Development and Construct Validity of the Work Instability Scale for People With Common Mental Disorders in a Sample of Depressed and Anxious Workers: A Rasch Analysis

Louise Danielsson[1,2] (iD), Robin Fornazar[3,4], Kristina Holmgren[2], Åsa Lundgren Nilsson[5] and Gunnel Hensing[4]

[1]Research Unit, Angered Hospital, Gothenburg, Sweden. [2]Department of Health and Rehabilitation, Institute of Neuroscience and Physiology, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden. [3]Administration for Allocation of Social Welfare, Gothenburg, Sweden. [4]Department of Public Health and Community Medicine, School of Public Health, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden. [5]Department of Clinical Neuroscience, Institute of Neuroscience and Physiology, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden.

ABSTRACT

BACKGROUND: Sick leave due to common mental disorders, encompassing depression and anxiety disorders, is high. Capturing early signs of reduced function could aid adjustments of work tasks and environment and, thus, endorse a pro-active approach to occupational and health care interventions to prevent long-term sick-leave spells. However, few measurements exist to identify early signs of imbalance, and none that is illness-specific. The aim of this study was to develop a work instability scale for people with common mental disorders and to test the fundamental psychometric properties of the scale.

METHODS: Participants were working adults 18-65 years old with depression or anxiety. The scale development started with qualitative interviews (n = 27) which informed the drafting of a dichotomous, self-report questionnaire. Cognitive debriefing (n = 12) was used to check face validity and modify the draft. Internal construct validity of the draft was tested using Rasch analysis (n = 128). The work ability index was used as a comparator measure.

RESULTS: The initial 63-item draft showed poor fit to Rasch model expectations. Items displaying poor fit or local response dependency were stepwise removed, resulting in a unidimensional 34-item scale fitting the model expectations, and with no differential item functioning. Person-item threshold distribution showed that the scale is better suited to measure low to moderate work instability, than to measure high instability. Correlations between the newly developed scale and the work ability index showed a significant, moderately strong correlation.

CONCLUSIONS: In the initial target sample, the 34-item scale showed acceptable fundamental properties and internal construct validity. Further validation of the scale in a larger sample, including tests for external validity, is warranted.

KEYWORDS: Depression, anxiety, mental health, Rasch analysis, psychometrics, work capacity

Introduction

Depression and anxiety disorders, sometimes referred to as "common mental disorders," are leading diagnoses and reasons for sick leave.[1,2] Although there are some variations regarding the definition of common mental disorders (CMDs), the concept refers to highly prevalent mental health problems, including a range of symptoms (from mild to severe) and duration (from months to years), which affect the mood or feelings of the affected persons.[3] For this study, we use the concept as described by the World Health Organization[3] and the National Collaborating Centre for Mental Health (UK),[4] including the 2 diagnostic categories depressive disorders (including sub-threshold states) and anxiety disorders (including generalized anxiety disorder, panic disorder, social anxiety disorder, phobias, obsessive-compulsive disorder, and post-traumatic stress disorder).

Although the ability to work is often affected,[5] most people with CMDs struggle to work despite their illness,[6] trying to cope at their job on a day-to-day basis.[7] The day-to-day adaptation and wish to keep up a well-functioning appearance can be tiring and add to the burden of the depressive and anxiety symptoms,[8] such as fatigue, worry, and sleeping problems. This struggle to keep working can turn into a downward spiral where the worker cannot recover sufficiently,[9] constituting a risk for future sick-leave.[10] Moreover, as symptoms and distress at work occur long before sick leave[11-13] and vary over time,[14]

there can be great fluctuation in the worker's experience of managing their job. To support workers' mental health and to prevent long-term sick leave, early signs of negative change in the worker's well-being and work functioning need to be identified.[15,16] Acknowledging negative change might facilitate early and timely interventions,[17] such as workplace adjustments and eventual needs for vocational rehabilitation. A pro-active approach has been encouraged,[17] but there is paucity of structured ways to aid the identification of individuals at risk.

Sustained work participation is beneficial for people with CMDs,[18] calling for a need for support while working to promote mental health and functioning and to prevent, in particular, long sick-leave spells. Preventive support from occupational and primary health care and rehabilitation is important also in the large group of workers with mild symptoms, to avoid an increase of symptoms associated with higher risk of sickness absence.[19] Early support at the workplace is also important, calling for awareness among managers about depression and anxiety, and about adjustments to support the worker. Timely interventions to restore imbalance between resources (individual and organizational) and job demands are challenging,[20] and it should thus be essential to recognize when workers are on the verge of declining health and function.

In a previous qualitative study, depressed and anxious workers' descriptions were interpreted as a processual experience of "working in dissonance," a disturbed sense of the ordinary workflow involving changes related to time and energy, bodily discomfort, awareness, agency and meaning.[21] The negative experiences were often disregarded or denied and, for some, increased until they could not work anymore. Possibly, warning signs might be captured in dialogue with the worker, to guide adjustment and early support. This pro-active dialogue may take place in clinical settings such as primary or occupational health care, but could also be part of a manager-worker dialogue.

Work ability is a multifactorial concept, dynamic over time and depending on the interaction of several dimensions (ie, personal attributes, work, health care, and community systems) that influence each other through different structural levels (individual, organizational, and societal). Moreover, a person's general ability to work and their situational ability in the specific job environment can be vastly different. In line with this, the International Classification of Functioning, Disability and Health (ICF) addresses the complexity of work ability by considering both personal factors (ie, education, skills, coping) and the environment (ie, possibility to adapt tasks and the physical/psychosocial environment). Similar assumptions underpin other established models, such as the work ability model[22] and the person-environmental-occupational (PEO) model. These frameworks all suggest interplay between individual resources, occupational tasks, and physical/psychosocial environment and have inspired the development of several validated scales. Commonly used are the generic Work Ability Index (WAI)[23]; the Work Role Functioning Questionnaire[24] (WRFQ), with subdomains related to scheduling/output, flexibility, physical, mental, and social demands; and the Mini-ICF-APP,[25] which was developed to measure limitations of capacities in the context of mental disorders.[26]

Although these scales capture the crucial interaction between person capacities and job demands, we saw a need for an additional concept and scale for 3 reasons. First, we lacked a tool targeting the early, experience-based signs of change in job functioning that the depressed or anxious worker may encounter. Qualitative studies have shown that these workers struggle long and hard to compensate for difficulties at work.[7,9] We argue that there is an important window for increased awareness, support, and adjustments at this early phase, before capacities and job performance are visibly reduced. Rather, this phase starts when "things just don't seem right" for the worker—between functioning well and working while ill (presenteeism). These changes in the worker's job experience are important warning signs. Using the analogy of traffic lights, this in-between phase represents the potentially reversible shift from a "green zone" to an "orange zone," prior to a "red zone" of sick leave and disability.

Second, based on our understanding of the PEO model[27] and our previous qualitative studies,[8,9,13,21] we assume this phase to be a process of instability arising in the dynamics between individual (including severity of symptoms), environmental, and occupational dimensions. How depressed and anxious workers function at work will change, continually, depending on fluctuations in symptoms, but also depending on job tasks and work environment. The change is not always a declining process until the worker cannot work any longer—sometimes negative change stabilizes again and the worker can continue working, often with some adaptation. This calls for a process-oriented concept rather than a state of function. The instability can be stabilized if one or several of the 3 dimensions (person capacities, environmental, and occupational demands) are supported or adjusted. For example, stabilization may occur if the individual receives psychological or medical treatments from health care to alleviate symptoms or develop coping strategies, but stabilization may also occur if the worker is allowed different tasks for a period of time, or if there are supportive alterations in the work environment. However, for such adjustments and interventions to take place, it is of issue to recognize and capture early signs of instability.

Third, existing scales are theory-based, which provides a solid methodological foundation, but the item phrasing might not mirror the lived experiences of workers with depression and anxiety disorders. To help workers become aware of and articulate eventual problems in their job situation, a tool with a bottom-up perspective, starting in the experiences of the target group, would be desirable. Ideally, such tool would reflect common experiences among depressed and anxious workers. It would also be constructed in a way that enables use in both clinical settings (patient-health professional encounters) and workplace settings (worker-manager dialogue).

To summarize, the idea of introducing the concept of work instability in the field of mental health is to suggest a shift from the individual state due to illness toward a changeable process. It adds a slightly different perspective where "stabilization" may occur following adjustments or individual/environmental interventions.

The concept of work instability was originally coined in a study on workers with rheumatoid arthritis,[28] defined as a potentially reversible state in which the consequences of a mismatch between an individual's functional abilities and the demands of their job can threaten continuing employment if not resolved.[28] During the last 15 years, several self-rating scales to measure work instability in different diagnoses and occupational settings have been developed.[29-32] The researchers involved in this work suggest that each diagnosis brings unique features to the scale as well as some common elements,[30] possibly reflecting core elements of the construct. Thus, there is a need to study characteristics of work instability separately, in different diagnoses, to ensure that the scale captures the relevant elements specific to that disorder.

A self-report scale measuring work instability would help to identify essential signs of reduced function prior to sick leave, or among workers prone to instability during a return-to-work phase. The assessment could also give structured support to health care or rehabilitation professionals in the dialogue with patients, and possibly to managers in their dialogue with workers, about their situation. It could also enable following up of adjustments or interventions. As, to our knowledge, there are no such existing tools, the aim of this study was to develop a work instability scale for people with common mental disorders (WIS-CMD) and to test its fundamental psychometric properties in an initial target sample.

Methods

The development of the WIS-CMD was inspired by a stepwise model used in previous work instability scales,[28,29,32] starting with qualitative data collection, followed by drafting a dichotomous (true/not true) self-report questionnaire validated in the target population, ie, workers with depression or anxiety disorder. The research group consisted researchers from psychiatric rehabilitation, social medicine, and public health. The study was approved by the regional ethical review board in Gothenburg (registration number 810-15). All participants gave their informed, documented consent before taking part in the study.

Participants

Participants were working adults 18-65 years old with depression or anxiety disorder (International Classification of Diseases, 10th revision,[33] codes F32–33, F41, and F43). To include subthreshold states, which are also part of CMDs[4] and for which the future scale might be particularly useful, we also included a number of participants experiencing depressive or anxiety symptoms without a clinical diagnosis. The participants for the qualitative phase, presented in detail elsewhere,[21] were recruited from primary care centers via public lectures at the city library, and through a patient organization. Participants for the face validity phase were recruited from primary care. Participants for the construct validity phase were recruited from primary care centers, via advertisements in local newspapers and through a patient organization.

Data collection

In-depth interviews were conducted with 27 workers with depression or anxiety disorders, focusing on their experiences of change at work.[21] Comprehensive qualitative analyses are presented elsewhere,[9,21] describing themes relating to energy/time, bodily discomfort, lack of awareness, agency, and meaning. Next, statements were extracted from the interviews to develop an item pool of 74 statements. We tried to keep close to the direct statements, aiming to use colloquial language. To receive feedback and to increase the clarity of the items, face validity was checked through a cognitive debriefing with a new sample of 12 workers (9 women, 3 men) with depression or anxiety disorders. They received the draft scale and were asked to complete it and comment on all items that were hard to understand and/or ambiguous and/or that did not make sense. The feedback resulted in revised phrasing of 7 items, and the removal of 11 items which the workers found to be already covered by similar items and therefore to be redundant.

We recruited a new sample (see Table 1) for the validity testing of the 63-item draft. Background questions, the WIS-CMD draft, the 5-item World Health Organization (WHO-5) Wellbeing Index,[34] and the comparator measure the WAI[23] were distributed to the sample using an electronic questionnaire pack, constructed using the esMaker survey tool (Entergate AB, Halmstad, Sweden).

The WHO-5 Wellbeing Index was chosen as a brief scale measuring the participants' psychological well-being.[34] The scale has been validated in people with depression and anxiety, showing satisfactory psychometric qualities.[34] Scale categories are defined as different levels of psychological well-being, where 0 = minimal, 1-25 = poor, 26-50 = fair, 51-75 = good, 76-99 = very good, and 100 = excellent well-being.

The WAI was chosen as a well-established and validated assessment measuring the participants' self-reported work ability.[22,23] The 7 dimensions of the WAI are derived as the sum of 10 items, resulting in a total sum score of 7-49, with present work ability classified as follows: 7-27 = poor; 28-36 = moderate; 37-43 = good; and 44-59 = excellent.

Neither the WRFQ[24] nor the Mini-ICF-APP[25] had been validated in Swedish, and we therefore decided not to use them for comparator measures.

Table 1. Characteristics of participants included in the Rasch analysis, n = 126 (98% of the total sample).

CHARACTERISTIC	N
Age	
18-29 years	25
30-39 years	36
40-49 years	28
50-67 years	37
Sex	
Women/men	102/24
Highest education level	
Elementary education	0
High school	35
University <3 years	18
University ⩾3 years	73
Marital status	
Married/cohabiting	60
Single	47
Other	19
Children living at home	
Yes/no	40/86
Current sick leave[a]	
Not on sick leave	74
25% sick leave	13
50% sick leave	17
75% sick leave	5
100% sick leave	17
Diagnosis[b]	
Depression	66
Anxiety disorder	62
Exhaustion syndrome	32
Comorbid mental health problem[c]	11
No diagnosis	27
Work ability	
Work ability index, mean score (SD)	28.5 (7.0)
Poor work ability: 7-27	49
Moderate work ability: 28-36	50
Good work ability 37-43	15

(Continued)

Table 1. (Continued)

CHARACTERISTIC	N
Excellent work ability 44-59	1
Psychological well-being	
WHO-5 Mental well-being, mean score (SD)	33.5 (18.3)
Minimal to poor well-being: 1-25	49
Fair well-being 26-50	57
Good well-being: 51-75	19
Very good to excellent: 76-100	2

Abbreviations: SD, standard deviation; WHO-5, 5-item World Health Organization Mental Wellbeing Index.
[a]In Sweden, workers can be on sick leave 25%, 50%, 75%, or 100% of their ordinary work time.
[b]Diagnosis confirmed using the Mini-International Neuropsychiatric Interview (n = 31), or self-reported (n = 95) in answer to the question, "Have you been told by a medical doctor or psychologist that you have one of the following diagnoses?"
[c]Self-reported: personality disorder, stress reaction, social phobia, post-traumatic stress, obsessive-compulsive disorder, attention deficit hyperactivity disorder.

Psychometric analyses

Rasch analysis was used to test the internal construct validity of the scale and to guide item reduction. Rasch analysis, which is increasingly used in health research on patient-reported outcomes,[35] offers a method of investigating whether the required properties of a scale are confirmed or not, before arithmetic operations can be undertaken. The Rasch model is a unidimensional model with 2 main assertions, namely, (1) the easier an item, the more likely it will be affirmed, and (2) the "more" of the construct a patient has, the more likely they will affirm an item.[36] Here, this means that for the WIS-CMD, an easy item would be affirmed by many workers with depression or anxiety disorders, and also those who cope reasonably well at work, while a difficult item would be affirmed only by those who experience substantial instability. The more work instability the worker has, the more likely they will affirm an item. The following Rasch requirements were checked in the analysis:

(A) *Fit statistics and fit residuals*. Essentially, Rasch analysis looks at the deviation of the observed data from the model expectation, the fit to model. Results are reported as a series of χ^2 statistics and fit residuals demonstrating the discrepancy between expected and observed data. Where an item fits the model, the χ^2 probability is nonsignificant (Bonferroni-corrected *P* value) and fit residuals should be within a specified range (eg, ±2.5).[37]

(B) *Internal reliability*. Internal reliability was analyzed using the person separation index (PSI), a reliability statistic that is interpreted in the same way as Cronbach's α, suggesting that a minimum PSI value

of 0.7 is required for group use and 0.85 for use at the individual level.[38]

(C) *Item bias.* Differential item functioning was analyzed to check whether items worked in the same way across groups of patients regarding age, sex, and education.

(D) *Local dependency.* Local independence means that the entire correlation between the items has to be captured by the underlying construct (ie, the latent trait, here, work instability). By calculating the mean correlation between item residuals, we assessed each pair of items for local response dependency. Should there be correlation values above 0.3,[39] this would indicate local dependence, which is a concern as it may reflect multidimensionality or response dependency (that a response to one given item is not independent from the response to another item).

(E) *Unidimensionality.* The Rasch model assumes a single, unidimensional construct, a prerequisite to adding items into a sum score. Unidimensionality means that a single construct (eg, work instability in people with depression and anxiety disorders) is being measured by a set of items. In a principal components analysis of the draft scale, 2 subsets of items with the highest and lowest loadings were created. The person estimates from these 2 subsets of items were subjected to a series of *t* tests. A nonsignificant difference between the 2 person estimates would support the unidimensionality of the scale. The percentage of tests outside the –1.96 to 1.96 range should not exceed 5%.[38]

Statistical notes, software, and sample size

In line with previous work instability scales, 200 was the intended sample size to ensure stable calibrations. Basic requirements for Rasch analyses suggest that a sample size of 109 participants is needed to estimate item difficulty, with $\alpha = 0.01$, to within ± 0.7 logits.[40] To detect item bias ($\alpha = 0.01$), a difference of 0.5 standard deviations (SDs) within the residuals for any 2 groups with β of 0.20, a minimum sample of 120 participants would be required.

Test-retest data were collected in a subsample of the participants, 2 weeks after the first measurement. To achieve a larger sample in the construct validity test, we decided to follow the principles of stacking these data[41,42] with the original data in an anchoring reanalysis within the Rasch framework.

Descriptive statistics were used to analyze characteristics of the study participants, using IBM SPSS version 22.0 (IBM Corp, Armonk, NY, USA). For the Rasch analysis, we used the RUMM2030 software (RUMM Laboratory Pty Ltd, Perth, Australia). Bonferroni adjustments were used in the Rasch analysis, correcting for multiple testing. The correlation between WIS-CMD data and the comparator measure WAI was analyzed using bivariate correlations in SPSS.

Results

During April 2015 to December 2017, 128 participants agreed to take part in the construct validity phase. About one-third of these were recruited from primary care centers, one-third replied to advertisements, and one-third were recruited from a patient organization. Table 1 presents characteristics for the 126 participants (98%) who completed the questionnaire. They had a mean age of 41 years (SD = 12) and worked in different work sectors, including education, health and social welfare, the manufacturing industry, construction, finance, technology, communication and retail. Most (81%) were women.

In the Rasch analysis of the 63-item draft, the initial fit to model was poor (Table 2), with a significant item-trait interaction statistic ($\chi^2 = 299.2$, $P < .001$). The PSI was excellent, at 0.91. Person-item threshold distribution (Figure 1) indicated that relatively few of the items covered high work instability. This means that the scale is better suited to correctly measure work instability among workers with low to moderate work instability, while there is more uncertainty in measuring workers with high instability. None of the items displayed any differential item functioning based on age, sex, or education. This means that, given the same level of work instability, participants responded to the items in a similar way regardless of whether they were younger or older, woman or man, or had low or high education levels.

We found that 3 items ("I feel wound up," "I take fewer breaks than usual," and "Things run better at work than at home") had high fit residuals (between 2.28 and 5.72) and significant χ^2 probability ($P < .0005$). In the first step of item reduction, we therefore removed these 3 items and another 7 items (Figure 2) displaying major local response dependency.

Statistics of the 53-item draft showed that the scale had improved but still did not fit model expectations (Table 2). Local response dependency, the location of each specific item, and their qualitative meaning guided further item reduction, stepwise removing items (Figure 2) toward a 46- and, finally, a 34-item draft. Before removal, we discussed the meaning of each item in relation to theory generated from the qualitative work, ensuring that relevant thematic meaning were kept among the remaining items.

The 34-item draft of the scale showed good overall fit to the Rasch model, item residual mean –0.09 (SD = 0.88); person residual mean –0.05 (SD = 0.77); χ^2 probability .34. The scale had a PSI reliability of 0.84, indicating that the scale worked well to separate the persons. There was no differential item functioning or residual correlation above 0.3. Unidimensionality was acceptable, displaying 4.8% significant *t* tests (just below the expected limit of <5%), with a lower bound confidence interval of 1.8% (below the expected limit of <5%). The person-item distribution is presented in Figure 3. The 34-item parameters and fit statistics are presented in Table 3.

Finally, we used the stacked data, including both the original validity data (n = 126) and retest data (n = 76), resulting in a

Table 2. Rasch analysis of the Work Instability Scale for common mental disorders at the different stages of development.

SCALE STAGE	N	RESIDUAL-BASED FIT				χ² FIT		RELIABILITY	UNIDIMENSIONALITY	
		ITEM RESIDUAL MEAN	SD	PERSON RESIDUAL MEAN	SD	VALUE	P VALUE	PSI	T TEST %	% LBCI
63-item draft	126	−0.132	1.43	−0.022	0.97	299.2	<.001	0.91	17.5	13.7
53-item draft	126	−0.113	1.15	−0.036	0.92	147.1	.005	0.90	11.1	7.3
46-item draft	126	−0.091	1.06	−0.039	0.85	102.9	.206	0.88	14.3	4.1
34-item draft	126	−0.094	0.88	−0.053	0.77	72.1	.344	0.84	7.9	4.9
Anchored analysis	203	−0.103	1.18	−0.025	0.77	90.4	.04	0.84	6.9	3.9
Ideal values		0.0	1	0.0	1		>.002[a]	>0.70	<5.0	<5.0

Abbreviations: LBCI, lower bound confidence interval; PSI, person separation index; SD, standard deviation.
[a]Bonferroni-corrected value of 0.05 will vary by analysis, adjusted for the number of tests; this value refers to the final scale stage.
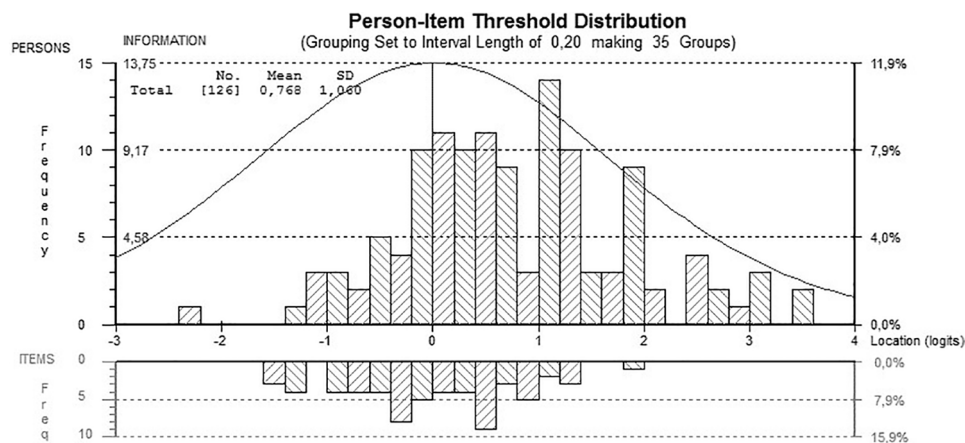
Figure 1. Person-item threshold distribution plots from the Rasch analysis of the initial 63-item work instability scale for people with depression and anxiety disorders.
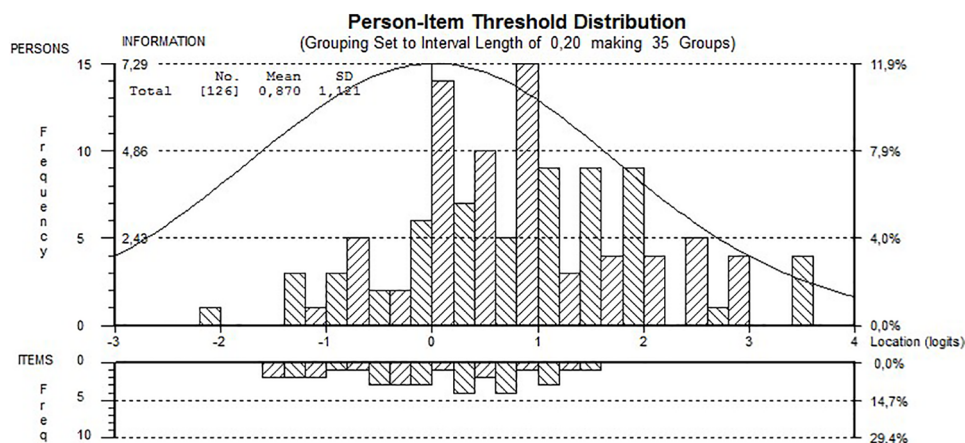
Figure 2. Items removed in the stepwise analysis.

combined sample (n = 203). In this anchored analysis, the 34-item scale still showed fit to model, a similar PSI value, and acceptable unidimensionality (Table 2). There was still no indication of differential item functioning or residual correlations above 0.3.

The 34-item WIS-CMD had a moderately strong correlation with the WAI, significant at the 0.01 level. The Spearman coefficient was −0.59 (*P* < .001). This indicates that poor work ability, as measured by the WAI, correlates with high work instability as measured by the WIS-CMD. The moderate
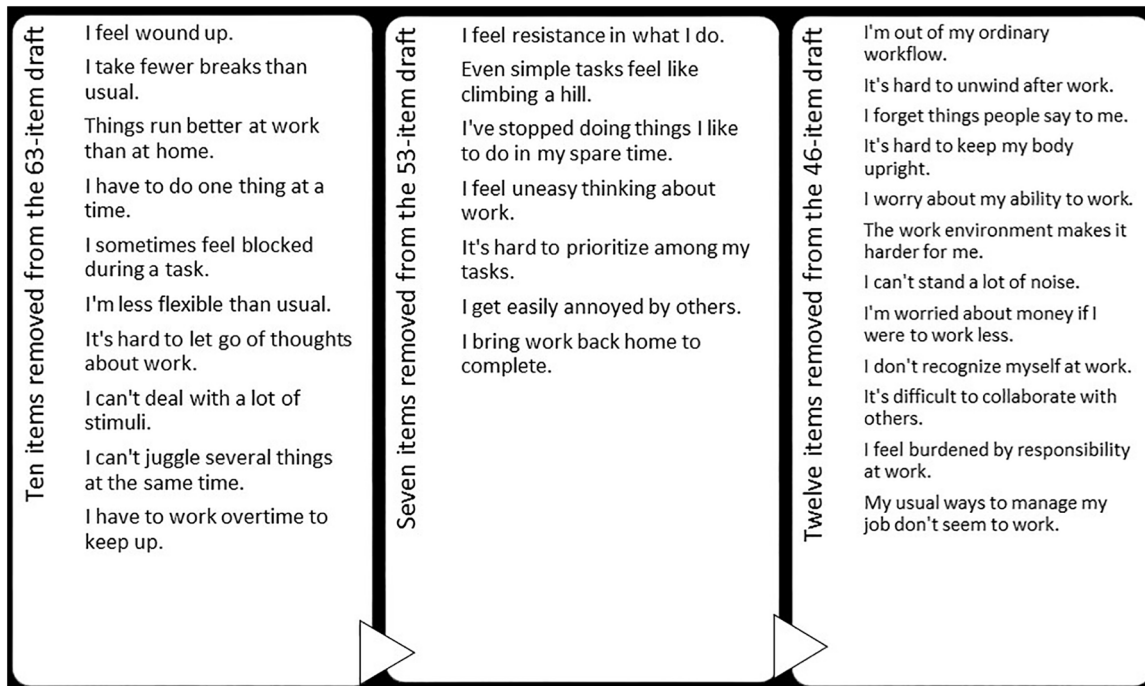
Ten items removed from the 63-item draft

I feel wound up.

I take fewer breaks than usual.

Things run better at work than at home.

I have to do one thing at a time.

I sometimes feel blocked during a task.

I'm less flexible than usual.

It's hard to let go of thoughts about work.

I can't deal with a lot of stimuli.

I can't juggle several things at the same time.

I have to work overtime to keep up.

Seven items removed from the 53-item draft

I feel resistance in what I do.

Even simple tasks feel like climbing a hill.

I've stopped doing things I like to do in my spare time.

I feel uneasy thinking about work.

It's hard to prioritize among my tasks.

I get easily annoyed by others.

I bring work back home to complete.

Twelve items removed from the 46-item draft

I'm out of my ordinary workflow.

It's hard to unwind after work.

I forget things people say to me.

It's hard to keep my body upright.

I worry about my ability to work.

The work environment makes it harder for me.

I can't stand a lot of noise.

I'm worried about money if I were to work less.

I don't recognize myself at work.

It's difficult to collaborate with others.

I feel burdened by responsibility at work.

My usual ways to manage my job don't seem to work.

Figure 3. Person-item threshold distribution plots from the Rasch analysis of the revised 34-item work instability scale for people with depression and anxiety disorders.

correlation suggests that the 2 scales measure similar, but not identical, constructs.

Discussion

This study developed and validated a scale to capture work instability in people with CMDs, resulting in a 34-item scale with experience-based statements, and true/not true responses. Grounded in lived experiences and with acceptable internal construct validity, the new scale provides a first step toward a psychometrically sound self-report measurement of work instability among depressed and anxious workers.

The targeting of the 34 items (Figure 3) showed satisfactory coverage of easy to moderately difficult items: ie, the present scale is constructed to measure workers with some, but not highly pronounced problems at work. This is important as our goal when developing the scale was to capture early signs of declining work capacity. As indicated by the skewed distribution between persons and items visualized in Figure 3, there was a lack of item thresholds at the person locations at the positive end of the vertical axis. This suggests that participants who affirmed many statements (ie, who had high instability, close to "inability," were not reliably measured by the present scale. To increase measurement reliability also among workers with high instability, items reflecting experiences closer to "inability" would need to be added to the scale. Additional items may be created by returning to the interview transcripts and searching for experiences of high instability, or by conducting more interviews with workers on the verge of taking sick leave.

Drawing on these results, we suggest that the main application of the present WIS-CMD is in the early phase of imbalance at the job. This phase, if identified, is important from a preventive perspective, as it may give possibilities to make adjustments while working.[19,20] The scale provides opportunities to identify imbalance rather than inability, and may thus be helpful in taking early measures to avoid sickness absence and its negative consequences.[17] More research is warranted to test the scale but also to further investigate the process of interconnection between depressive and anxiety symptoms and related changes in functioning.

The ordering of items (see item descriptors in Table 3) suggests that the lowest threshold items, ie, the earliest signs of work instability, relate to body discomfort and loss of energy, affecting work and life outside work. This is coherent with research suggesting that somatic complaints such as nonspecific musculoskeletal pain and fatigue are common early manifestations of mood disorders.[43] The mid-threshold items relate to social withdrawal, emotional and cognitive dysfunction, in line with more pronounced symptoms of depression and anxiety. The highest threshold items in the present scale reflect loss of meaning at the job, and a more visible impression of withdrawal and dysfunction ("Others tell me that I don't seem well"). The highest threshold item ("I could collapse any day now") reflects the experience of that work is barely possible, and seems like a logic accentuation of the early fatigue expressed in the low-threshold items ("I feel very tired after work").

The content of affirmed items may give an indication of what aspects of the job are problematic for the individual worker. We suggest that the scale can be a tool primarily for occupational and primary health care professionals,[44] but possibly also a screening tool at the workplace. It is, however,

Table 3. Item parameters and fit statistics for the 34-item Work Instability Scale for common mental disorders.

ITEM	ITEM DESCRIPTOR	LOCATION (ITEM DIFFICULTY EXPRESSED IN LOGITS)	STANDARD ERROR OF THE DIFFICULTY ESTIMATE	FIT RESIDUAL	χ^2	P^a
35	Body doesn't feel right	−1.502	0.288	−0.977	2.136	.344
12	More tired than usual	−1.443	0.283	0.355	0.921	.631
13	Need to shield myself off	−1.284	0.271	−0.061	0.914	.633
18	Very tired after work	−1.231	0.267	0.108	2.098	.350
33	Things are getting to me	−1.168	0.262	−0.929	1.926	.382
51	Worry about disappointing	−1.141	0.260	0.412	5.059	.078
59	Up to me to solve things	−0.831	0.241	0.454	1.792	.408
39	Commit less than usual	−0.633	0.231	−1.021	3.583	.167
61	Varies a lot at work	−0.584	0.229	1.377	2.853	.240
34	Hard to control emotions	−0.515	0.226	−0.878	1.374	.503
20	Strain affects private life	−0.460	0.224	0.511	2.688	.261
2	Prefer routine tasks	−0.368	0.221	0.699	2.070	.355
45	Put up a front at work	−0.304	0.218	0.462	0.114	.944
62	Recovery doesn't help	−0.236	0.216	−0.532	1.739	.419
43	More quiet at work	−0.168	0.214	−0.899	1.278	.528
54	Push myself to work	−0.149	0.213	−0.146	0.356	.837
40	Don't enjoy socializing	−0.107	0.212	−0.349	0.060	.738
50	Wonder if I'm worse off	0.023	0.208	−0.069	0.603	.738
48	Afraid to lose control	0.201	0.205	−0.048	0.231	.891
26	More mistakes at work	0.266	0.204	−1.258	1.950	.377
32	Clumsy or butterfingered	0.266	0.204	0.939	3.570	.168
14	Hard to ask for help	0.369	0.202	2.162	8.034	.018
49	Sick leave means giving up	0.516	0.202	0.667	4.230	.121
31	Difficult make decisions	0.594	0.200	−1.349	1.384	.501
3	Trouble completing tasks	0.610	0.200	0.955	2.454	.293
23	Feel like "in a bubble"	0.644	0.200	0.473	0.483	.785
28	Not mentally present	0.743	0.201	−1.479	4.022	.134
53	Reaching goals mean nothing	0.751	0.200	0.047	1.154	.561
42	Misread people	0.930	0.200	−0.502	1.772	.412
56	Don't feel part of things	1.092	0.201	−1.438	1.428	.490
57	Tasks feel overpowering	1.098	0.201	−1.126	4.250	.119
52	Job not meaningful	1.133	0.201	−0.817	4.861	.088
46	Others tell me I'm not well	1.395	0.204	0.363	0.293	.864
37	Could collapse any day	1.495	0.206	0.688	0.410	.815

aA Bonferroni-corrected χ^2 P value of .002 was applied.

important to point out that the WIS-CMD is not meant to replace a more thorough clinical evaluation for workers in need of medical treatment or disability assessment.

Local response dependency was the main issue that needed to be resolved during the Rasch analysis, and that guided item reduction. The assumption of local independence is central, as violations lead to inflated estimates of reliability (PSI) and may reflect multidimensionality or response dependence.[45] We considered the item location and the qualitative meaning of the items, so that important findings from the qualitative work and its theoretical conceptualization would not be dismissed,[46] while reducing items due to local dependency. In the solution obtained with the 34-item scale, the values of residual correlations indicated independence between items, supporting that a unidimensional construct had been extracted.[38]

The WIS-CMD score had a moderately strong correlation with the WAI score. The WAI is a widely used generic assessment tool in epidemiological studies to measure work ability, including perceived prognosis of future ability.[23] Although the conceptualizations underpinning the 2 scales (work ability vs work instability) are likely connected, the WIS-CMD provides a different measure using items specific to depression and anxiety, developed to capture early signs of change and fluctuations experienced by the individual worker.

While relevant scales such as the WRFQ[24] or the Mini ICF-APP[25,26] are not yet validated in Swedish and were therefore not used for comparator measure, it will be important to compare these more established scales with the new WIS-CMD in future studies. The Mini-ICF-APP,[25] developed in the context of mental disorders, is based on the ICF and measures limitations of capacities related to adherence, planning/structuring of tasks, flexibility, competency, endurance, assertiveness, contact with others, group integration, intimate relationships, nonwork activities, self-care, mobility, and competence to judge and decide.[26] Compared with the items of the WIS-CMD, based on qualitative themes related to changes in energy/time at and outside work, bodily discomfort, awareness (including social support), agency, and meaning, there are similarities but also differences regarding content and scale construction. Moreover, a comparison and discussion concerning the underlying constructs and practical use of the respective scales will be useful to discern between scales. A recent Scandinavian cross-cultural adaptation[47] could aid future comparisons of WRFQ and WIS-CMD.

Strengths and limitations

Drafting of the scale using thematic content from in-depth interviews to retrieve relevant items was considered a strength. Moreover, cognitive debriefing enhanced face validity and justified the modification of items. The use of Rasch analysis enabled systematic item reduction to produce a unidimensional and logically constructed scale. The low dropout rate suggests that the electronic questionnaire format facilitated participation among the depressed and anxious workers.

Despite several recruitment sources, the sample size was small, which is the main limitation of the study. While the present sample size fulfills the basic requirements for Rasch analysis,[40] recent recommendations suggest sample sizes around 200, to reduce the risk of inappropriate modifications of the scale, such as needlessly discarded candidate items.[48] The small sample size increases the risk of less precise estimates (bigger standard errors), less powerful fit analysis, and less robust estimates (more likely that accidents in the data will distort them).[40] Our anchored analysis stacking data from the first round of data collection with the available retest data was an attempt to increase the sample for more stable calibrations.[41,42] Increasing the sample in further validation tests will be important.

The combined recruitment strategy opted to reach a representative sample, including patients as well as workers not seeking care. However, the sample was biased toward more women and workers with higher education. For further validation of the WIS-CMD, collaboration with large companies could be considered, preferably with diverse education levels among employees, to be representative of the working population. It would also be useful to increase the sample size according to recommendations.[32,48] Moreover, external validity tests, eg, comparing WIS-CMD scores with clinician judgments[29,32] (such as psychologists or occupational therapists experienced in vocational assessments), will be important in the further development of the WIS-CMD.

Conclusions

Grounded in workers' experiences, this study developed and drafted a scale to capture work instability related to depression and anxiety disorders to be used in pro-active management in the workforce. In an initial target sample, internal construct validity was analyzed, which guided item reduction and resulted in the 34-item WIS-CMD with satisfactory fundamental properties and unidimensionality. Further validity tests in a larger sample representative to a more diverse working population, with high difficulty items added, are warranted to refine the scale.

Author Contributions

The idea for this study was developed by GH, who designed and planned the study together with LD, KH, and ÅLN. LD carried out the recruitment and data collection and drafted the manuscript. LD, RF, and ÅLN performed and interpreted the

data analyses. All 5 authors contributed with methodological input and refined the manuscript.

Data Availability

Data are available from the corresponding author upon reasonable request.

ORCID iD

Louise Danielsson [iD] https://orcid.org/0000-0002-7089-3006

REFERENCES

1. Organisation for Economic Co-operation and Development (OECD). *Sick on the Job? Myths and Realities about Mental Health and Work*. Paris, France: OECD; 2012.
2. Vos T, Barber RM, Bell B, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. *The Lancet*. 2015;386:743-800.
3. World Health Organization. *Depression and Other Common Mental Disorders: Global Health Estimates*. Geneva, Switzerland: World Health Organization; 2017.
4. National Collaboration Centre for Mental Health. *Common Mental Health Disorders: Identification and Pathways to Care* (NICE clinical guidelines no. 123). Leicester, UK: British Psychological Society; 2011.
5. Thisted CN, Nielsen CV, Bjerrum M. Work participation among employees with common mental disorders: a meta-synthesis. *J Occup Rehabil*. 2018;28:452-464.
6. Aronsson G, Gustafsson K. Sickness presenteeism: prevalence, attendance-pressure factors, and an outline of a model for research. *J Occup Environ Med*. 2005;47:958-966.
7. Hjarsbech PU, Nielsen MBD, Andersen MF, Rugulies R, Christensen U. Struggling at work—a qualitative study of working Danes with depressive symptoms. *Disabil Rehabil*. 2015;37:1674-1682.
8. Bertilsson M, Petersson EL, Ostlund G, Waern M, Hensing G. Capacity to work while depressed and anxious—a phenomenological study. *Disabil Rehabil*. 2013;35:1705-1711.
9. Danielsson L, Elf M, Hensing G. Strategies to keep working among workers with common mental disorders—a grounded theory study. *Disabil Rehabil*. 2019;41:786-795.
10. Janssens H, Clays E, De Clercq B, De Bacquer D, Braeckman L. The relation between presenteeism and different types of future sickness absence. *J Occup Health*. 2013;55:132-141.
11. Adamsson A, Bernhardsson S. Symptoms that may be stress-related and lead to exhaustion disorder: a retrospective medical chart review in Swedish primary care. *BMC Fam Pract*. 2018;19:172.
12. de Vries H, Fishta A, Weikert B, Rodriguez Sanchez A, Wegewitz U. Determinants of sickness absence and return to work among employees with common mental disorders: a scoping review. *J Occup Rehabil*. 2018;28:393-417.
13. Holmgren K, Ivanoff SD. Women on sickness absence—views of possibilities and obstacles for returning to work. A focus group study. *Disabil Rehabil*. 2004;26:213-222.
14. Moise N, Khodneva Y, Jannat-Khah DP, et al. Observational study of the differential impact of time-varying depressive symptoms on all-cause and cause-specific mortality by health status in community-dwelling adults: the REGARDS study. *BMJ Open*. 2018;8:e017385.
15. Hjarsbech PU, Andersen RV, Christensen KB, Aust B, Borg V, Rugulies R. Clinical and non-clinical depressive symptoms and risk of long-term sickness absence among female employees in the Danish eldercare sector. *J Affect Disord*. 2011;129:87-93.
16. von Vultée P. The Munik—reteaming method is a way to rehabilitate psychosomatic and work-related diseases and a way to create healthy psychosocial workplaces. *Rehabil Process Outcome*. 2017;6:1179572717742661.
17. Organisation for Economic Co-operation and Development (OECD). *Fit Mind, Fit Job: From Evidence to Practice in Mental Health at Work, Mental Health at Work*. Paris, France: OECD; 2015.
18. Olsen IB, Øverland S, Reme SE, Løvvik C. Exploring work-related causal attributions of common mental disorders. *J Occup Rehabil*. 2015;25:493-505.
19. Mauramo E, Lallukka T, Lahelma E, Pietiläinen O, Rahkonen O. Common mental disorders and sickness absence. *J Occup Environ Med*. 2018;60:569-575.
20. Gensby U, Limborg HJ, Mahood Q, Ståhl C, Albertsen K. Employer strategies for preventing mental health related work disability: a scoping review. *Nordic J Working Life Stud*. 2019;9:57-101.
21. Danielsson L, Bertilsson M, Holmgren K, Hensing G. Working in dissonance: experiences of work instability in workers with common mental disorders. *BMC Public Health*. 2017;17:472.
22. Ilmarinen J. Work ability—a comprehensive concept for occupational health research and prevention. *Scand J Work Environ Health*. 2009;35:1-5.
23. Ilmarinen J, Tuomi K, Klockars M. Changes in the work ability of active employees over an 11-year period. *Scand J Work Environ Health*. 1997;23:49-57.
24. Abma FI, Bültmann U, Amick Iii BC, et al. The Work Role Functioning Questionnaire v2.0 showed consistent factor structure across six working samples. *J Occup Rehabil*. 2018;28:465-474.
25. Linden M, Baron S, Muschalla B. Relationship between work-related attitudes, performance and capacities according to the ICF in patients with mental disorders. *Psychopathology*. 2010;43:262-267.
26. Molodynski A, Linden M, Juckel G, et al. The reliability, validity, and applicability of an English language version of the Mini-ICF-APP. *Soc Psychiatry Psychiatr Epidemiol*. 2013;48:1347-1354.
27. Law M, Cooper B, Strong S, Stewart D, Rigby P, Letts L. The person-environment-occupation model: a transactive approach to occupational performance. *Can J Occup Ther*. 1996;63:9-23.
28. Gilworth G, Chamberlain MA, Harvey A, et al. Development of a work instability scale for rheumatoid arthritis. *Arthritis Care Res*. 2003;49:349-354.
29. Gilworth G, Carey A, Eyres S, et al. Screening for job loss: development of a work instability scale for traumatic brain injury. *Brain Inj*. 2006;20:835-843.
30. Gilworth G, Emery P, Barkham N, Smyth MG, Helliwell P, Tennant A. Reducing work disability in ankylosing spondylitis—development of a work instability scale for AS. *BMC Musculoskelet Disord*. 2009;10:68.
31. Gilworth G, Smyth MG, Smith J, Tennant A. The manual work instability scale: development and validation. *Occup Med (London)*. 2016;66:300-304.
32. McFadden E, Horton MC, Ford HL, Gilworth G, McFadden M, Tennant A. Screening for the risk of job loss in multiple sclerosis (MS): development of an MS-specific Work Instability Scale (MS-WIS). *Mult Scler*. 2012; 18:862-870.
33. World Health Organization. *The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines*. Geneva, Switzerland: World Health Organization, 1992.
34. Bech P, Olsen LR, Kjoller M, Rasmussen NK. Measuring well-being rather than the absence of distress symptoms: a comparison of the SF-36 mental health subscale and the WHO-five well-being scale. *Int J Methods Psychiatr Res*. 2003;12:85-91.
35. 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;47:141-148.
36. Tennant A, McKenna SP, Hagell P. Application of Rasch analysis in the development and application of quality of life instruments. *Value Health*. 2004;7: S22-S26.
37. 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.
38. Tennant A, Conaghan PG. The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis Care Res*. 2007;57:1358-1362.
39. Andrich D. *Rasch models for measurement*. London, England: SAGE; 1988.
40. Linacre JM. Sample size and item calibration stability. *Rasch Meas Trans*. 1994;7:328.
41. van der Beek NA, Hagemans ML, van der Ploeg AT, van Doorn PA, Merkies IS. The Rasch-built Pompe-specific activity (R-PAct) scale. *Neuromuscul Disord*. 2013;23:256-264.
42. Wright B. Rack and stack: time 1 vs. time 2. *Rasch Meas Trans*. 2003;17: 905-906.
43. Escobar JI, Interian A, Diaz-Martinez A, Gara M. Idiopathic physical symptoms: a common manifestation of psychiatric disorders in primary care. *CNS Spectr*. 2006;11:201-210.
44. Newcomb RD, Steffen MW, Breeher LE, et al. Screening for depression in the occupational health setting. *Occup Med (London)*. 2016;66:390-393.
45. Christensen KB, Makransky G, Horton M. Critical values for Yen's Q_3: identification of local dependence in the Rasch model using residual correlations. *Appl Psychol Meas*. 2017;41:178-194.
46. Miller KJ, Pollock CL, Brouwer B, Garland SJ. Use of Rasch analysis to evaluate and refine the community balance and mobility scale for use in ambulatory community-dwelling adults following stroke. *Phys Ther*. 2016;96:1648-1657.
47. Johansen T, Lund T, Jensen C, et al. Cross-cultural adaptation of the Work Role Functioning Questionnaire 2.0 to Norwegian and Danish. *Work*. 2018;59: 471-478.
48. Houts CR, Edwards MC, Wirth RJ, Deal LS. A review of empirical research related to the use of small quantitative samples in clinical outcome scale development. *Qual Life Res*. 2016;25:2685-2691.