Our finding is more consistent with a major depression study that reported that agitation could be dropped from the definition of major depression with no loss of validity (68).

The Feighner symptom of "self-reproach or guilt" was expanded in DSM-III to include "feelings of worthlessness …" as part of the DSM trend to broaden the criteria for major depression (69). There is a clear semantic difference between guilt (worry) about past misdeeds and anxiety (worry) about future threats and perhaps feeling helpless, but this distinction might not be meaningful for people with common mental disorders, high comorbidity, or high distress (70–73). Our results replicate findings of a previous Rasch analysis of the PHQ-9 scale that guilt was not coherent with the model of depression (51).

We eliminated fatigue because of higher than expected correlation with cognitive problems. This could be the result of similar wording: both symptoms are presented as diminished ability. An alternative to eliminating this item is rewording either fatigue or cognitive problems. Retaining this item is probably the more prudent course of action. Although one study reported that fatigue is not unidimensional with the other depression symptoms (74), several other studies reported that it satisfies the Rasch model (11, 52, 75).

It should be emphasized that the misfitting items are frequently experienced by patients with major depression. What is in question is whether they are central to the network of symptoms comprising the depressive syndrome and whether they are useful in its assessment. The search for underlying biological or psychological aberrations (76) or treatment (77) for depression is probably hampered by a heterogeneous cluster of symptoms.

Our study has several limitations. First, the data are based on retrospective recall with all of the disadvantages of this method (78). People who are depressed have a general negative recall bias that might affect reporting of symptoms (78). Second, the CIS-R was designed to elicit the ICD-10 criteria for depression, although the symptoms are very similar to those of DSM-V. Third, we did not consider duration criteria for the individual symptoms, and thus cannot be certain that all symptoms occurred at the same time. However, all duration criteria of major depression and similar groupings are arbitrary and, theoretically, symptoms may occur sequentially and still indicate the same syndrome (55, 79). Fourth, we studied MI and irritability, but other common symptoms such as anxiety, rumination and physical symptoms should be studied (73, 80). Fifth, we performed DIF analysis only by age and gender in a British sample, so further analysis is required to determine if MI and irritability are part of the syndrome across cultures. The differential item status of sleep, irritability, and MI should be addressed by future work. Sixth, while MI and irritability were shown to load on a single factor with depression, we did not have an independent external criterion to serve as a reference. Finally, the PMS did not assess bipolar disorder so it is possible that some of the respondents had bipolar, instead of unipolar depression. We do not think this limitation undermines our findings because MI is a feature of a wide range of psychiatric disorders (22, 81) and, second, the prevalence of people with bipolar depression in the sample in comparison to unipolar depression is likely to have been small. It would be beneficial in replication studies to correlate the scores of patients in our proposed 7-item scale to standard psychometric questionnaires, such as the Mood Disorders Questionnaire or the Affective Lability Scale.

A strength of our study is that it was based on empirical data obtained from an epidemiological sample of the population. Accordingly, it was not constrained by a pre-selected sample with major depression as diagnosed by the criteria being studied.

CONCLUSION

Mood instability and irritability are candidate core symptoms of the depressive syndrome and should be part of its clinical assessment.

AUTHOR CONTRIBUTIONS

LB, RB, MB, and SM conceptualized the study. SM and LB acquired the data. LB, RB, MB, and SM searched the literature on major depression, irritability, and mood instability. LB analyzed the data. LB and RB wrote the initial draft. MB and SM critically reviewed and commented on the initial draft. LB, RB, MB, and SM interpreted the results. LB, RB, MB, and SM approved the final draft for submission and are accountable for the accuracy and integrity of the work.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://journal.frontiersin.org/article/10.3389/ fpsyt.2016.00174

REFERENCES

1. Kessler RC, Berglund P, Demler O, Jin R, Koretz D, Merikangas KR, et al. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *JAMA* (2003) 289(23):3095–105. doi:10.1001/jama.289.23.3095

2. Mrazek DA, Hornberger JC, Altar CA, Degtiar I. A review of the clinical, economic, and societal burden of treatment-resistant depression: 1996-2013. *Psychiatr Serv* (2014) 65(8):977–87. doi:10.1176/appi.ps.201300059

3. Wells KB, Stewart A, Hays RD, Burnam MA, Rogers W, Daniels M, et al. The functioning and well-being of depressed patients. Results from the Medical Outcomes Study. *JAMA* (1989) 262(7):914–9. doi:10.1001/jama.1989.03430230083030

4. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders: DSM-5*. Washington, DC: American Psychiatric Association (2013).

5. Karsten J, Hartman CA, Ormel J, Nolen WA, Penninx BW. Subthreshold depression based on functional impairment better defined by symptom severity than by number of DSM-IV symptoms. *J Affect Disord* (2010) 123(1–3):230–7. doi:10.1016/j.jad.2009.10.013

6. Rai D, Skapinakis P, Wiles N, Lewis G, Araya R. Common mental disorders, subthreshold symptoms and disability: longitudinal study. *Br J Psychiatry* (2010) 197(5):411–2. doi:10.1192/bjp.bp.110.079244

7. Navaneelan T. *Suicide Rates: An Overview*. (2012). Available from: http://www.statcan.gc.ca/pub/82-624-x/2012001/article/11696-eng.htm

8. Ghaemi SN. Why antidepressants are not antidepressants: STEP-BD, STAR*D, and the return of neurotic depression. *Bipolar Disord* (2008) 10(8):957–68. doi:10.1111/j.1399-5618.2008.00639.x

9. Lieblich SM, Castle DJ, Pantelis C, Hopwood M, Young AH, Everall IP. High heterogeneity and low reliability in the diagnosis of major depression will impair the development of new drugs. *BJPsych Open* (2015) 1:e5–7. doi:10.1192/bjpo.bp.115.000786

10. Parker G. Beyond major depression. *Psychol Med* (2005) 35(4):467–74. doi:10.1017/S0033291704004210

11. da Rocha NS, Chachamovich E, de Almeida Fleck MP, Tennant A. An introduction to Rasch analysis for psychiatric practice and research. *J Psychiatr Res* (2013) 47(2):141–8. doi:10.1016/j.jpsychires.2012.09.014

12. Ostergaard SD, Jensen SO, Bech P. The heterogeneity of the depressive syndrome: when numbers get serious. *Acta Psychiatr Scand* (2011) 124(6):495–6. doi:10.1111/j.1600-0447.2011.01744.x

13. Trivedi MH. Major depressive disorder: remission of associated symptoms. *J Clin Psychiatry* (2006) 67(Suppl 6):27–32.

14. Lux V, Kendler KS. Deconstructing major depression: a validation study of the DSM-IV symptomatic criteria. *Psychol Med* (2010) 40(10):1679–90. doi:10.1017/S0033291709992157

15. Pallant JF, Tennant A. An introduction to the Rasch measurement model: an example using the Hospital Anxiety and Depression Scale (HADS). *Br J Clin Psychol* (2007) 46:1–18. doi:10.1348/014466506X96931

16. Olbert CM, Gala GJ, Tupler LA. Quantifying heterogeneity attributable to polythetic diagnostic criteria: theoretical framework and empirical application. *J Abnorm Psychol* (2014) 123(2):452–62. doi:10.1037/a0036068

17. Ferrari AJ, Somerville AJ, Baxter AJ, Norman R, Patten SB, Vos T, et al. Global variation in the prevalence and incidence of major depressive disorder: a systematic review of the epidemiological literature. *Psychol Med* (2013) 43(3):471–81. doi:10.1017/S0033291712001511

18. National Collaborating Centre for Mental Health (Great Britain), National Institute for Health and Clinical Excellence (Great Britain), Royal College of Psychiatrists, British Psychological Society. *Depression: The Treatment and Management of Depression in Adults*. Updated ed. London: British Psychological Society and the Royal College of Psychiatrists (2010). 705 p.

19. Andrews G, Anderson TM, Slade T, Sunderland M. Classification of anxiety and depressive disorders: problems and solutions. *Depress Anxiety* (2008) 25(4):274–81. doi:10.1002/da.20489

20. Kendler KS. The phenomenology of major depression and the representativeness and nature of DSM criteria. *Am J Psychiatry* (2016) 173(8):771–80. doi:10.1176/appi.ajp.2016.15121509

21. Vidal-Ribas P, Brotman MA, Valdivieso I, Leibenluft E, Stringaris A. The status of irritability in psychiatry: a conceptual and quantitative review. *J Am Acad Child Adolesc Psychiatry* (2016) 55(7):556–70. doi:10.1016/j.jaac.2016.04.014

22. Marwaha S, He Z, Broome M, Singh SP, Scott J, Eyden J, et al. How is affective instability defined and measured? A systematic review. *Psychol Med* (2014) 44(9):1793–808. doi:10.1017/S0033291713002407

23. Bowen RC, Wang Y, Balbuena L, Houmphan A, Baetz M. The relationship between mood instability and depression: implications for studying and treating depression. *Med Hypotheses* (2013) 81(3):459–62. doi:10.1016/j.mehy.2013.06.010

24. Marwaha S, Balbuena L, Winsper C, Bowen R. Mood instability as a precursor to depressive illness: a prospective and mediational analysis. *Aust N Z J Psychiatry* (2015) 49(6):557–65. doi:10.1177/0004867415579920

25. Marwaha S, Parsons N, Flanagan S, Broome M. The prevalence and clinical associations of mood instability in adults living in England: results from the Adult Psychiatric Morbidity Survey 2007. *Psychiatry Res* (2013) 205(3):262–8. doi:10.1016/j.psychres.2012.09.036

26. Bowen R, Balbuena L, Leuschen C, Baetz M. Mood instability is the distinctive feature of neuroticism. Results from the British Health and Lifestyle Study (HALS). *Pers Indiv Differ* (2012) 53(7):896–900. doi:10.1016/j.paid.2012.07.003

27. Kendler KS, Kuhn J, Prescott CA. The interrelationship of neuroticism, sex, and stressful life events in the prediction of episodes of major depression. *Am J Psychiatry* (2004) 161(4):631–6. doi:10.1176/appi.ajp.161.4.631

28. World Health Organization. *The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines*. Geneva: World Health Organization (1992). xii, 362 p.

29. Fava M, Hwang I, Rush AJ, Sampson N, Walters EE, Kessler RC. The importance of irritability as a symptom of major depressive disorder: results from the National Comorbidity Survey Replication. *Mol Psychiatry* (2010) 15(8):856–67. doi:10.1038/mp.2009.20

30. Judd LL, Schettler PJ, Coryell W, Akiskal HS, Fiedorowicz JG. Overt irritability/anger in unipolar major depressive episodes past and current characteristics and implications for long-term course. *JAMA Psychiatry* (2013) 70(11):1171–80. doi:10.1001/jamapsychiatry.2013.1957

31. Stringaris A, Maughan B, Copeland WS, Costello EJ, Angold A. Irritable mood as a symptom of depression in youth: prevalence, developmental, and clinical correlates in the Great Smoky Mountains Study. *J Am Acad Child Adolesc Psychiatry* (2013) 52(8):831–40. doi:10.1016/j.jaac.2013.05.017

32. Singleton N, Lewis G. *Better or Worse: A Longitudinal Study of the Mental Health of Adults Living in Private Households in Great Britain*. London: Office of National Statistics, UK (2003).

33. Lewis G, Pelosi AJ, Araya R, Dunn G. Measuring psychiatric disorder in the community: a standardized assessment for use by lay interviewers. *Psychol Med* (1992) 22(2):465–86. doi:10.1017/S0033291700030415

34. Uher R, Payne JL, Pavlova B, Perlis RH. Major depressive disorder in DSM-5: implications for clinical practice and research of changes from DSM-IV. *Depress Anxiety* (2014) 31(6):459–71. doi:10.1002/da.22217

35. Moller HJ. The consequences of DSM-5 for psychiatric diagnosis and psycho-pharmacotherapy. *Int J Psychiatry Clin Pract* (2014) 18(2):78–85. doi:10.3109/13651501.2014.890228

36. Moller HJ, Bandelow B, Bauer M, Hampel H, Herpertz SC, Soyka M, et al. DSM-5 reviewed from different angles: goal attainment, rationality, use of evidence, consequences – part 1: general aspects and paradigmatic discussion of depressive disorders. *Eur Arch Psychiatry Clin Neurosci* (2015) 265(1):5–18. doi:10.1007/s00406-014-0520-x

37. First MB. *User's Guide for the Structured Clinical Interview for DSM-IV Axis II Personality Disorders: SCID-II*. Washington, DC: American Psychiatric Press (1997). 91 p.

38. Duncan PW, Bode RK, Lai SM, Perera S, Neuroprotection GA. Rasch analysis of a new stroke-specific outcome scale: the Stroke Impact Scale. *Arch Phys Med Rehabil* (2003) 84(7):950–63. doi:10.1016/S0003-9993(03)00035-2

39. Pesudovs K, Noble BA. Improving subjective scaling of pain using Rasch analysis. *J Pain* (2005) 6(9):630–6. doi:10.1016/j.jpain.2005.04.001

40. Bond TG, Fox CM. *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*. 2nd ed. Mahwah, NJ: Lawrence Erlbaum Associates Publishers (2007). 340 p.

41. Elhan AH, Kutlay S, Kucukdeveci AA, Cotuk C, Ozturk G, Tesio L, et al. Psychometric properties of the Mini-Mental state examination in patients with acquired brain injury in Turkey. *J Rehabil Med* (2005) 37(5):306–11. doi:10.1080/16501970510037573

42. Horn JL. A rationale and test for the number of factors in factor-analysis. *Psychometrika* (1965) 30(2):179–85. doi:10.1007/BF02289447

43. Dinno A. *paran: Horn's Test of Principal Components/Factors*. R package version 1.5.1 (2012). Available from: https://CRAN.R-project.org/package=paran

44. Smith RM, Schumacker RE, Bush MJ. Using item mean squares to evaluate fit to the Rasch model. *J Outcome Meas* (1998) 2(1):66–78.

45. Chen JS, Torre J, Zhang Z. Relative and absolute fit evaluation in cognitive diagnosis modeling. *J Educ Meas* (2013) 50(2):123–40. doi:10.1111/j.1745-3984.2012.00185.x

46. Chen WH, Thissen D. Local dependence indexes for item pairs: using item response theory. *J Educ Behav Stat* (1997) 22(3):265–89. doi:10.2307/1165285

47. Penfield RD, Algina J. Applying the Liu-Agresti estimator of the cumulative common odds ratio to DIF detection in polytomous items. *J Educ Meas* (2003) 40(4):353–70. doi:10.1111/j.1745-3984.2003.tb01151.x

48. Penfield RD. DIFAS: differential item functioning analysis system. *Appl Psych Meas* (2005) 29(2):150–1. doi:10.1177/0146621603260686

49. Borsboom D, Cramer AO. Network analysis: an integrative approach to the structure of psychopathology. *Annu Rev Clin Psychol* (2013) 9:91–121. doi:10.1146/annurev-clinpsy-050212-185608

50. Crane PK, Gibbons LE, Jolley L, van Belle G. Differential item function-ing analysis with ordinal logistic regression techniques. DIFdetect and difwithpar. *Med Care* (2006) 44(11 Suppl 3):S115–23. doi:10.1097/01.mlr.0000245183.28384.ed

51. Kendel F, Wirtz M, Dunkel A, Lehmkuhl E, Hetzer R, Regitz-Zagrosek V. Screening for depression: Rasch analysis of the dimensional structure of the PHQ-9 and the HADS-D. *J Affect Disord* (2010) 122(3):241–6. doi:10.1016/j.jad.2009.07.004

52. Lerdal A, Kottorp A, Gay CL, Grov EK, Lee KA. Rasch analysis of the Beck Depression Inventory-II in stroke survivors: a cross-sectional study. *J Affect Disord* (2014) 158:48–52. doi:10.1016/j.jad.2014.01.013

53. Olino TM, Yu L, Klein DN, Rohde P, Seeley JR, Pilkonis PA, et al. Measuring depression using item response theory: an examination of three measures of depressive symptomatology. *Int J Methods Psychiatr Res* (2012) 21(1):76–85. doi:10.1002/mpr.1348

54. Koukopoulos A, Sani G, Koukopoulos AE, Manfredi G, Pacchiarotti I, Girardi P. Melancholia agitata and mixed depression. *Acta Psychiatr Scand Suppl* (2007) 115(433):50–7. doi:10.1111/j.1600-0447.2007.00963.x

55. Kendler KS, Gardner CO Jr. Boundaries of major depression: an evaluation of DSM-IV criteria. *Am J Psychiatry* (1998) 155(2):172–7.

56. Marwaha S, Parsons N, Broome M. Mood instability, mental illness and sui-cidal ideas: results from a household survey. *Soc Psychiatry Psychiatr Epidemiol* (2013) 48(9):1431–7. doi:10.1007/s00127-013-0653-7

57. Wessman AE, Ricks DF. *Mood and Personality*. New York: Holt (1966). xi,317 p.

58. Bowen R, Baetz M, Hawkes J, Bowen A. Mood variability in anxiety disorders. *J Affect Disord* (2006) 91(2–3):165–70. doi:10.1016/j.jad.2005.12.050

59. Palmier-Claus JE, Taylor PJ, Varese F, Pratt D. Does unstable mood increase risk of suicide? Theory, research and practice. *J Affect Disord* (2012) 143(1–3):5–15. doi:10.1016/j.jad.2012.05.030

60. Broome MR, He Z, Iftikhar M, Eyden J, Marwaha S. Neurobiological and behavioural studies of affective instability in clinical populations: a systematic review. *Neurosci Biobehav Rev* (2015) 51:243–54. doi:10.1016/j.neubiorev.2015.01.021

61. Tremblay PF, Ewart LA. The Buss and Perry Aggression Questionnaire and its relations to values, the Big Five, provoking hypothetical situations, alcohol consumption patterns, and alcohol expectancies. *Pers Indiv Differ* (2005) 38(2):337–46. doi:10.1016/j.paid.2004.04.012

62. Koukopoulos A, Sani G. DSM-5 criteria for depression with mixed features: a farewell to mixed depression. *Acta Psychiatr Scand* (2014) 129(1):4–16. doi:10.1111/acps.12140

63. Perlis RH, Fraguas R, Fava M, Trivedi MH, Luther JF, Wisniewski SR, et al. Prevalence and clinical correlates of irritability in major depressive disorder: a preliminary report from the sequenced treatment alternatives to relieve depression study. *J Clin Psychiatry* (2005) 66(2):159–66;quiz 47, 273-4. doi:10.4088/JCP.v66n0202

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

65. Kovess-Masfety V, Alonso J, Angermeyer M, Bromet E, de Girolamo G, de Jonge P, et al. Irritable mood in adult major depressive disorder: results from the world mental health surveys. *Depress Anxiety* (2013) 30(4):395–406. doi:10.1002/da.22033

66. Angst J, Gamma A, Benazzi F, Ajdacic V, Rossler W. Does psychomotor agitation in major depressive episodes indicate bipolarity? Evidence from the Zurich study. *Eur Arch Psychiatry Clin Neurosci* (2009) 259(1):55–63. doi:10.1007/s00406-008-0834-7

67. Thase ME. Recognition and diagnosis of atypical depression. *J Clin Psychiatry* (2007) 68(Suppl 8):11–6.

68. Zimmerman M, Galione JN, Chelminski I, McGlinchey JB, Young D, Dalrymple K, et al. A simpler definition of major depressive disorder. *Psychol Med* (2010) 40(3):451–7. doi:10.1017/S0033291709990572

69. Decker HS. *The Making of DSM-III: A Diagnostic Manual's Conquest of American Psychiatry*. New York: Oxford University Press (2013). xxii, 443 p.

70. Kendler KS. Major depression and generalised anxiety disorder. Same genes, (partly) different environments – revisited. *Brit J Psychiatry* (1996) 168:68–75.

71. Krueger RF. The structure of common mental disorders. *Arch Gen Psychiatry* (1999) 56(10):921–6. doi:10.1001/archpsyc.56.10.921

72. Preisig M, Merikangas KR, Angst J. Clinical significance and comorbidity of subthreshold depression and anxiety in the community. *Acta Psychiatr Scand* (2001) 104(2):96–103. doi:10.1034/j.1600-0447.2001.00284.x

73. Spiers N, Bebbington P, McManus S, Brugha TS, Jenkins R, Meltzer H. Age and birth cohort differences in the prevalence of common mental disorder in England: National Psychiatric Morbidity Surveys 1993-2007. *Br J Psychiatry* (2011) 198(6):479–84. doi:10.1192/bjp.bp.110.084269

74. Forkmann T, Boecker M, Wirtz M, Glaesmer H, Brahler E, Norra C, et al. Validation of the Rasch-based depression screening in a large scale German general population sample. *Health Qual Life Outcomes* (2010) 8:105. doi:10.1186/1477-7525-8-105

75. Siegert RJ, Tennant A, Turner-Stokes L. Rasch analysis of the Beck Depression Inventory-II in a neurological rehabilitation sample. *Disabil Rehabil* (2010) 32(1):8–17. doi:10.3109/09638280902971398

76. Kapur S, Phillips AG, Insel TR. Why has it taken so long for biological psy-chiatry to develop clinical tests and what to do about it? *Mol Psychiatry* (2012) 17(12):1174–9. doi:10.1038/mp.2012.105

77. Sinyor M, Schaffer A, Levitt A. The sequenced treatment alternatives to relieve depression (STAR*D) trial: a review. *Can J Psychiatry* (2010) 55(3):126–35.

78. Ben-Zeev D, Young MA. Accuracy of hospitalized depressed patients' and healthy controls' retrospective symptom reports: an experience sampling study. *J Nerv Ment Dis* (2010) 198(4):280–5. doi:10.1097/NMD.0b013e3181d6141f

79. Sutin AR, Terracciano A, Milaneschi Y, An Y, Ferrucci L, Zonderman AB. The trajectory of depressive symptoms across the adult life span. *JAMA Psychiatry* (2013) 70(8):803–11. doi:10.1001/jamapsychiatry.2013.193

80. Martin LA, Neighbors HW, Griffith DM. The experience of symptoms of depression in men vs women: analysis of the National Comorbidity Survey Replication. *JAMA Psychiatry* (2013) 70(10):1100–6. doi:10.1001/jamapsychiatry.2013.1985

81. Broome MR, Saunders KE, Harrison PJ, Marwaha S. Mood instability: significance, definition and measurement. *Br J Psychiatry* (2015) 207(4):283–5. doi:10.1192/bjp.bp.114.158543

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.