A Systematic Review of Health-related Work Outcome Measures, and Quality Criteria based Evaluations of their Psychometric Properties

Abstract – 291 Words

Objective: To examine the state of psychometric validation in the health-related work outcome literature.

Data Sources: We searched Pubmed, PMC, CINAHL, EMBASE [+ EMBASE Classic], and PsycINFO, from inception to January 2016., using the search terms: Stroke, Multiple Sclerosis, Epilepsy, Spinal Cord Injury, Brain Injury, Musculoskeletal Disease, Work, Absenteeism, Presenteeism, Occupation, Employment, Job, Outcome measure, Assessment, Work Capacity Evaluation, Scale, and Questionnaire.

Study Selection & Data Extraction: 597 outcome measures were identified from the 22,676 retrieved abstracts. Inclusion was based on content analysis. 95 health-related work outcome measures were retained, of which two were treated as outliers and therefore are discussed separately. All six authors individually organized the 93 remaining scales based on their content. A follow-up search using the same sources, and time period, with the name of the outcome measures and the following terms: Psychometric, Reliability, Validity, Responsiveness, identified 263 unique Classical Test Theory (CTT) psychometric property datasets for the 93 tools. An assessment criterion for psychometric properties was applied to each manuscript, and where consensus was not achieved, the rating delivered by the majority of the assessors was reported.

Data Synthesis: 18 of the manuscripts reporting psychometric data were not accessible and therefore could not be assessed. 39 scored less than 20% of the maximum achievable score, 106 between 20-40%, 82 between 40- 60%, 15 scored between 60-80%, and only 1 scored above 80%. The three outcome measures associated with the highest scoring datasets were the Sheehan Disability Scale, the Fear Avoidance Beliefs Questionnaire, and the Assessment of Subjective Handicap of Epilepsy scales. And finally, only 2 psychometric validation datasets reported the complete set of baseline psychometric properties.

Conclusion: This systematic review highlights the current limitations of the health-related work outcome measure literature, including the limited number of robust tools available.

Key Terms: Psychometrics, Reliability, Validity, Vocational Rehabilitation, Work Instability

Introduction

The importance of work, and its role in maintaining the health and well-being of an individual has been increasingly recognised [1], as has the destabilizing effect of unemployment [1]. As such, the Universal Declaration of Human Rights (UDHR) codifies each and every person's right to work [2]. The services that aid people in retaining work, or altering their job specifications so that they can continue working regardless of their disability or chronic condition, can be examined at two levels: societal/state and individual service/intervention.

At the state level, governments employ inclusive policies, such as reasonable adjustment guidance in the 2010 Equality act [3], to support those with medical conditions and disability in achieving as similar a level of function as possible to their pre-morbid state [4]. At the level of the individual, specific interventions are utilized. The intervention in this circumstance is referred to as vocational rehabilitation; the multifaceted process that enables people with health conditions to overcome barriers to accessing, maintaining or returning to meaningful occupation [5]. States can evaluate the efficacy of specific policies, by utilizing macro-economic data such as 'number in work' (employment rate) [5,6]. However, the macroscopic perspective provides little insight into the effectiveness of specific individual-level interventions. Furthermore, whilst it is relatively easy to identify a patient who is returning to work from unemployment, it is more difficult to identify the impact of supportive interventions to remain in work. For example, an intervention for an individual with multiple sclerosis may target anxiety, low self efficacy, fatigue, and difficulties with attention and memory [7]. These are just two examples of why selecting appropriate outcomes to measure to illustrate the effectiveness of an intervention can be difficult in vocational rehabilitation.

The tools used to assess the efficacy of these interventions are referred to as health-related work outcome measures [8]. Health-related work outcome measures are typically described as tools capable of capturing the interplay between an individual's health and work performance. A narrative review of these tools, and their uses, describes several areas they may capture, including: the work

status of the individual; how well they are working; and how many hours they are fully effective [8]. Given the breadth of the literature in this area of research, the second important obstacle when attempting to illustrate the effectiveness of an intervention is that there are many potential tools available for each measurement construct, and it can be very difficult to discern which ones are objectively better.

The most recent systematic review of health-related work-outcome measures, focused on the presenteeism sub-set of measures, which attempt to quantify the effect of attending work whilst unwell, and the associated decrease in productivity [9]. Roy and colleagues concluded that there was insufficient psychometric evidence to determine which of the instruments they identified was preferable. Whilst some tools were associated with high quality studies, and promising results, it was clear that future studies would need to focus on fixing the lack of evidence available surrounding the reliability and responsiveness of presenteeism tools to be able to make such a decision [9]. Many believe this limitation is applicable to the other groups of health-related work outcome measures available in the literature, however, it is currently an unsubstantiated claim.

The aims of this systematic review are three fold: (1) to identify the tools available in the health-related work outcome literature. Secondly (2), to collate the psychometric properties of each of the previously identified tools. And finally (3), by utilizing validated quality criteria, to assess the quality of psychometric validation in the health-related work outcome literature. This information should allow the vocational rehabilitation community to know the available range, evaluate, and select the most appropriate measure for use in their service.

Method

Stage 1 (Primary Data Sources, Study Selection & Data Extraction) –

Pairs of the authors of this review (hereafter referred to as reviewers), conducted a search of Pubmed, PMC, CINAHL, EMBASE (+ EMBASE Classic), and PsycINFO, using the search phrases outlined in figure 1, from the inception of the databases to January 2016. A total of 22,676 abstracts were found (not corrected for duplicates), from which 597 outcome measures were identified. Of the total 597 outcome measures, after the application of the inclusion and exclusion criteria (detailed in figure 1), 93 outcome measures remained for further analysis (+ 2 outliers). Work-environment measures, have been purposely excluded from this review, as they are treated separately to measures of an individual's capacity in the research literature, and form a substantial body of work in their own right. Additionally, tools that do not specifically reference work have also been excluded, as they measure an individual's capability in uncontexualised scenarios, which do not give direct insight into an individual's ability to complete the activities relevant to their vocation/trade. Generic measures of physical and mental function also form a significant part of the literature in their own right. However, it should be noted that for individual's who have been out of work for extended periods of time, the latter (uncontexualised) measures may be more appropriate, and is therefore a potential limitation of the search strategy.

Stage 2 (Organising the Tools) –

The content (questions) of each work related outcome measure was examined by a multi-disciplinary team (MDT) consisting of three occupational therapists, a clinical psychologist, a physician specializing in neuro-rehabilitation, and a researcher, all of whom work in the field of vocational rehabilitation. The MDT met twice to assess the content of each questionnaire, and each member individually assigned one, or several of the codes listed in figure 2, to the tool. In situations where consensus was not achieved, codes were only used to described the tool if a majority of the panel agreed it was appropriate. Based on the results of this exercise, the tools were organized by content (Tables 1 -5).

Stage 3 (Psychometric Data: Data Sources, Study Selection) –

The psychometric quantities of interest were then defined using the the **CO**nsensus-based **S**tandards for the selection of health **M**easurement **IN**struments (COSMIN) Taxonomy [10] of measurement properties and the COSMIN checklist [11]. In total, nine domains were identified: internal consistency, reliability (test-retest reliability, inter-rater reliability and intra-rater reliability), measurement error, content/face validity, structural validity, hypotheses testing, cross-cultural validity, criterion validity, and responsiveness. Definitions for these psychometric properties can be found in figure 3. All of these psychometric measurements are examples of Classical Test Theory (CTT) properties. Rasch analysis, an example of an Item Response Theory (IRT) based psychometric analysis, is sometimes used instead of the more traditional CTT, however, given the breadth of this review, we feel justified in excluding this subset of validation studies, especially since they require their own unique quality criteria.

The psychometric data for each outcome measure was retrieved by using the search phrase: (("Name of Outcome Measure") AND (Psychometric OR Reliability OR Validity OR Responsiveness)), in all four aforementioned databases. A total of 3,449 abstracts were returned by the search parameters. Additionally, psychometric data reported in the paper describing the development of the tools has also been reported. These original citations were found using a less structured approach, which included searching reference lists of the validation papers and direct searching in the four aforementioned databases. After inclusion of the psychometric data published in the original development studies, and the already identified psychometric validation studies (corrected for duplicated), in total, 263 sets of psychometric data corresponding to the 93 work-related outcome measures were identified. Inclusion and exclusion criteria for the psychometric validation studies is described in detail in figure 1. In situations where the manuscript was inaccessible, and the authors unreachable, but the abstract explicitly stated that psychometric validation was conducted, or specific properties were reported, the study was included in the results. As such, 18 inaccessible manuscripts were retained. Their inclusion illustrates an important distinction between those tools for which no validation studies were identified,

and those for which some exist but were inaccessible. Similarly, the original citations for tools which did not report any psychometric properties were also retained, as this provides valuable information about the development of the tool"

Stage 4 – Assessing the Products of Psychometric Analysis

There are several studies in the literature that describe quality criteria for assessing psychometric properties [12-14]. We adapted the Terwee et al., criteria [12], which was selected as it most closely resembled the selection of psychometric properties identified from the COSMIN checklist [11], outlined in figure 3. Where necessary it was complemented with the Scientific Advisory Committee of the Medical Outcomes Trust criteria [14]. One specific property was excluded from the assessment, predictive validity, as the methods by which it can be assessed were deemed too varied to assess effectively. The resulting assessment criteria (figure 3) resulted in a score ranging from 0-2 for each psychometric property produced in each validation study. Two reviewers independently applied the criteria to each study, and where consensus could not be achieved, a third rater was recruited, and the reported score was that conferred by the majority of raters. The maximum achievable score was 16 for non-cross cultural validation studies, and 18 for cross cultural validation studies. The score for each individual property, the total scores, as well as the percentage of the maximum achievable scores are reported in Tables 1-5.

Results

95 work-related outcome measures were identified in our search of the literature. 93 tools were subsequently organised into groups based on their content, and two specific tools were identified as notable outlier and are discussed separately. Table 1 details the psychometric properties of the 30 tools that were identified as relating to the physical aspects of work, including physical capacity, presenteeism, absenteeism, and performance. Table 2 details the measurement properties of the 25 tools related to the psychosocial aspects of work: personality, stress, satisfaction, boredom, and well-being. Table 3 includes the psychometric properties of the 10 self-efficacy, 5 work status, and 1 work demands questionnaire. Table 4 details the psychometric properties of the 14 work instability scales identified. And finally, table 5 details the psychometric properties relating to the 8 work ability scales identified. Work ability [271] is the combined psychosocial and physical work capacities of an individual, and therefore is a combination of any of the previous areas (except work instability, which is a result of work ability being exceeded by job demands). Figure 4 illustrates which areas each of the work ability scales captures based on the content analysis conducted by the MDT.

The State of Psychometric Validation Studies in the Vocational Rehabilitation Literature

The search for CTT based psychometric properties associated with the aforementioned tools, identified 263 datasets that reported one or more of the relevant measurement properties described in figure 3. Below is an overview of the number of scales upon which each of the types of psychometric analysis had been conducted. The number of times a specific form of analysis was conducted on an individual scale, can be inferred from tables 1-5.

Examination of Validity Data

All 93 (100%) of the scales appear to have been examined with respect to their content/face validity. However, the threshold for scoring 1 of 2 points in the quality criteria did not require a group of stakeholders to examine the content, which is the most widely recognized test for assessing face and

content validity. On closer examination, only 32 (34.4%) of outcome measures had a panel of experts and/or patients examine the content. With regards to cross-cultural validation, only 30 (32.3%) outcome measures were translated and the results subsequently compared to the original language. For criterion validity, 66 (71.0%) scales underwent some form of analysis comparing their correlation with other measures of work-related outcomes and a variety of other factors. And finally, when considering hypothesis testing, and examination of structural validity using either principal component analysis, confirmatory or exploratory factor analysis, both properties were scrutinized in 46 (49.5%) of the total 93 outcomes measures.

Examination of Reliability & Responsiveness Data

Of the 93 outcome measures, internal consistency was reported for 62 (66.7%), and the vast majority of these scales met the minimum requirement of Cronbach's alpha > 0.7 [296]. Comparatively, 37 (40.0%) outcome measures had test-retest analysis conducted, to establish their stability over long and short periods of time. The two reliability quantities that were the least frequently investigated were inter-rater consistency, and the Standard Error of Measurement (SEM – the relationship between the minimally detectable change and the smallest clinically important change in score), with 7 (7.5%), and 8 (8.6%) outcome measures, respectively, having this data reported as part of the psychometric battery of tests. The low rate of inter-rater consistency analysis is likely to be a reflection of the fact that the vast majority of the tools examined are patient reported outcome measures, and therefore this property is not relevant.

Responsiveness analysis sat between the two extremes described for reliability analysis, with a total of 17 outcome measures (18.3%) of the 93 identified, having been tested.

<u>Application of the Quality Criteria to Evaluate the Identified Psychometric Properties</u>

The range for the percentage of the maximum possible score achieved by the validation studies ranged from 6.3% to 81.3%. Assessment of the measurement properties demonstrated that 39 studies

scored less than 20% of the maximum achievable score, 106 scored between 20 and 40%, 82 scored between 40 and 60%, 15 scored between 60 and 80%, and only 1 scored in excess of 80%. 18 of the manuscripts reporting psychometric data were not accessible and therefore could not be assessed, but have been identified in tables 1-5. Of the 263 datasets, only 2 (0.76%) had values for all of the baseline properties [151,167].

Notable Measures

Although a single best tool cannot be acknowledged due to the prevalence of incomplete psychometric datasets, the following outcome measures have been identified, as they appear to be the most reliable, valid and responsive tools we have identified in this review. Only one outcome measure had a validation study which reported psychometric properties that scored more than 80%, the Sheehan Disability Scale [167]. The Sheehan scale [160] also had validation studies associated with it that scored between 70 and 80%. However, two other outcome measures also had validation studies that illustrated similar quality measurement properties: The Fear Avoidance Beliefs Questionnaire [107], and the Assessment of Subjective Handicap of Epilepsy [294]. Descriptions, and the strengths/limitations of each tools are discussed below.

Sheehan Disability Scale (SDS) [160] – A self-reported assessment of functional impairment, consisting of 5 items. The first three are global rating scales which assess impairment in work, home and family responsibilities due to symptoms on 10-point numerical rating scales. There are two additional questions which measure presenteeism and absenteeism over the preceding 7 days. The main limitation surrounding the use of this tool is that the validation studies have all occurred in psychological disease patient populations. Therefore, its utility in populations where vocational rehabilitation is more commonly deployed, e.g. musculoskeletal diseases and neurological diseases, is currently unknown.

Assessment of Subjective Handicap of Epilepsy Scale (SHE) [294] – A self-reported 32 item questionnaire, with a 5 point likert scale for each question. The items are organized into six subscales: 1

- Work and activity (eight items); 2 - Social and personal (four items); 3 - Physical (four items); 4 -Self-perception (five items); 5 - Life-satisfaction (four items), and 6 -Change (seven items). Each subscale's scores can be linearly transformed onto a 0-100 scale to produce a score indicating the degree of handicap/satisfaction. Whilst the tool is psychometrically robust, and as a measure of work ability it captures a variety of work-related areas, it is currently only available and validated for use in people with epilepsy, which limits its utility to in clinical practice.

Fear & Avoidance Beliefs Questionnaire (FABQ) [107] – A self-reported, 16 item questionnaire, with a 6-point ordinal scale superimposed on a 3-point likert scale. It is based on the "Fear-Avoidance Model of Exaggerated Pain Perception", measuring a patient's fear of pain, and the subsequent avoidance of two areas: physical activity and work. Compared to the other two psychometrically robust tools identified by this review, the FABQ is the most widely validated tool, making it the obvious suggestion for use as a generic work-related measure. However, the information captured by the tool is limited to the physical aspects of work, and those in pain, therefore a holistic assessment is likely to require the use of other tools in tandem.

<u>Novel Measures</u>

The first of the previously identified notable outliers is the Return to Work Questionnaire [297]. Although it has not been assessed for its psychometric properties, and its lack of scoring would suggest it is not an outcome measure at all, it is still unique in what it aims to achieve. The tool aims to facilitate conversations about an employees' work ability in the context of their current environment, and therefore make the employer aware of any issues that may exist due to an existing or progressing medical condition. The over arching aim is for both parties to present solutions to any identified issues. The instrument does not fall under remit of the traditional vocational rehabilitation setting, which exists between an Occupational therapist (OT)/specialist and a patient, or potentially an OT and the employer. Rather it exposes an area that tools are rarely made to service, the employer-employee relationship.

This particular tool has a significant focus on stress, and so may not be useful in a large majority of potential situations, but the underlying principle is still worth considering.

The second notable outlier is the Job Accommodation Scale (JAS) [298], as it presents a standardized method of assessing supervisors' willingness/ability to engage with specific physical activity, environment, and demands related accommodations. The novelty of this tool includes its role in supporting communication between the vocational rehabilitation specialist and the employer.

Discussion

Assessment of the state of the validation literature appears to suggest that work-related outcome measures are not as rigorously tested as they should be, which echoes the findings of previous studies [145]. There are only two validation studies that have all the baselines quantities measures in a single paper. Moreover, even if we were to aggregate all the data for the other tools, there is still not an additional outcome measure with a complete set of baseline reliability and validity values, disregarding the obvious problems with compiling a complete data set from several different papers. The incompleteness of the meta-data means that any comparison between the tools would be incomplete, and therefore, as we suggested previously, it is not possible to recommend any single tool.

With regards to which properties are most commonly considered. The most frequently reported CTT reliability property is the Cronbach's alpha for internal consistency, whereas the most commonly actively investigated validity related property was criterion validity. The properties that are very rarely reported, include measurement error, responsiveness, and genuine investigations of content validity by experts and/or stakeholders. The final key finding, is that it appears as though personality, demands, boredom and fatigue have very few outcome measures that are solely focused on their specific measurement in a work specific manner. This peculiarity is worth noting, as it identifies a significant gap in the literature, and therefore avenues for future research in work-related outcome measure development.

In this systematic review we have collated a comprehensive set of the psychometric data for a variety of tools, with regards to several populations of interest (see reference section – all citations coded to reflect the population studied). However, based on the results, a purely quantitative solution to the problem of health-related work outcome measure selection is unlikely to be feasible. As such, rehabilitation specialists will need to apply some level of clinical judgment [299], to navigate the incompleteness and equivocalness of the psychometric data currently available. A holistic assessment of

the data presented in this review, based on individual service specifications and clinical judgment, is how we envision the problem of outcome measure selection to be solved in the status quo.

Limitations of the Quality Criteria (Figure 3)

The quality criteria utilized in this review have clearly identified several scales whose psychometric properties are objectively better than other potential tools on offer. However, there are certain criteria that we felt were either too restrictive or not included at all, which should be considered in the future. Firstly, the magnitude of the correlation co-efficient necessary to score 2 points when assessing criterion validity is inappropriate in this situation. The use of 0.7 when a definitive gold standard measure exists is reasonable, however, there is no accepted gold standard measure for many of the domains that the health-related work tools were compared to, as many studies noted in their methods. Therefore, based on our experience with this review, we would propose the following change; a modest reduction to 0.5 as the threshold for 2 points. Only a handful of studies had coefficients consistently in excess of 0.7. Given that the literature lacks a consensus on a gold standard measure, the very high (0.7) criteria is not appropriate for these circumstances, as each potential comparator will have a unique impact on the strength of the correlation. Therefore, the effect of this change would be to provide a greater margin for error, thus accounting for this variability, and so, preventing unfair penalization due to the lack of a gold standard measure. The additional point many studies would score would reduce the homogeneity of the concurrent validity scores, and thus, allow us to better distinguish between the tools.

Moreover, areas such as known-groups validity were excluded altogether as they were not included in the COSMIN taxonomy [10], and therefore deemed accessory properties. Whilst we maintain that this description is reasonable, it is the additional properties such as known groups validity and predictive validity that distinguish excellent outcome measures from the acceptable. The limitation to including the latter two proposed validity properties is that a quality criterion is necessary which

captures the breadth of statistical tools and comparisons that could potentially be made, which unfortunately is currently lacking.

Furthermore, producing an overall (summed) score of the measurement properties for each tool is not without controversy. It assumes that each of the measurement properties is equally weighted, which all of the authors of this study principally disagree with. Unfortunately, the authors of the criteria used in this study, only provide qualitative evidence for why the scores should not be summed [12]. However, other validated quality criteria have demonstrated quantitatively that comparisons based on summary scores can be valid and reliable, as it was in their criteria for RCTs [300]. As such, we have presented the summary score as an additional descriptive tool to be used alongside the individual scores for each property.

Strengths and Weaknesses of the Study

There were many outcome measures identified in the original search, that were not published, and so we had to contact authors to source the original questionnaires. The sourcing of these unpublished outcome measures was one of the methods by which we sought to minimize publication bias. Moreover, anecdotal evidence suggests that there are many tools commonly used in research which are rarely utilized in clinical practice. The best example of this is the Occupational Outcomes Questionnaire [199]. Although it has been utilized once in the scientific literature [301], it is used by a number of institutions throughout the United States. Our attempts to capture a variety of scales, has identified tools that have been previously overlooked, but could provide substantial value to practitioners, given that most of them are freely available for use, from the authors. Despite a comprehensive search strategy, the main limitation of the study is that the retrieval process for both the outcome measures, and the psychometric data, was not exhaustive. However, the inclusion of relevant, but inaccessibly manuscripts provides an additional element of comprehensiveness, creating a more accurate reflection of the literature. The limitations of the search strategy are compounded by fact that the choice of descriptive headings applied to indexed research sometimes did not include the word

'work', or another related term (e.g. 302). Moreover, the use of specific conditions in search 1, and 2 means that we are more likely to have excluded generic (non-condition) specific tools. This was a limitation foreseen during the planning of the review, but was deemed acceptable, given that the abstract count increases exponentially when unbound by these additional condition-specific terms, which would have rendered the review unmanageable.

Comparison to Health-related Work Outcome Measurement Literature

Our search strategy captured all the scales and psychometric properties identified by Roy [9], suggesting that the search strategy used in this review is appropriate. Roy and colleagues' conclusion is very similar to our results, in that future studies should focus on improving the body of psychometric evidence available in the literature. However, there are several key differences between the sub-set of presenteeism scales that exists in this paper, and Roy's original systematic review. Firstly, we captured several additional scales that were found to capture presenteeism as one of their measurements, which is likely to be a result of our more extensive search strategy. Furthermore, the seven presenteeism scales identified by Roy et al. [9] are not all classified as presenteeism scales in this study. Roy and colleagues relied on the original authors' classification of the scales. Similar to a previous study examining the content validity of a single fatigue related outcome measure [303], in this study the stakeholders were given a range of tools, and complete freedom to label them as they saw fit, giving an additional level of insight into the utility of these tools from the perspective of practitioners. Thus we would argue that this study is the first to provide a contextual analysis of a wide range of work-related outcome measures, and so explains why our results differ from Roy's [9]. Finally, given that the sample of tools examined in this systematic review includes a range of other tools, of which presenteeism is one example, we are quite confident that within the literature, this is the most comprehensive review of work-related outcome measures.

Meaning of study; mechanisms and implications for policy makers and clinicians

Our review presents a comprehensive list of the available outcome measures in vocational rehabilitation, and the relevant psychometric data to aid policy makers, clinicians, and occupational therapists in identifying the most appropriate tools for their services. The assessment of the psychometric properties allows for quantitative discrimination between the performance of these tools. Future research should focus on the several conceptual and practical gaps identified in the literature surrounding measurement properties in vocational rehabilitation.

Conclusion

In clinical practice there are a wide range of measures that can be used to evaluate the effectiveness of vocational rehabilitation (VR) services and interventions. However, the psychometric data sets associated with these tools are commonly incomplete, which has contributed to our inability to identify which tools perform best. The key inference from this study is that it is not currently possible to determine the best method by which to measure work status, especially for those individuals attempting to remain in work despite increased disability.

References

* - No specific psychometric properties reported
• - General Work Population/Without a specific identified health condition/Healthy Participants, •• - Musculoskeletal Disorders and Pain, ••• - Spinal Cord Injury, •••• - Brain Injury, ••••• - Neurological and Psychiatric Disorders, and •••••• - Other Disorders/ Combination of Disorders

[1*] - Waddell G, Burton A. Is work good for your health and well-being? London: The Stationery Office, 2006

[2*] - UN General Assembly, Universal Declaration of Human Rights, 10 December 1948, Resolution 217A (III), available at: http://www.un.org/en/universal-declaration-human-rights/ [accessed 27 January 2016]

[3*] - Equality Act 2010, c. 15 (Eng.).

[4*] - Waddell, G., Burton, A. and Kendall, N. (2008). *Vocational rehabilitation*. [Great Britain]: TSO.

[5*] - Jones JR, Huxtable CS, Hodgson JT, Price MJ, Self-reported work-related illness in 2001/02: Results from a household survey, 2003. Published on the Internet at www.hse.gov.uk/statistics/causdis/swi0102.pdf

[6*] - Jones JR, Huxtable CS and Hodgson JT, Self-reported work-related illness in 2003/04: Results from the Labour Force Survey, 2005. Published on the Internet at www.hse.gov.uk/statistics/causdis/swi0304.pdf

[7*] - Bevan S, Zheltoukhova K, McGee R, Blazey L (2011) Ready to Work? Meeting the Employment and Career Aspirations of People with Multiple Sclerosis. London: The Work Foundation.

[8*] - Amick, B., Lerner, D., Rogers, W., Rooney, T. and Katz, J. (2000). A Review of Health-Related Work Outcome Measures and Their Uses, and Recommended Measures. Spine, 25(24), pp.3152-3160.

[9*] - Roy J, Desmeules F, MacDermid J. (2011). Psychometric properties of presenteeism scales for musculoskeletal disorders: A systematic review. J Rehabil Med . 43(1):23-31.

[10*] - Mokkink, L., Terwee, C., Patrick, D., Alonso, J., Stratford, P., Knol, D., Bouter, L. and de Vet, H. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology*, 63(7), pp.737-745.

[11*] - Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., ... & De Vet, H. C. (2010). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Quality of Life Research*, *19*(4), 539-549.

[12*] – Terwee, C., Bot, S., de Boer, M., van der Windt, D., Knol, D., Dekker, J., Bouter, L. and de Vet, H. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology*, 60(1), pp.34-42.

[13*] - Andresen, E. (2000). Criteria for assessing the tools of disability outcomes research. *Archives of Physical Medicine and Rehabilitation*, 81, pp.S15-S20.

[14*] - Scientific Advisory Committee of the Medical Outcomes Trust (2002) Assessing health status and quality of life instruments: Attributes and review criteria. Qual. Life Res. 11, 193–205

[15*] - Middel, B., & van Sonderen, E. (2002). Statistical significant change versus relevant or important change in (quasi) experimental design: some conceptual and methodological problems in estimating magnitude of intervention-related change in health services research. *International Journal of Integrated Care*, *2*, e15.

[16•] – Rantanen, J., Feldt, T., Hakanen, J., Kokko, K., Huhtala, M., Pulkkinen, L., and Schaufeli, W. (2015). Cross-national and longitudinal investigation of a short measure of workaholism. *Industrial Health*, 53(2), pp.113-123.

[17•] - Schaufeli, W., Shimazu, A. and Taris, T. (2009). Being Driven to Work Excessively Hard: The Evaluation of a Two-Factor Measure of Workaholism in The Netherlands and Japan. *Cross-Cultural Research*, 43(4), pp.320-348.

[18••••••] - Bolton B, Roessler R. (1986). The work personality profile: factor scales, reliability, validity and norms. Vocat Eval Work Adjustment Bull. 19:143-149.

[19•] - Siu, A. M. H., Yau, M. K., & Lam, P. C. W. (1998). The Chinese Work Personality Profile: Factor scales, reliability, and norms. *Vocational Evaluation and Work Adjustment Bulletin, 31*, 87-92

[20•••••] - Law, C., Siu, A., Lee, J. and Lee, S. (2006). Prediction of Work Rehabilitation Placements Using the Chinese Work Personality Profile. *Psychiatric Rehabilitation Journal*, 30(2), pp.120-128.

[21••••••] - Williams ER (1997) Work Personality Profile: Validation within the supported employment environment. The Journal of Rehabilitation 63(2), 26-30.

[22•] - Spector, P. E., (1988). Development of the Work Locus of Control Scale. Journal of Occupational Psychology. 61:335-340.

[23•] – Oliver J., Jose P., and Brough P., (2006). Confirmatory Factor Analysis of the Work Locus of Control Scale, Educational and Psychological Measurement. 66(5):835-851.

[24•••] - Crewe N, Krause S. An eleven-year follow-up of adjustment to spinal cord injury. (1991) Rehabil Psychol. 35:205–210

[25•••] - Krause, J. (1998). Dimensions of subjective well-being after spinal cord injury: An empirical analysis by gender and race/ethnicity. *Archives of Physical Medicine and Rehabilitation*, 79(8), pp.900-909.

[26•••] – Krause, J. (1992). Life satisfaction after spinal cord injury: A descriptive study. *Rehabilitation Psychology*, 37(1), pp.61-70.

[27•] - Holmgren K, Hensing G, Dahlin-Ivanoff S. (2009), Development of a questionnaire assessing work-related stress in women – identifying individuals who risk being put on sick leave. Disability & Rehabilitation . 31(4):284-292.

[28•] - Khader Y, Airan D, Al-Faouri I. (2009). Work Stress Inventory for Dental Assistants: Development and Psychometric Evaluation. Journal of Public Health Dentistry . 69(1):56-61.

[29•] - Rosnawati M.R., Moe H., Masilamani R., and A. Darus, (2010). The Bahasa Melayu Version of the Nursing Stress Scale Among Nurses: A Reliability Study in Malaysia; Asia Pac J Public Health, 22: 501

[30•] - Lee, M., Holzemer, W. and Faucett, J. (2007). Psychometric Evaluation of the Nursing Stress Scale (NSS) Among Chinese Nurses in Taiwan. *J Nurs Measure*, 15(2), pp.133-144.

[31•] - Gray-Toft P., Anderson J.G., (1981). The Nursing Stress Scale: Development of an Instrument; Journal of Behavioural Assessment; 3(1):11-23

[32•] - French S.E., Lenton R., Walters V., Eyles J., (2000). An empirical evaluation of an Expanded Nursing Stress Scale; Journal of Nursing Measurement, 8(2):161-178

[33•] - Hurrell J.J. Jr, McLaney M.A. (1988). Exposure to job stress--a new psychometric instrument; Scand J Work Environ Health; 14 suppl 1:27-28

[34•] - Stamps, P., Piedmont, E., Slavitt, D. and Haase, A. (1978). Measurement of Work Satisfaction among Health Professionals. *Medical Care*, 16(4), pp.337-352.

[35•] - Mueller, C. W., & McCloskey, J. C. (1990). Nurses' job satisfaction: A proposed measure. *Nursing Research*, *39*:113-117.

[36•] - Prosen, M. and Piskar, F. (2013). Job satisfaction of Slovenian hospital nursing workforce. *J Nurs Manag*, 23(2),

pp.242-251.

[37•] - Tourangeau, A. E.; McGillis Hall, L.; Doran, D. M.; Petch, T., (2006). Measurement of Nurse Job Satisfaction Using the McCloskey/Mueller Satisfaction Scale, Nursing Research. 55(2):128–136

[38•] - Gillet, B. and Schwab, D. (1975). Convergent and discriminant validities of corresponding Job Descriptive Index and Minnesota Satisfaction Questionnaire scales. *Journal of Applied Psychology*, 60(3), pp.313-317.

[39•] - Asegid, A., Belachew, T. and Yimam, E. (2014). Factors Influencing Job Satisfaction and Anticipated Turnover among Nurses in Sidama Zone Public Health Facilities, South Ethiopia. *Nursing Research and Practice*, 2014, pp.1-26.

[40•] - Weng, R., Huang, C., Tsai, W., Chang, L., Lin, S. and Lee, M. (2010). Exploring the impact of mentoring functions on job satisfaction and organizational commitment of new staff nurses. *BMC Health Services Research*, 10(1), p.240.

[41•] – Weiss D.J., Dawis, R.V., England, G.W., & Lofquist, L.H. (1967). Manual for the Minnesota Satisfaction Questionnaire. Minneapolis: University of Minnesota.

[42•] - Hirschfeld, R. (2000). Does Revising the Intrinsic and Extrinsic Subscales of the Minnesota Satisfaction Questionnaire Short Form Make a Difference?. *Educational and Psychological Measurement*, 60(2), pp.255-270.

[43•] –Hancer M., and George R.T. (2003). Job Satisfaction Of Restaurant Employees: An Empirical Investigation Using The Minnesota Satisfaction ; Questionnaire Journal of Hospitality & Tourism Research. 27:85- 100

[44•] - Lee, T. W. (1986). Toward the development and validation of a measure of job boredom. *Manhattan College Journal of Business, 15*:22-28.

[45•] – Tsutsumi A., Ishitake T., Peter R., Siegrist J. & T. Matoba. (2001). The Japanese version of the Effort-Reward Imbalance Questionnaire: A study in dental technicians, Work & Stress: An International Journal of Work, Health & Organisations, 15(1):86-96

[46•] – Lundkvist E., Stenling A., Gustafsson H. & P. Hassmén. (2014). How to Measure Coach Burnout: An Evaluation of Three Burnout Measures, Measurement in Physical Education and Exercise Science, 18(3):209-226

[47•] - Martin Lekutle & Jan Alewyn Nel (2012) Psychometric Evaluation of The Utrecht Work Engagement Scale (UWES) and Oldenburg Burnout Inventory (OLBI) within a Cement Factory, Journal of Psychology in Africa, 22:4, 641-647

[48•] – Peterson U, Bergström G, Demerouti E, (2011). Burnout levels and self-rated health prospectively predict future long-term sickness absence: a study among female health professionals, J Occup Environ Med. 53(7):788-93

[49•] - Demerouti E, Bakker AB, Nachreiner F, Schaufeli WB. (2001). The job demands–resources model of burnout. J Appl Psychol, 86:499–512

[50•] - Qiao H, Schaufeli W. (2010). The Convergent Validity of Four Burnout Measures in a Chinese Sample: A Confirmatory Factor-Analytic Approach. Applied Psychology. 60(1):87-111.

[51•] - Halbesleben J, Demerouti E. (2005). The construct validity of an alternative measure of burnout: Investigating the English translation of the Oldenburg Burnout Inventory. Work & Stress. 19(3):208-220.

[52•] - Katwyk Van PT, Fox S, Spector PE, Kelloway K. (2000). Using the Job-Related Affective Well-Being Scale (JAWS) to investigate affective responses to work stressors. J Occup Health Psychol. 5:219-30.

[53•] - Basińska, B., Gruszczyńska, E. and Schaufeli, W. (2014). Psychometric properties of the polish version of the Job-related Affective Well-being Scale. *IJOMEH*, 27(6), pp.993-1004.

[54•] - Mäkikangas, A., Feldt, T. and Kinnunen, U. (2007). Warr's scale of job-related affective well-being: A longitudinal examination of its structure and relationships with work characteristics. *Work & Stress*, 21(3), pp.197-219.

[55•] - Hancer, M. (2005). Dimensions of the turkish version of the psychological empowerment scale. *Psychological Reports*, 97(6), p.645.

[56•] - Hancer, M., George, R. and Kim, B. (2005). An examination of dimensions of psychological empowerment scale for service employees 1. *Psychological Reports*, 97(2), pp.667-672.

[57•] - Kraimer, M., Seibert, S. and Liden, R. (1999). Psychological Empowerment as a Multidimensional Construct: A Test of Construct Validity. *Educational and Psychological Measurement*, 59(1), pp.127-142.

[58•] - Uner S, Turan S. (2010). The construct validity and reliability of the Turkish version of Spreitzer's psychological empowerment scale. BMC Public Health . 10(1):117.

[59•] - Albar M, García-Ramírez M, Jiménez A, Garrido R. (2010). Spanish Adaptation of the Scale of Psychological Empowerment in the Workplace. Span. j. psychol. 15(02):793-800.

[60•] - Spreitzer, G. (1995). Psychological empowerment in the workplace: Dimensions, measurement, and validation. Academy of Management Journal. 38(5), 1442-1465

[61•] - Seppälä, P., Mauno, S., Feldt, T., Hakanen, J., Kinnunen, U., Tolvanen, A. and Schaufeli, W. (2008). The Construct Validity of the Utrecht Work Engagement Scale: Multisample and Longitudinal Evidence. *Journal of Happiness Studies*, 10(4), pp.459-481.

[62•] - Nerstad C, Richardsen A, Martinussen M. (2010). Factorial validity of the Utrecht Work Engagement Scale (UWES) across occupational groups in Norway. Scandinavian Journal of Psychology. Aug;51(4):326-33

[63•] - Simbula, S., Guglielmi, D., Schaufeli, W. B. & Depolo, M. (2013). An Italian validation of the Utrecht Work Engagement Scale: Characterization of engaged groups in a sample of schoolteachers. Bollettino Di Psicologia Applicata, 268, 43-54

[64•] - Mills, M., Culbertson, S. and Fullagar, C. (2011). Conceptualizing and Measuring Engagement: An Analysis of the Utrecht Work Engagement Scale. *Journal of Happiness Studies*, 13(3), pp.519-545.

[65•] - Fong T, Ng S. (2011). Measuring Engagement at Work: Validation of the Chinese Version of the Utrecht Work Engagement Scale. Int.J. Behav. Med. 19(3):391-397.

[66•] - Zecca, G., Györkös, C., Becker, J., Massoudi, K., de Bruin, G. and Rossier, J. (2015). Validation of the French Utrecht Work Engagement Scale and its relationship with personality traits and impulsivity. *Revue Européenne de Psychologie Appliquée/European Review of Applied Psychology*, 65(1), pp.19-28.

[67•] - Schaufeli W., & Bakker A. (2003). Utrecht Work Engagement Scale Preliminary Manual, Version 1. Occupational Health Psychology Unit, Utrecht University

[68•] - Shimazu, A., Schaufeli, W., Kosugi, S., Suzuki, A., Nashiwa, H., Kato, A., Sakamoto, M., Irimajiri, H., Amano, S., Hirohata, K., Goto, R. and Kitaoka-Higashiguchi, K. (2008). Work Engagement in Japan: Validation of the Japanese Version of the Utrecht Work Engagement Scale. *Applied Psychology*, 57(3), pp.510-523.

[69•] - Viljevac, A., Cooper-Thomas, H. and Saks, A. (2012). An investigation into the validity of two measures of work engagement. *The International Journal of Human Resource Management*, 23(17), pp.3692-3709.

[70•] - de Bruin, G. and Henn, C. (2013). Dimensionality of the 9-item Utrecht Work Engagement Scale (UWES–9). *Psychological Reports*, p.130729135325001.

[71•••••] - Villotti, P., Balducci, C., Zaniboni, S., Corbiere, M. and Fraccaroli, F. (2013). An Analysis of Work Engagement Among Workers With Mental Disorders Recently Integrated to Work. *Journal of Career Assessment*, 22(1), pp.18-27.

[72•] - Fong T, Ng S. (2011). Measuring Engagement at Work: Validation of the Chinese Version of the Utrecht Work Engagement Scale. Int.J. Behav. Med. 19(3):391-397.

[73•] - Schaufeli, W. B., Bakker, A. B., & Salanova, M. (2006) The measurement of work engagement with a short questionnaire: a cross-national study. *Educational and Psychological Measurement,* 66(4), 701-716

[74•] - Wefald, A., Mills, M., Smith, M. and Downey, R. (2011). A Comparison of Three Job Engagement Measures: Examining their Factorial and Criterion-Related Validity. *Applied Psychology: Health and Well-Being*, 4(1), pp.67-90.

[75•] - Panthee, B., Shimazu, A. and Kawakami, N. (2014). Validation of Nepalese Version of Utrecht Work Engagement Scale. *Journal of Occupational Health*, 56(6), pp.421-429.

[76•] - Klassen, R., Aldhafri, S., Mansfield, C., Purwanto, E., Siu, A., Wong, M. and Woods-McConney, A. (2012). Teachers' Engagement at Work: An International Validation Study. *The Journal of Experimental Education*, 80(4), pp.317-337.

[77•] - Balducci, C., Fraccaroli, F. and Schaufeli, W. (2010). Psychometric Properties of the Italian Version of the Utrecht Work Engagement Scale (UWES-9). *European Journal of Psychological Assessment*, 26(2), pp.143-149.

[78•] - Littman-Ovadia, H. and Balducci, C. (2013). Psychometric Properties of the Hebrew Version of the Utrecht Work Engagement Scale (UWES-9). *European Journal of Psychological Assessment*, 29(1), pp.58-63.

[79•] - Clark, Cynthia M.; Landrum, R. Eric; and Nguyen, Danh T. (2013). Development and Description of the Organizational Civility Scale (OCS) *Journal of Theory Construction & Testing, 17*(1), 11-17.

[80*] - The EEF: The Manufacturer's Organisation. Work Organisation Assessment Questionnaire: A Tool for the Risk Management of Stress. London: EEF; 2004.

[81•] - Wynne-Jones, G., Varnava, A., and Buck, R., (2009). Examination of the Work Organization Assessment Questionnaire in Public Sector Workers, Journal of Occupational and Environmental Medicine, Issue: Volume 51(5), pp 586-593

[82•] - Karimi, L. and Meyer, D. (2015). Validity and model-based reliability of the Work Organisation Assessment Questionnaire among nurses. *Nursing Outlook*, 63(3), pp.318-330.

[83•] - Griffiths A, Cox T, Karanika M, Khan S, Tomás JM. (2006). Work design and management in the manufacturing sector: development and validation of the Work Organisation Assessment Questionnaire. *Occup Environ Med*. 63:669–675

[84••] - Symonds T, Burton A, Tillotson K, Main C. (1996). Do attitudes and beliefs influence work loss due to low back trouble? Occupational Medicine. 46:25-32.

[85•] - Chungkham, H., Ingre, M., Karasek, R., Westerlund, H. and Theorell, T. (2013). Factor Structure and Longitudinal Measurement Invariance of the Demand Control Support Model: An Evidence from the Swedish Longitudinal Occupational Survey of Health (SLOSH). *PLoS ONE*, 8(8), p.e70541.

[86•] - Hökerberg, Y., Reichenheim, M., Faerstein, E., Passos, S., Fritzell, J., Toivanen, S. and Westerlund, H. (2014). Cross-cultural validity of the demand-control questionnaire: Swedish and Brazilian workers. *Revista de Saúde Pública*, 48(3), pp.486-496.

[87•] - Sanne, B., Torp, S., Mykletun, A. and Dahl, A. (2005). The Swedish Demand-Control-Support Questionnaire (DCSQ): Factor structure, item analyses, and internal consistency in a large population. *Scandinavian Journal of Public Health*, 33(3), pp.166-174.

[88•] - Mase, J., Ota, A., Inoue, K., Iida, T., Tsutsumi, A., Yatsuya, H. and Ono, Y. (2012). Reliability and Validity of the Japanese Translated Version of the Swedish Demand-Control-Support Questionnaire. *Industrial Health*, 50(6), pp.467-475.

[89•] - Abma, F., Brouwer, S., de Vries, H., Arends, I., Robroek, S., Cuijpers, M., van der Wilt, G., Bültmann, U. and van der Klink, J. (2016). The capability set for work: development and validation of a new questionnaire. *Scandinavian Journal of Work, Environment & Health*, 42(1), pp.34-42.

[90•] - Comer, J.M., Ramsey, R., Lassk, F.G. and Marshall, G.W. (1995). Methods in sales research: A critical evaluation of a measure of job involvement: The use of the Lodahl and Kejner (1965) scale with salespeople. *Journal of personal selling & Sales Management*, *15*(3), pp.65-74.

[91•] - Lodahl T, Kejnar M. (1965). The definition and measurement of job involvement. Journal of Applied Psychology . 49(1):24-33.

[92••] - Holdena J., Davidson M., and Tam J. (2010). Can the Fear-Avoidance Beliefs Questionnaire predict work status in people with work-related musculoskeltal disorders?, Journal of Back and Musculoskeletal Rehabilitation. 23:201–208

[93••] - Ketenci A., Kesiktas N., Sindel D., and Disci R. (2014). No. 244 Validation of the Turkish Version of the Fear Avoidance Belief Questionnaire. *PM&R*, 6(8), p.S140.

[94••] - George, S., Valencia, C. and Beneciuk, J. (2010). A Psychometric Investigation of Fear-Avoidance Model Measures in Patients With Chronic Low Back Pain. *J Orthop Sports Phys Ther*, 40(4), pp.197-205.

[95••••••] - Inrig T, Amey B, Borthwick C, Beaton D. (2011). Validity and Reliability of the Fear-Avoidance Beliefs Questionnaire (FABQ) in Workers with Upper Extremity Injuries. Journal of Occupational Rehabilitation. 22(1):59-70.

[96••] - Swinkels-Meewisse, E.J.C.M., Swinkels, R.A.H.M., Verbeek, A.L.M., Vlaeyen, J.W.S. and Oostendorp, R.A.B. (2003) Psychometric properties of the Tampa Scale for kinesiophobia and the fear-avoidance beliefs questionnaire in acute low back pain. *Manual therapy*, *8*(1), pp.29-36.

[97•••] - Dedering, Å. and Börjesson, T. (2012). Assessing Fear-avoidance Beliefs in Patients with Cervical Radiculopathy. *Physiotherapy Research International*, 18(4), pp.193-202.

[98••] - Kovacs F, Muriel A, Medina J, Abraira V, Sanchez M, Jauregui J. (2006). Psychometric Characteristics of the Spanish Version of the FAB Questionnaire. Spine. 31(1):104-110.

[99••] - de Souza, F., da Silva Marinho, C., Siqueira, F., Maher, C. and Costa, L. (2008). Psychometric Testing Confirms That the Brazilian-Portuguese Adaptations, the Original Versions of the Fear-Avoidance Beliefs Questionnaire, and the Tampa Scale of Kinesiophobia Have Similar Measurement Properties. *Spine*, 33(9), pp.1028-1033.

[100••] -Cleland JA, Fritz JM, Childs JD, (2008). Validation of the French Version of the Fear Avoidance Belief Questionnaire, Am J Phys Med Rehabil. 87(2):109-17.

[101••] - Mintken, P., Cleland, J., Whitman, J. and George, S. (2010). Psychometric Properties of the Fear-Avoidance Beliefs Questionnaire and Tampa Scale of Kinesiophobia in Patients With Shoulder Pain. *Archives of Physical Medicine and Rehabilitation*, 91(7), pp.1128-1136.

[102••] - Pfingsten M, Kröner-Herwig B, Leibing E, Kronshage U, Hildebrandt J. (2004). Validation of the German version of the Fear-Avoidance Beliefs Questionnaire (FABQ). European Journal of Pain. 4(3):259-266.

[103••] - Laufer Y., Abu Elheiga-Na'amne B., and Rozen N. (2012). Translation and validation of the Arab version of the fear avoidance beliefs questionnaire, Journal of Back and Musculoskeletal Rehabilitation. 25:201–208

[104••] - Askary-Ashtiani A, Ebrahimi-Takamejani I, Torkaman G, Amiri M, Mousavi S. (2014). Reliability and Validity of the Persian Versions of the Fear Avoidance Beliefs Questionnaire and Tampa Scale of Kinesiophobia in Patients With Neck Pain. Spine. 39(18):E1095-E1102.

[105••] - Terho, H., Haapea, M., Paananen, M., Korniloff, K., Häkkinen, A. and Karppinen, J. (2016). Translation and validation of the Finnish version of the Fear-Avoidance Beliefs Questionnaire (FABQ). *Scandinavian Journal of Pain*, 10, pp.113-118.

[106••] - Chaory, K., Fayad, F., Rannou, F., Lefevre-Colau, M.M., Fermanian, J., Revel, M. and Poiraudeau, S., (2004) Validation of the French version of the fear avoidance belief questionnaire. *Spine*, *29*(8), pp.908-913.

[107••] - Waddell, G., Newton, M., Henderson, I., Somerville, D. and Main, C.J. (1993) A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. *Pain*, *52*(2), pp.157-168.

[108••] - Matsudaira K, Kikuchi N, Murakami A, Isomura T. (2013). Psychometric properties of the Japanese version of the Fear-Avoidance Beliefs Questionnaire (FABQ). Journal of Orthopaedic Science. 19(1):26-32.

[109••] - Korkmaz, N., Akinci, A., Yörükan, S., Sürücü, H.S., Saraçbaşi, O. and Ozçakar, L., (2009) Validation and reliability of the Turkish version of the fear avoidance beliefs questionnaire in patients with low back pain. *European journal of physical and rehabilitation medicine*, *45*(4), pp.527-535.

[110••] - Pei L.B., Xia J.J. and Yan J.L. (2010).Cross-Cultural Adaptation, Reliability and Validity of the Chinese Version of the Fear Avoidance Beliefs Questionnaire, Journal of International Medical Research. 38:1985-1996

[111••] - Rostami M., Noorian N., Ali Mansourni M., and Sharafi E., (2014). Validation of the Persian version of the fear avoidance belief questionnaire in patients with low back pain, Journal of Back and Musculoskeletal Rehabilitation. 27:213–221

[112••] - Staerkle, R., Mannion, A.F., Elfering, A., Junge, A., Semmer, N.K., Jacobshagen, N., Grob, D., Dvorak, J. and Boos, N., (2004) Longitudinal validation of the fear-avoidance beliefs questionnaire (FABQ) in a Swiss-German sample of low back pain patients. *European Spine Journal*, *13*(4), pp.332-340.

[113••] - Monticone, M., Baiardi, P., Bonetti, F., Ferrari, S., Foti, C., Pillastrini, P., Rocca, B., Vanti, C. and Zanoli, G., (2012) The Italian version of the Fear-Avoidance Beliefs Questionnaire (FABQ-I): cross-cultural adaptation, factor analysis, reliability, validity, and sensitivity to change. *Spine*, *37*(6), pp.E374-E380.

[114••] - Grotle, M., Brox, J.I. and Vøllestad, N.K. (2006) Reliability, validity and responsiveness of the fear-avoidance beliefs questionnaire: methodological aspects of the Norwegian version. *Journal of Rehabilitation Medicine*, *38*(6), pp.346-353.

[115••] – Lee K., Chiu T., and Lam T. (2006). Psychometric properties of the Fear-Avoidance Beliefs Questionnaire in patients with neck pain, Clin Rehabil. 20: 909-920

[116••] - Hildebrandt V, Bongers P, van Dijk F, Kemper H, Dul J. (2001). Dutch Musculoskeletal Questionnaire: description and basic qualities. Ergonomics. 44(12):1038-1055.

[117••] - Hedge, A., Morimoto, S. And McCrobie, D. (1999) Effects of keyboard tray geometry on upper body posture and comfort, Ergonomics, 42 (10), 1333-1349.

[118•] - Erdinc, O., Hot, K. and Ozkaya, M. (2011). Turkish version of the Cornell Musculoskeletal Discomfort Questionnaire: Cross-cultural adaptation and validation. *Work*, *39*(3), pp.251-260.

[119•] - Lalic H., and Hromin M., (2012). Presenteeism Towards Absenteeism: Manual Work Versus Sedentary Work, Private Versus Governmental – A Croatian Review. Coll. Antropol. 36(1): 111–116

[120••] - Truchon, M., Schmouth, M.È., Côté, D., Fillion, L., Rossignol, M. and Durand, M.J. (2012). Absenteeism screening questionnaire (ASQ): a new tool for predicting long-term absenteeism among workers with low back pain. *Journal of occupational rehabilitation*, *22*(1), pp.27-50.

[121•] - Gawlicki, M., Reilly, M., Popielnicki, A. and Reilly, K. (2006). Linguistic Validation of the US Spanish Work Productivity and Activity Impairment Questionnaire, General Health Version. *Value in Health*, 9(3), pp.199-204.

[122••] - Zhang W., Bansback N., Boonen A. (2010). Validity of the work productivity and activity impairment questionnaire - general health version in patients with rheumatoid arthritis; *Arthritis Research & Therapy,* 12(5):R177

[123•••••] - Reilly MC, Zbrozek AS, Dukes EM. (1993). The validity and reproducibility of a work productivity and activity impairment instrument. Pharmacoeconomics, 4:353-365)

[124••] - Ciconelli R, Soárez P, Kowalski C, Ferraz M. (2006). The Brazilian Portuguese version of the Work Productivity and Activity Impairment: General Health (WPAI-GH) Questionnaire. Sao Paulo Med. J. 124(6):325-332.

[125•] - Giovannetti E,, Wolff J., Frick K., and Boult C. (2009). Construct Validity of the Work Productivity and Activity Impairment Questionnaire across Informal Caregivers of Chronically Ill Older Patients; Value Health. 12(6): 1011–1017

[126••••••] - Wahlqvist, P., Carlsson, J., Stålhammar, N. and Wiklund, I. (2002). Validity of a Work Productivity and Activity Impairment Questionnaire for Patients with Symptoms of Gastro-Esophageal Reflux Disease (WPAI-GERD)—Results from a Cross-Sectional Study. *Value in Health*, 5(2), pp.106-113.

[127••] - Lambert, J., Hansen, B., Arnould, B., Grataloup, G., Guillemin, I., Højbjerre, L., Strandberg-Larsen, M. and Reilly, M. (2014). Linguistic Validation into 20 Languages and Content Validity of the Rheumatoid Arthritis-Specific Work Productivity and Activity Impairment Questionnaire. *The Patient - Patient-Centered Outcomes Research*, 7(2), pp.171-176.

[128••••••] - Vergara, M., Montserrat, A., Casellas, F., Villoria, A., Suarez, D., Maudsley, M., Gallardo, O., Ricart, E. and Calvet, X. (2011). A new validation of the Spanish Work Productivity and Activity Impairment Questionnaire—Crohn's disease version. *Value in Health*, *14*(6), pp.859-861.

[129••] - Ozcan, E., Specialist, S., Sen, E., Rezvani, A., Baysak, T. and Prof, A. (2014). Poster 67 Turkish Adaptation and Validity of the Work Productivity and Activity Impairment Questionnaire in Ankylosing Spondylitis. *PM&R*, 6(9), p.S205.

[130••••••] - Wahlqvist, P., Guyatt, G., Armstrong, D., Degl'Innocenti, A., Heels-Ansdell, D., El-Dika, S., Wiklund, I., Fallone, C., Tanser, L., Veldhuyzen van Zanten, S., Austin, P., Barkun, A., Chiba, N. and Scheunemann, H. (2007). The Work Productivity and Activity Impairment Questionnaire for Patients with Gastroesophageal Reflux Disease (WPAI-GERD). *PharmacoEconomics*, 25(5), pp.385-396.

[131••] - Reilly, M.C., Gooch, K.L., Wong, R.L., Kupper, H. and Van der Heijde, D. (2010). Validity, reliability and responsiveness of the Work Productivity and Activity Impairment Questionnaire in ankylosing spondylitis. *Rheumatology*, *49*(4), pp.812-819.

[132••] - Reilly, M., Gooch, K., Wong, R., Kupper, H. and van der Heijde, D. (2010). Validity, reliability and responsiveness of the Work Productivity and Activity Impairment Questionnaire in ankylosing spondylitis. *Rheumatology*, 49(4), pp.812-819.

[133••••••] - Vergara, M., Montserrat, A., Casellas, F., Maudsley, M., Gallardo, O., Ricart, E. and Calvet, X. (2009). Validation of the Spanish Work Productivity and Activity impairment questionnaire: Crohn's disease version. *European Journal of Gastroenterology & Hepatology*, 21(7), pp.809-815.

[134••••••] - Reilly, M., Bracco, A., Ricci, J., Santoro, J. and Stevens, T. (2004). The validity and accuracy of the Work Productivity and Activity Impairment questionnaire - irritable bowel syndrome version (WPAI:IBS). *Alimentary Pharmacology and Therapeutics*, 20(4), pp.459-467.

[135••••••] - Reily M., Gerlier L., Brabant Y., and Brown M. (2008). Validity, reliability, and responsiveness of the work productivity and activity impairment questionnaire in Crohn's disease; Clinical Therapeutics. 30(2): 393–404

[136••] - Kigozi J, Lewis M, Jowett S, Barton P, Coast J. (2014). Construct Validity and Responsiveness of the Single-Item Presenteeism Question in Patients With Lower Back Pain for the Measurement of Presenteeism. Spine. 39(5):409-416.

[137••••••] - Pumpaisalchai, W., Ruengorn, C., Karahong, K., Jamroenkhajonsuk, P., Pongdoung, T. and Udombhornprabha, A. (2013). Reliability and Validity of a Thai Version of Lam Employment Absence and Productivity Scale (LEAPS). *Value in Health*, 16(7), p.A596.

[138•••••] - Lam, R. (2014). Lam employment absence and productivity scale (LEAPS): Further validation studies in major depressive disorder. *Value in Health*, 17(3), p.A195

[139•••••] - Lam R., Michalak E., Yatham L., (2009). A new clinical rating scale for work absence and productivity: validation in patients with major depressive disorder; BMC Psychiatry. 9(1):78.

[140••••••] - Steiner TJ. (2007). The HALT and HART indices. J Headache Pain. 8 suppl 1: S22-S25

[140••] - Meerding W, IJzelenberg W, Koopmanschap M, Severens J, Burdorf A. (2005). Health problems lead to considerable productivity loss at work among workers with high physical load jobs. Journal of Clinical Epidemiology. 58(5): 517-523.

[141••••••] - Brouwer WB, Koopmanschap MA, Rutten FF. (1999). Productivity losses without absence: measurement validation and empirical evidence. Health Policy. 48:13–27.

[143••] Puolakka K, Kautiainen H, Mottonen T, Hannonen P, Korpela M, Hakala M, Viikari-Juntura E, Solovieva S, Arkela-Kautiainen M, Leirisalo-Repo M. (2009). A mismatch between self-reported physical work load and the HAQ: early identification of rheumatoid arthritis patients at risk for loss of work productivity. Clin Exp Rheumatol. 27:422–429.

[144••] - Durand, M., Vachon, B., Hong, Q., Imbeau, D., Amick, B. and Loisel, P. (2004). The cross-cultural adaptation of the Work Role Functioning Questionnaire in Canadian French. *International Journal of Rehabilitation Research*, 27(4), pp.261-268.

[145••••••] - Abma FI, Amick BC III, Brouwer S, van der Klink JJL, Bültmann U. (2012). The cross-cultural adaptation of the Work Role Functioning Questionnaire to Dutch. Work. 43;203–210

[146••] - Gallasch, C., Alexandre, N. and Amick, B. (2007). Cross-cultural Adaptation, Reliability, and Validity of the Work Role Functioning Questionnaire to Brazilian Portuguese. *Journal of Occupational Rehabilitation*, 17(4), pp.701-711.

[147••••••] - Ramada, J., Serra, C., Amick III, B., Castaño, J. and Delclos, G. (2013). Cross-Cultural Adaptation of the Work Role Functioning Questionnaire to Spanish Spoken in Spain. *Journal of Occupational Rehabilitation*, 23(4), pp.566-575.

[148•] - Ramada, J., Delclos, G., Amick, B., Abma, F., Pidemunt, G., Castaño, J., Bültmann, U. and Serra, C. (2014). Responsiveness of the Work Role Functioning Questionnaire (Spanish Version) in a General Working Population. *Journal of Occupational and Environmental Medicine*, 56(2), pp.189-194.

[149••••••] - Ramada, J., Serra, C., Amick, B., Abma, F., Castaño, J., Pidemunt, G., Bültmann, U. and Delclos, G. (2014). Reliability and Validity of the Work Role Functioning Questionnaire (Spanish Version). *Journal of Occupational Rehabilitation*, 24(4), pp.640-649.

[150••••••] - Abma, F., Dorland, H., Amick, B. and Bueltmann, U., (2015). Validation of the Work Role Functioning Questionnaire 2.0 in Cancer Patients. In *Psycho-Oncology* (vol. 24, pp. 151-151). 111 River st, Hoboken 07030-5774, NJ USA: Wiley-Blackwell.

[151•] - Abma F., Van der Klink J., (2013). The Work Role Functioning Questionnaire 2.0 (Dutch Version): Examination of its Reliability, Validity and Responsiveness in the General Working Population, J Occup Rehabil. 23:135–147

[152•••••] - Greig T., Nicholls S., Bryson G., and Bell M. (2004).The Vocational Cognitive Rating Scale: A scale for the assessment of cognitive functioning at work for clients with severe mental illness; Journal of Vocational Rehabilitation. 21: 71–81

[153•••••] - Hannula JA, Lahtela K, Jarvikoski A, Salminen JK, Makela P. (2006). Occupational Functioning Scale (OFS) An instrument for assessment of work ability in psychiatric disorders. Nord J Psychiatry. 60:372378

[154•••••] - Griffiths R. (1973). A Standardized Assessment of the Work Behaviour of Psychiatric Patients. The British Journal of Psychiatry. 123(4):403-408.

[155•] - Millington M, Leierer S, Abadie M. (2000). Validity and the Employment Expectation Questionnaire: Do Disability-Related Attitudes Affect Employment Selection Outcomes?. Rehabilitation Counseling Bulletin. 44(1):39-47.

[156••••••] - Finger, M.E., Glässel, A., Erhart, P., Gradinger, F., Klipstein, A., Rivier, G., Schröer, M., Wenk, C., Gmünder, H.P., Stucki, G. and Escorpizo, R. (2011). Identification of relevant ICF categories in vocational rehabilitation: a cross sectional study evaluating the clinical perspective. *Journal of occupational rehabilitation*, *21*(2), pp.156-166.

[157••] - Osterhaus J, Purcaru O, Richard L. (2009). Discriminant validity, responsiveness and reliability of the rheumatoid arthritis-specific Work Productivity Survey (WPS-RA). Arthritis Res Ther. 11(3):R73.

[158••] - Osterhaus, J. and Purcaru, O. (2014). Discriminant validity, responsiveness and reliability of the arthritis-specific Work Productivity Survey assessing workplace and household productivity within and outside the home in patients with axial spondyloarthritis, including nonradiographic axial spondyloarthritis and ankylosing spondylitis. *Arthritis Res Ther*, 16(4), p.R164

[159••] - Osterhaus, J. and Purcaru, O. (2014). Discriminant validity, responsiveness and reliability of the arthritis-specific Work Productivity Survey assessing workplace and household productivity in patients with psoriatic arthritis. *Arthritis Res Ther*, 16(4), p.R140.

[160*] - Sheehan DV. The Sheehan Disability Scales. In The Anxiety Disease and How to Overcome It. New York: Charles Scribner and Sons, 1983, p. 151

[161••••••] - Hambrick J, Turk C, Heimberg R, Schneier F, Liebowitz M. (2004). Psychometric properties of disability measures among patients with social anxiety disorder. Journal of Anxiety Disorders. 18(6):825-839.

[162••••••] - Amin-Esmaeili, M., Motevalian, A., Rahimi-Movaghar, A., Hajebi, A., Hefazi, M., Radgoodarzi, R. and Sharifi, V., (2014). The Translation and Psychometric Assessment of the Persian Version of the Sheehan Disability Scale. *Iranian journal of psychiatry*, *9*(3), p.125.

[163••••••] - Hodgins, D. (2013). Reliability and validity of the Sheehan disability scale modified for pathological gambling. *BMC Psychiatry*, 13(1), p.177.

[164•••••] - Cole T, Coon C, DeMuro C, McLeod L, Gnanasakthy A. (2014). Psychometric evaluation of the Sheehan Disability Scale in adult patients with attention-deficit/hyperactivity disorder. Neuropsychiatr Dis Treat. 19;10:887-95

[165••••••] - Luciano, J.V., Bertsch, J., Salvador-Carulla, L., Tomás, J.M., Fernández, A., Pinto-Meza, A., Haro, J.M., Palao, D.J. and Serrano-Blanco, A., (2010). Factor structure, internal consistency and construct validity of the Sheehan Disability Scale in a Spanish primary care sample. *Journal of evaluation in clinical practice*, *16*(5), pp.895-901.

[166••••••] - Leu, S.-H., Chou, J.-Y., Lee, P.-c., Cheng, H.-C., Shao, W.-C., Hsien, W.-L., Huang, C.-L. and Chen, V. C.-H. (2015), Validity and reliability of the Chinese version of the Sheehan Disability Scale (SDS-C). Asia-Pacific Psychiatry, 7: 215–222.

[167••••••] - Arbuckle, R., Frye, M.A., Brecher, M., Paulsson, B., Rajagopalan, K., Palmer, S. and Degl'Innocenti, A., (2009). The psychometric validation of the Sheehan Disability Scale (SDS) in patients with bipolar disorder. *Psychiatry research*, *165*(1), pp.163-174.

[168*] - Hakkaart-vanRoijen L, Short Form- Health and Labour Questionnaire. Institute for Medical Technology Assessment, Erasmus University Rotterdam.

[169•] - David B, Morgan K. (2007). Measuring the Occupational Impact of Sleep Quality: the OISQ. Sleep Biol Rhythms 5, A160.

[170••••••] - Halpern M., Shikiar R., Rentz A., Khan Z. (2001). Impact of smoking status on workplace absenteeism and productivity, Tobacco Control. 10:233–238

[171•] - von Thiele Schwarz, U., Sjöberg, A., Hasson, H. and Tafvelin, S. (2014). Measuring Self-Rated Productivity. *Journal of Occupational and Environmental Medicine*, 56(12), pp.1302-1307.

[172••••••] - Shikiar R, Rentz A, Halpern M, Khan Z. (2001). The health and work questionnaire (HWQ): An instrument for assessing workplace productivity in relation to worker health. Value in Health. 4(2):181.

[173••••] - Malec, J, Smigielski, J, DePompolo, R, and Thompson, J. (1993). Outcome evaluation and prediction in a comprehensive-integrated postacute outpatient brain injury rehabilitation program. *Brain Inj*. 7: 15–29

[174*] - Vanti C, Prosperi D, Boschi M. (2013. The Prolo Scale: history, evolution and psychometric properties. J Orthopaed Traumatol. 14(4):235-245.

[175•] – Kessler RC, Barber C, Beck A, Berglund PA, Cleary PD, McKenas D, Pronk N, Simon G, Stang P, Ustün TB, Wang PS. (2003). The World Health Organization Health and Work Performance Questionnaire (HPQ). J Occup Environ Med. 45: 156–74.

[176•••••] – Ellenberger D., Friede T., Sterz C., Flachenecke P. (2015). Validation of the German version of the multiple sclerosis work difficulties questionnaire (MSWDQ). *Multiple Sclerosis Journal*, 21(11 Suppl), pp.76-653.

[177•••••] - Honan, C.A., Brown, R.F., Hine, D.W., Vowels, L., Wollin, J.A., Simmons, R.D. and Pollard, J.D. (2012). The multiple sclerosis work difficulties questionnaire. *Multiple Sclerosis Journal*, *18*(6), pp.871-880.

[178•••••] - Honan, C., Brown, R. and Hine, D. (2014). The Multiple Sclerosis Work Difficulties Questionnaire (MSWDQ): development of a shortened scale. *Disability and Rehabilitation*, 36(8), pp.635-641.

[179*] - Bouwmans C, Krol M, Brouwer W, Severens J, Koopmanschap M, Hakkaart L.(2014). IMTA Productivity Cost Questionnaire (IPCQ). Value in Health. 17(7):A550.

[180••••••] - Lam C, Wiley A, Siu A, Emmett J., (2010). Assessing readiness to work from a stages of change perspective: implications for return to work, Work. 37(3):321-9

[181••••••] - Chan H, Li-Tsang C, Chan C, Lam C, Hui K, Bard C. (2006). Validation of Lam assessment of employment readiness (C-LASER) for Chinese injured workers. Journal of Occupational Rehabilitation. 16(4):697-705.

[182••] - Franche R, Corbière M, Lee H, Breslin F, Hepburn C. (2007). The Readiness for Return-To-Work (RRTW) scale: Development and Validation of a Self-report Staging Scale in Lost-time Claimants with Musculoskeletal Disorders. Journal of Occupational Rehabilitation. 17(3):450-472.

[183••••••] - Braathen T, Brage S, Tellnes G, Eftedal M. (2012). Psychometric Properties of the Readiness for Return to Work Scale in Inpatient Occupational Rehabilitation in Norway. Journal of Occupational Rehabilitation. 23(3):371-380.

[184••] - Laranjeira C. (2013). Validation of the Portuguese version of the Stanford Presenteeism Scale in nurses. International Journal of Nursing Practice. 19(6):644-650.

[185••] - Frauendorf R, de Medeiros Pinheiro M, Ciconelli R. (2013) Translation into Brazilian Portuguese, cross-cultural adaptation and validation of the Stanford presenteeism scale-6 and work instability scale for ankylosing spondylitis. Clin Rheumatol. 33(12):1751-1757.

[186•] - Koopman, C., Pelletier, K.R., Murray, J.F., Sharda, C.E., Berger, M.L., Turpin, R.S., Hackleman, P., Gibson, P., Holmes, D.M. and Bendel, T. (2002). Stanford presenteeism scale: health status and employee productivity. *Journal of Occupational and Environmental Medicine*, *44*(1), pp.14-20.

[187••] - Hutting N, Engels J, Heerkens Y, Staal J, Nijhuis-van der Sanden M. Development and Measurement Properties of the Dutch Version of the Stanford Presenteeism Scale (SPS-6). Journal of Occupational Rehabilitation . 2013;24(2):268-277.

[188••••••] - Turpin, R.S., Ozminkowski, R.J., Sharda, C.E., Collins, J.J., Berger, M.L., Billotti, G.M., Baase, C.M., Olson, M.J. and Nicholson, S. (2004). Reliability and validity of the Stanford Presenteeism Scale. *Journal of Occupational and Environmental Medicine*, *46*(11), pp.1123-1133.

[189••] - Marhold C., Linton S., Melin L., (2002). Identification of Obstacles for Chronic Pain Patients to Return to Work: Evaluation of a Questionnaire, Journal of Occupational Rehabilitation. 12(2):65-75.

[190•] - Schyns B, von Collani G. (2002). A new occupational self-efficacy scale and its relation to personality constructs and organizational variables. European Journal of Work and Organizational Psychology. 11(2):219-241.

[191••••••] - Endicott J., Nee J., (1997). Endicott Work Productivity Scale (EWPS): A New Measure to Assess Treatment Effects; Psychopharmacology Bulletin. 33(1): pp 13-16

[192•] - Gärtner, F., Nieuwenhuijsen, K., van Dijk, F. and Sluiter, J. (2012). Interpretability of change in the Nurses Work Functioning Questionnaire: minimal important change and smallest detectable change. *Journal of Clinical Epidemiology*, 65(12), pp.1337-1347.

[193•] - Gärtner, F., Nieuwenhuijsen, K., van Dijk, F. and Sluiter, J. (2011). Psychometric Properties of the Nurses Work Functioning Questionnaire (NWFQ). *PLoS ONE*, 6(11), p.e26565.

[194•] - Gärtner, F., Nieuwenhuijsen, K., van Dijk, F. and Sluiter, J. (2011). Impaired work functioning due to common mental disorders in nurses and allied health professionals: the Nurses Work Functioning Questionnaire. *Int Arch Occup Environ Health*, 85(2), pp.125-138.

[195•••••] - Lerner, D.J., Amick III, B.C., Malspeis, S., Rogers, W.H., Santanello, N.C., Gerth, W.C. and Lipton, R.B., (1999). The migraine work and productivity loss questionnaire: concepts and design. *Quality of Life Research*, *8*(8), pp.699-710.

[196••••] - Buffington A, Malec J. (1997). The vocational rehabilitation continuum: maximizing outcomes through bridging the gap from hospital to community-based services. J Head Trauma Rehabil;12(5):1-13

[197*] - Centre for Health Evaluation and Outcome Sciences. Valuation of Lost Productivity Questionnaire. Available at: www.thevolp.com. Accessed on : 01/07/2014

[198••] - Zhang W, Bansback N, Kopec J, Anis AH. (2011). Measuring time input loss among patients with rheumatoid arthritis: validity and reliability of the Valuation of Lost Productivity questionnaire. J Occup Environ Med;53(5):530-6.

[199••••] - Kuipers P, Kendall M, Fleming J, Tate R. (2004). Comparison of the Sydney Psychosocial Reintegration Scale (SPRS) with the Community Integration Questionnaire (CIQ): psychometric properties. Brain Inj. 18(2):161-177.

[200••••] - Tate R, Hodgkinson A, Veerabangsa A, Maggiotto S. (1999). Measuring Psychosocial Recovery after Traumatic Brain Injury: Psychometric Properties of a New Scale. Journal of Head Trauma Rehabilitation. 14(6):543-557.

[201•••] - De Wolf A, Lane-Brown A, Tate R, Middleton J, Cameron I. (2010). Measuring community integration after spinal cord injury: validation of the Sydney psychosocial reintegration scale and community integration measure. Qual Life Res. 19(8):1185-1193.

[202••] - Addley K, Burke C, McQuillan P. (2010). Impact of a direct access occupational physiotherapy treatment service. Occupational Medicine. 60(8):651-653.

[203*] - Wolf T.,Occupational Outcome Questionnaire, Program in Occupational Therapy, Washington University School of Medicine, St. Louis, MO, USA

[204•] - Bova, N., De Jonge, J. and Guglielmi, D. (2013). The Demand-Induced Strain Compensation Questionnaire: A Cross-national Validation Study. *Stress Health*, 31(3), pp.236-244.

[205••] - Gilworth, G., Chamberlain, M.A., Harvey, A., Woodhouse, A., Smith, J., Smyth, M.G. and Tennant, A., (2003). Development of a work instability scale for rheumatoid arthritis. *Arthritis Care & Research*, *49*(3), pp.349-354.

[206••] – Beaton, D.E., Tang, K., Gignac, M.A., Lacaille, D., Badley, E.M., Anis, A.H. and Bombardier, C. (2010). Reliability, validity, and responsiveness of five at-work productivity measures in patients with rheumatoid arthritis or osteoarthritis. *Arthritis care & research*, *62*(1), pp.28-37.

[207••] - Roy, J.S., MacDermid, J.C., Amick, B.C., Shannon, H.S., McMurtry, R., Roth, J.H., Grewal, R., Tang, K. and Beaton, D., (2011). Validity and responsiveness of presenteeism scales in chronic work-related upper-extremity disorders. *Physical therapy*, *91*(2), pp.254-266.

[208••] - Revicki, D., Ganguli, A., Kimel, M., Roy, S., Chen, N., Safikhani, S. and Cifaldi, M. (2015). Reliability and Validity of the Work Instability Scale for Rheumatoid Arthritis. *Value in Health*, 18(8), pp.1008-1015.

[209••] - Gilworth, G., Emery, P., Barkham, N., Smyth, M.G., Helliwell, P. and Tennant, A., (2009). Reducing work disability in Ankylosing Spondylitis–development of a work instability scale for AS. *BMC musculoskeletal disorders*, *10*(1), p.1.

[210•] - Gilworth, G., Bhakta, B., Eyres, S., Carey, A., Anne Chamberlain, M. and Tennant, A., (2007). Keeping nurses working: development and psychometric testing of the Nurse-Work Instability Scale (Nurse-WIS). *Journal of advanced nursing*, *57*(5), pp.543-551.

[211•] - Harling, M., Schablon, A., Peters, C. and Nienhaus, A. (2014). Predictive values and other quality criteria of the German version of the Nurse-Work Instability Scale (Nurse-WIS) - follow-up survey findings of a prospective study of a cohort of geriatric care workers. *Journal of Occupational Medicine and Toxicology*, 9(1).

[212•] - Harling M, Schablon A, Nienhaus A. (2013). Validation of the German version of the Nurse-Work Instability Scale: baseline survey findings of a prospective study of a cohort of geriatric care workers. Journal of Occupational Medicine and Toxicology. 8(1):33.

[213•] - Harling, M., Schablon, A. and Nienhaus, A. (2013). Validation of the German version of the Nurse-Work Instability Scale: baseline survey findings of a prospective study of a cohort of geriatric care workers. *Journal of Occupational Medicine and Toxicology*, 8(1), p.33.

[214••••] - Gilworth, G., Carey, A., Eyres, S., Sloan, J., Rainford, B., Bodenham, D., Neumann, V. and Tennant, A., (2006). Screening for job loss: development of a work instability scale for traumatic brain injury. *Brain Injury*, *20*(8), pp.835-843.

[215•••••] - McFadden E, Horton M, Ford H, Gilworth G, McFadden M, Tennant A. (2011). Screening for the risk of job loss in multiple sclerosis (MS): development of an MS-specific Work Instability Scale (MS-WIS). Multiple Sclerosis Journal . 18(6):862-870.

[216*] - Epilepsy - Work Instability Scale, Psychometric Laboratory for Health Sciences, University of Leeds

[217*] - Manual Work Instability Scale, Psychometric Laboratory for Health Sciences, University of Leeds

[218*] - The Workplace Stress Scale, The Marlin Company, North Haven, CT, and the American Institute of Stress, Yonkers, NY ; Accessed at [http://teorionline.files.wordpress.com/2011/04/unit-3-the-workplace-stress-scale.pdf], Accessed on 20/07/2014

[219•] – Tabatabaee Jabil S., Ghaffari M, Pournik O, Ghalichi L, Tehrani Y., Motevalian S. (2013). Reliability and Validity of Persian Version of Job Content Questionnaire in Health Care Workers In Iran, Int J Occup Environ Med. 4(2):96-101

[220••••••] - Karasek R, Brisson C, Kawakami N, Houtman I, Bongers P, Amick B. (1998). The Job Content Questionnaire (JCQ): An instrument for internationally comparative assessments of psychosocial job characteristics. Journal of Occupational Health Psychology. 3(4):322-355.

[221•] - de Araújo T, Karasek R. (2008). Validity and reliability of the job content questionnaire in formal and informal jobs in Brazil. SJWEH Suppl. (6):52–59.

[222•] - Phakthongsuk P, Apakupakul N. (2008). Psychometric Properties of the Thai Version of the 22-Item and 45-Item Karasek Job Content Questionnaire. International Journal of Occupational Medicine and Environmental Health. 21(4).

[223•] - Poanta, L.I.; Zdrenghea, D.; Albu, A. (2006). Psychometric evaluation of Romanian version of job content questionnaire in physicians. Rom. J. Int. Med. 44, 183–199.

[224•] - Choobineh A., Ghaem H., Ahmedinejad P., (2011). Validity and reliability of the Persian (Farsi) version of the Job Content Questionnaire: a study among hospital nurses, Eastern Mediterranean Health Journal. 17:4; 335-341

[225•] - Hoang, T., Corbière, M., Negrini, A., Pham, M. and Reinharz, D. (2013). Validation of the karasek–job content questionnaire to measure job strain in Vietnam 1. *Psychological Reports*, 113(2), pp.363-379.

[226•] - Amin, N.A., Quek, K.F., Oxley, J.A., Noah, R.M. and Nordin, R., (2015). Validity and Reliability of Malay Version of the Job Content Questionnaire among Public Hospital Female Nurses in Malaysia. *The international journal of occupational and environmental medicine*, *6*(4 October), pp.632-232.

[227•] - Eum, K.D., Li, J., Jhun, H.J., Park, J.T., Tak, S.W., Karasek, R. and Cho, S.I., (2007). Psychometric properties of the Korean version of the job content questionnaire: data from health care workers. *International archives of occupational and environmental health*, *80*(6), pp.497-504.

[228•] – Kawakami N, Fujigaki Y. (1996). Reliability and Validity of the Japanese Version of Job Content Questionnaire: Replication and Extension in Computer Company Employees. Industrial Health. 34(4):295-306.

[229•] - Brisson, C., Blanchette, C., Guimont, C., Dion, G., Moisan, J., Vézina, M., Dagenais, G.R. and M [acaron] sse, L., 1998. Reliability and validity of the French version of the 18-item Karasek Job Content Questionnaire. *Work & Stress*, *12*(4), pp.322-336.

[230•] - Niedhammer, I. (2002). Psychometric properties of the French version of the Karasek Job Content Questionnaire: a study of the scales of decision latitude, psychological demands, social support, and physical demands in the GAZEL cohort. *International Archives of Occupational and Environmental Health*, 75(3), pp.129-144.

[231•] - Li W, Zhang JQ, Sun J, Tan PF, Wang S. (2007). Reliability and validity of Job Content Questionnaire in Chinese petrochemical employees. Psychol Rep. 2007;100:35–46.

[232•] - Żołnierczyk-Zreda, D. and Bedyńska, S. (2014). Psychometric Properties of the Polish Version of Karasek's Job Content Questionnaire. *International Journal of Occupational Safety and Ergonomics*, 20(4), pp.583-593.

[233•] - Alexopoulos, E., Argyriou, E., Bourna, V. and Bakoyannis, G. (2015). Reliability and Validity of the Greek Version of the Job Content Questionnaire in Greek Health Care Workers. *Safety and Health at Work*, 6(3), pp.233-239.

[234•] - Li J, Yang W, Liu P, Xu Z, Cho S. (2004). Psychometric Evaluation of the Chinese (Mainland) Version of Job Content Questionnaire: A Study in University Hospitals. Industrial Health. 42(2):260-267.

[235•] - Cheng Y, Luh W, Guo Y. (2003). Reliability and validity of the chinese version of the job content questionnaire in Taiwanese workers. Int. J. Behav. Med. 10(1):15-30

[236•] - Maizura H, Masilamani R, Aris T. (2009). Reliability (Internal Consistency) of the Job Content Questionnaire on Job Stress Among Office Workers of a Multinational Company in Kuala Lumpur. Asia-Pacific Journal of Public Health. 21(2):216-222.

[237•] - Hadi A, Naing N, Daud A, Nordin R.(2006). Reliability and construct validity of the Malay version of the Job Content Questionnaire (JCQ) among secondary school teachers in Kota Bharu, Kelantan, Malaysia. *Southeast Asian J Trop Med Publ Health*. 37(6):1254-1259.

[285•] - Nehzat F, Huda BZ, Tajuddin SH. (2014). Reliability and validity of job content questionnaire for university research laboratory staff in Malaysia. Mar;45(2):481-9.

[239•] - Choi, B., Ko, S., Dobson, M., Schnall, P.L., Garcia-Rivas, J., Israel, L. and Baker, D., 2014. Short-term test–retest reliability of the Job Content Questionnaire and Effort–Reward Imbalance Questionnaire items and scales among professional firefighters. *Ergonomics*, *57*(6), pp.897-911.

[240•] - d'Errico A, Punnett L, Gold J, Gore R. (2008). JCQ scale reliability and responsiveness to changes in manufacturing process. Am. J. Ind. Med. 51(2):138-147.

[241•] - Sale, J. and Kerr, M. (2002). The psychometric properties of Karasek's demand and control scales within a single sector: data from a large teaching hospital. *International Archives of Occupational and Environmental Health*, 75(3), pp.145-152.

[242•] - Choi B, Kurowski A , Bond M , Baker D , Clays E , De Bacquer D, Punnett L. (2012). Occupation-differential construct validity of the Job Content Questionnaire (JCQ) psychological job demands scale with physical job demands items: a mixed methods research, Ergonomics. 55:4, 425-439

[243•] - Pejtersen J, Kristensen T, Borg V, Bjorner J. (2009). The second version of the Copenhagen Psychosocial Questionnaire. Scandinavian Journal of Public Health. 38(3 Suppl):8-24.

[244•] - Moncada, S., Utzet, M., Molinero, E., Llorens, C., Moreno, N., Galtés, A. and Navarro, A., (2014). The copenhagen psychosocial questionnaire II (COPSOQ II) in Spain—A tool for psychosocial risk assessment at the workplace. *American journal of industrial medicine*, *57*(1), pp.97-107.

[245•] - Hanson, E., Schaufeli, W., Vrijkotte, T., Plomp, N. and Godaert, G. (2000). The validity and reliability of the Dutch Effort–Reward Imbalance Questionnaire. *Journal of Occupational Health Psychology*, 5(1), pp.142-155.

[246•] - Tsutsumi A, Ishitake T, Peter R, Siegrist J, Matoba T. (2001). The Japanese version of the Effort-Reward Imbalance Questionnaire: a study in dental technicians. Work & Stress. 15(1):86-96.

[247•] - Almadi T, Cathers I, Chow CM, (2013). An Arabic version of the effort-reward imbalance questionnaire: translation and validation study, Psychol Rep. Aug;113(1):1287-302.

[248•] - Yadegarfar, G., Alinia, T., Hosseini, R., Hassannejad, R., Fayaz, M., Sanati, J., Sanati, K., Harandi, J., Hajnoorozali, V., Baghi, M.R. and Mirzavand, E., (2013). Psychometric properties of the Farsi version of effort-reward imbalance questionnaire: a longitudinal study in employees of a synthetic fibre factory in Iran. *International archives of occupational and environmental health*, *86*(2), pp.147-155.

[248•] - Msaouel P., Keramaris N., Apostolopoulos A., (2012). The Effort-reward Imbalance Questionnaire in Greek: Translation, Validation and Psychometric Properties in Health Professionals, J Occup Health. 54: 119–130

[250••••••] - Aboa-Éboulé, C., Brisson, C., Blanchette, C., Maunsell, E., Bourbonnais, R., Abdous, B., Vézina, M., Milot, A. and Dagenais, G.R., (2011). Effort-reward imbalance at work and psychological distress: a validation study of post-myocardial infarction patients. *Psychosomatic medicine*, *73*(6), pp.448-455.

[251•] - Zurlo M, Pes D, Siegrist J. (2010). Validity and reliability of the effort-reward imbalance questionnaire in a sample of 673 Italian teachers. Int Arch Occup Environ Health . 83(6):665-674.

[252•] - Chor D, Werneck G, Faerstein E, Alves M, Rotenberg L. (2008). The Brazilian version of the effort-reward imbalance questionnaire to assess job stress. Cadernos de Saúde Pública . 24(1):219-224.

[253•] - Griep R, Rotenberg L, Vasconcellos A, Landsbergis P, Comaru C, Alves M. (2009). The psychometric properties of demand-control and effort–reward imbalance scales among Brazilian nurses. Int Arch Occup Environ Health 82(10):1163-1172.

[254•] - Buapetch A, Lagampan S, Faucett J, Kalampakorn S. (2008). The Thai Version of Effort-Reward Imbalance Questionnaire (Thai ERIQ): A Study of Psychometric Properties in Garment Workers. Journal of Occupational Health 50(6):480-491.

[255•] - Siegrist, J., Starke, D., Chandola, T., Godin, I., Marmot, M., Niedhammer, I. and Peter, R., (2004). The measurement of effort–reward imbalance at work: European comparisons. *Social science & medicine*, *58*(8), pp.1483-1499.

[256•] - Juárez-García, A., Vera-Calzaretta, A., Blanco-Gomez, G., Gómez-Ortíz, V., Hernández-Mendoza, E., Jacinto-Ubillus, J. and Choi, B. (2015). Validity of the effort/reward imbalance questionnaire in health professionals from six Latin-American countries. *Am. J. Ind. Med.*, 58(6), pp.636-649.

[257•] - Eum, K., Li, J., Lee, H., Kim, S., Paek, D., Siegrist, J. and Cho, S. (2007). Psychometric properties of the Korean version of the effort–reward imbalance questionnaire: a study in a petrochemical company. *Int Arch Occup Environ Health*, 80(8), pp.653-661.

[258•] - Gómez V. (2010) Assessment of psychosocial stressors at work: psychometric properties of the Spanish version of the ERI (Effort-Reward Imbalance) Questionnaire in Colombian Workers. R Psictrabajo Org 26, 147–56.

[259•] - Magnavita, N., Garbarino, S. and Siegrist, J. (2012). The Use of Parsimonious Questionnaires in Occupational Health Surveillance: Psychometric Properties of the Short Italian Version of the Effort/Reward Imbalance Questionnaire. *The Scientific World Journal*, 2012, pp.1-7.

[260•] - Dallner M, Elo A-L, Gamberale F, Hottinen V, Knardahl S, Lindström K, Skogstad A & Örhede E. (2000). Validation of the General Nordic Questionnaire (QPSNordic) for Psychological and Social Factors at Work. Nordic Council if Ministers, Copenhagen. Nord: 2000:12.

[261•] - Wännström I, Peterson U, Åsberg M, Nygren Å, Gustavsson J. (2009). Psychometric properties of scales in the General Nordic Questionnaire for Psychological and Social Factors at Work (QPS Nordic): Confirmatory factor analysis and prediction of certified long-term sickness absence. Scandinavian Journal of Psychology. 50(3):231-244.

[262••] - Gignac M. (2005). Arthritis and employment: An examination of behavioral coping efforts to manage workplace activity limitations. Arthritis Rheum. 53(3):328-336.

[263*] - Leiter MP, Maslach C. (1999). Six Areas of worklife: A model of the organizational context of burnout. J Health Hum Resour Adm. 21:472–489.

[264•] - Gascón, S., Leiter, M., Stright, N., Santed, M., Montero-Marín, J., Andrés, E., Asensio-Martínez, A. and García-Campayo, J. (2013). A factor confirmation and convergent validity of the "areas of worklife scale" (AWS) to Spanish translation. *Health and Quality of Life Outcomes*, 11(1), p.63.

[265*] - Linda Koopmans, Individual Work Performance Questionnaire, 2013, TNO Innovation for Life – VU University Medical Center

[266•] - Koopmans, L., Bernaards, C., Hildebrandt, V., de Vet, H. and van der Beek, A. (2014). Construct Validity of the Individual Work Performance Questionnaire. *Journal of Occupational and Environmental Medicine*, 56(3), pp.331-337.

[271*] - Ilmarinen J. The Work Ability Index (WAI). Occup Med. 2007;57(2):160

[272••] - Leggett, S., van der Zee-Neuen, A., Boonen, A., Beaton, D., Bojinca, M., Bosworth, A., Dadoun, S., Fautrel, B., Hagel, S., Hofstetter, C., Lacaille, D., Linton, D., Mihai, C., Petersson, I., Rogers, P., Sergeant, J., Scire, C. and Verstappen, S. (2015). Test-retest Reliability and Correlations of 5 Global Measures Addressing At-work Productivity Loss in Patients with Rheumatic Diseases. *The Journal of Rheumatology*.

[273•] - Yang, D.J., Kang, D., Kim, Y.K., Kim, Y.H., Yang, Y.A., Cha, S.M., Eom, I.K. and Kim, J.E., (2013). Reliability of self-administered Work Ability Index questionnaire among Korean workers. *Ergonomics*, *56*(11), pp.1652-1657.

[274•] - Martus, P., Jakob, O., Rose, U., Seibt, R. and Freude, G. (2010). A comparative analysis of the Work Ability Index. *Occupational Medicine*, 60(7), pp.517-524.

[275•] - Martinez MC, Latorre MRDO, Fischer FM. Validity and reliability of the Brazilian version of the Work Ability Index questionnaire. Revista de Saúde Publica. 2009;43(3):525–32

[276•] - Kaewboonchoo, O. and Ratanasiripong, P. (2015). Psychometric properties of the Thai version of the work ability index (Thai WAI). *Journal of Occupational Health*, 57(4), pp.371-377.

[277•] - Abdolalizadeh M, Arastoo A, Ghsemzadeh R, Montazeri A, Ahmadi K, Azizi A. The Psychometric Properties of an Iranian Translation of the Work Ability Index (WAI) questionnaire. Journal of Occupational Rehabilitation . 2012;22(3):401-408.

[278*] - Cooper, C., Sloan, S., Williams, S. (1988). Occupational Stress Indicator. Windsor: NFER-Nelson.

[279•] - Russinova, V., Vassileva, L., Randev, P., Jiliova, S., Cooper, C.L., (1997). Psychometric analysis of the first Bulgarian version of the Occupational Stress. Indicator. International Journal of Stress Management 4, 111–119.

[280•] - Davis, A. (1996). A re-analysis of the occupational stress indicator. *Work &Stress*, 10(2), pp.174-182.

[281•] - Swan, J., De Moraes, L. and Cooper, C. (1993). Developing the occupational stress indicator (OSI) for use in Brazil: A report on the reliability and validity of the translated OSI. *Stress Medicine*, 9(4), pp.247-253.

[282•] - Lyne K, Barrett P, Williams C, Coaley K. (2000). A psychometric evaluation of the Occupational Stress Indicator. Journal of Occupational and Organizational Psychology . 73(2):195-220.

[283•] - Steiler D, Paty B. (2009). Developing a French version of the Occupational Stress Indicator (OSI). Revue Européenne de Psychologie Appliquée/European Review of Applied Psychology . 59(2):113-122.

[284••••••] - Kumar R, Hass S, Li J, Nickens D, Daenzer C, Wathen L. (2003). Validation of the Health-Related Productivity Questionnaire Diary (HRPQ-D) on a Sample of Patients with Infectious Mononucleosis: Results from a Phase 1 Multicenter Clinical Trial. Journal of Occupational and Environmental Medicine . 45(8):899-907.

[285•••] - Abbott J., Employment opportunities for Albertans with spinal cord injuries, 1996, University of Alberta.

[286••••••] - Tang K, Beaton D, Amick B, Hogg-Johnson S, Côté P, Loisel P. (2012). Confirmatory Factor Analysis of the Work Limitations Questionnaire (WLQ-25) in Workers' Compensation Claimants with Chronic Upper-Limb Disorders. Journal of Occupational Rehabilitation . 23(2):228-238.

[287••] - Walker, N., Michaud, K. and Wolfe, F., (2005). Work limitations among working persons with rheumatoid arthritis: results, reliability, and validity of the work limitations questionnaire in 836 patients. *The Journal of rheumatology*, *32*(6), pp.1006-1012.

[288••] - Lerner D, Reed J, Massarotti E, Wester L, Burke T. (2002). The Work Limitations Questionnaire's validity and reliability among patients with osteoarthritis. Journal of Clinical Epidemiology . 55(2):197-208.

[289•] - Kono Y, Matsushima E, Uji M. (2014). Psychometric Properties of the 25-Item Work Limitations Questionnaire in Japan. Journal of Occupational and Environmental Medicine . 56(2):184-188.

[290••] - Verhoef J, Miedema H, Bramsen I, Roebroeck M. (2012). Using the Work Limitations Questionnaire in Patients With A Chronic Condition in the Netherlands. Journal of Occupational and Environmental Medicine . 54(10):1293-1299.

[291••••••] - Lerner D, Amick B, Rogers W, Malspeis S, Bungay K, Cynn D. (2001). The work limitations questionnaire. Med Care. 39:72–85.

[292•] - Takegami M, Yamazaki S, Greenhill A, Chang H, Fukuhara S. (2014). Work Performance Assessed by a Newly Developed Japanese Version of the Work Limitation Questionnaire in a General Japanese Adult Population. Journal of Occupational Health . 56(2):124-133.

[293••••••] - Tamminga S, Verbeek J, Frings-Dresen M, De Boer A. (2013). Measurement properties of the Work Limitations Questionnaire were sufficient among cancer survivors. Qual Life Res . 23(2):515-525.

[294•••••] - M. F. O'Donoghue, J. S. Duncan and J. W. A. S. Sander, (1998). The subjective handicap of epilepsy: A new approach to measuring treatment outcome, Brain. 121:317–343

[295••] - Hammond A, Allaire S, Grant M, Woodbridge S, O'Brien R, Culley J (2011). Modifying the Work Environment Survey- Rheumatic conditions for use in the UK. Annals of the Rheumatic Diseases 2011;70(Suppl 3):763

[296*] - Nunnally JC (1978). *Psychometric Theory, 2nd ed*. New York: McGraw-Hill.

[297*] - Hse.gov.uk, (2016). *Returning to work after sickness absence*. [online] Available at: http://www.hse.gov.uk/stress/furtheradvice/returntowork.htm [Accessed 29 Feb. 2016].

[298*] - Shaw, W., Kristman, V., Williams-Whitt, K., Soklaridis, S., Huang, Y., Côté, P. and Loisel, P. (2014). The Job Accommodation Scale (JAS): Psychometric Evaluation of a New Measure of Employer Support for Temporary Job Modifications. Journal of Occupational Rehabilitation, 24(4), pp.755-765.

[299*] - Fava, G., Tomba, E. and Sonino, N. (2011). Clinimetrics: the science of clinical measurements. *International Journal of Clinical Practice*, 66(1), pp.11-15.

[300*] - Jadad, A. R., Moore, R. A., Carroll, D., Jenkinson, C., Reynolds, D. J. M., Gavaghan, D. J., & McQuay, H. J. (1996). Assessing the quality of reports of randomized clinical trials: is blinding necessary?. *Controlled clinical trials*, *17*(1), 1-12.

[301*] - O'Brien AN, Wolf TJ. (2010). Determining work outcomes in mild to moderate stroke survivors. Work. 36(4):441–447

[302*] – Silvia P, Maria Luisa F, Francesco A, Michele V. (2010). Work Self-Efficacy Scale and Search for Work Self-Efficacy Scale: A Validation Study in Spanish and Italian Cultural Contexts. Rev psicol trab organ. 26(3):201-210.

[303*] - Hobart, J., Cano, S., Baron, R., Thompson, A., Schwid, S., Zajicek, J. and Andrich, D., 2013. Achieving valid patient-reported outcomes measurement: a lesson from fatigue in multiple sclerosis. *Multiple Sclerosis Journal*, *19*(13), pp.1773-1783.

Fig. 1 - Study Selection, Inclusion/Exclusion Criteria, and Scale Organization

Search 1

((Stroke OR Multiple Sclerosis OR Epilepsy OR Spinal Cord Injury OR Brain Injury OR Musculoskeletal Disease) AND ((Occupation OR Employment OR Job))) AND ((Outcome measure Or Assessment OR Work Capacity Evaluation OR Scale Or Questionnaire))

PubMed MeSH Terms: Stroke, Multiple Sclerosis, Epilepsy, Spinal Cord Injury, Brain Injury, Musculoskeletal Disease & Work Capacity Evaluation.

Search 2

(Work) AND (Absenteeism OR Presenteeism) AND (outcome measure OR questionnaire OR assessment OR scale)

PubMed MeSH Terms: Absenteeism

# of Abstracts (Search 1) - Pubmed (4074), PMC (4693), EMBASE (3499), CINAHL (1804), PsycINFO (748)

# of Abstracts (Search 2) Pubmed (1557), PMC (1312), EMBASE [+ EMBASE Classic] (3369), CINAHL (849), PsycINFO (771)

597 Outcome Measures identified

Apply Inclusion/Exclusion Criteria

Inclusion:
Substantial and significant portion of the assessment focusing on work/occupation

Exclusion:
Scales that do not refer to work, Work-environment measures, Non-English language instruments, Incomplete or inaccessible instruments (where the author was not contactable), Inventories of several outcome measures

A) 502 instruments failed criteria and rejected

- Unable to contact author/locate tool [5]
- Non-English [10]: German (2), Spanish (3), Chinese (2), Thai (1) and Polish (2)
- Not Work Related (General Health/Specific disease activity (182), Other (72), Disability/Function (39), Psychosocial/Anxiety/Depression (38), Pain (35), Sleep (21), Quality Of Life (19), Beliefs/Views/Self-Efficacy (15), Physical function (14), Stress (12), Coping (8), Sexual behaviors (7), Environment (6), Mental (5), Burden (5), Voice (5), Migraine/Headache (4))

B) 93 measures reviewed (+2 novel measures)

Content Analysis of the 93 reviewed Outcome Measures by a Multi-Disciplinary Group

Psychosocial aspects of work-related Scales (25)
Personality Scales (4), Stress Scales (5), Satisfaction Scales (4), Boredom Scales (1), Well-Being Scales (11)

Physical aspects of work-related Scales (30)
Physical Capacity Scales (3), Absenteeism/Presenteeism Scales (7), Performance (Productivity and Quality) Scales (20)

Work Status Scales (5), Self Efficacy Scales (10), Work Instability Scales (14), Work Ability Scales (8), Job Demands (1)

Retrieval of Psychometric Validation Studies
# of Abstracts – Pubmed & PMC (2055), EMBASE & PsycINFO (1105), CINAHL (289)

Search Strategy:
("Name of Outcome Measure" AND (Psychometric OR Reliability OR Validity OR Responsiveness))

Inclusion Criteria: (n =528)
Studies reporting at least one of the psychometric properties identified in Fig.2

Exclusion Criteria:
Non-English language manuscripts, Rasch model based validation, Validation studies solely focused on predictive validity.

A) 2,921 abstracts rejected (No relevant psychometric data reported)

B) 528 abstracts retained

- 52 abstracts excluded due to Rasch Model-based validation
- 48 abstracts excluded due to language: Spanish (15), Italian (8), German (5), Portuguese (5), Chinese (3), Dutch (2), Turkish (2), French (2), Hebrew (2), Hungarian (2), Romanian (1), Japanese (1)

C) After correction for duplicates (post-application of exclusion criteria), and inclusion of original citations (for the outcome measures) which also reported psychometric properties, 263 datasets were assessed using the quality criteria.

Note: Some manuscripts are counted twice in the dataset count as they analyzed two different outcome measures.

Application of Quality Criteria (Fig. 3)

0 – 20% : 39 Datasets
20 – 40% : 106 Datasets
40 – 60% : 82 Datasets
60 – 80% : 15 Datasets
80 – 100% : 1 Dataset

Number of inaccessible manuscripts: 18

Figure 3 – Definitions for the Baseline Psychometric Properties and the Quality Criteria used for Evaluation

	Description[10,11]	**Quality Criteria**[12,14]
Reliability – "The extent to which scores for patients who have not changed are the same for repeated measurement under several conditions: e.g. using different sets of items from the same work related-patient reported outcomes (WR-PRO) (internal consistency); over time (test-retest); by different persons on the same occasion (inter-rater); or by the same persons (i.e. raters or responders) on different occasions (intra-rater)" [6]		
Internal Consistency	The degree to which items in a scale/outcome measure or a subscale of that measure are homogeneous, and the extent to which they measure various aspects of the same construct.	[2] - Cronbach's alpha(s) calculated per dimension AND Cronbach's alpha(s) between 0.70 and 0.95; [1] - Cronbach's alpha(s) <0.70 or >0.95, despite adequate design and method; [0] – Insufficient sample size, doubtful design or method, or not attempted;
Test-Retest Consistency	The degree to which a scale/outcome measure is stable and produces similar results when administered at two different time points, on the same individual, with no interceding intervention.	[2] - ICC or weighted Kappa ≥ 0.70; [1] - ICC or weighted Kappa < 0.70, despite adequate design and method; [0] - Doubtful design or method (e.g., time interval not mentioned) , or not attempted;
Intra-rater Consistency	The degree to which a scale/outcome measure is stable and produces similar results when conducted by two different administrators (inter), on the same individual, with no interceding intervention.	
Measurement Error	The systematic and random error of a patient's score that is not attributed to true changes in the construct to be measured	[2] - MIC (minimal important change) < SDC (Smallest Detectable Change) OR MIC outside the LOA (Bland-Altman Limits of agreement[A]) OR convincing arguments that the standard error of measurement (SEM) is acceptable; [A] A method to study the mean difference between two quantitative measurements, and therefore to estimate an agreement interval. [1] - MIC ≥ SDC OR MIC equals or inside LOA, OR only MIC reported OR no convincing arguments concerning the acceptability of the SEM, despite adequate design and method; [0] - Doubtful design or method OR (MIC not defined AND no convincing arguments that the agreement is acceptable);
Validity Property - The degree to which a WR-PRO (work-related patient reported outcome) instrument measures the construct(s) it purports to measure		
Content & Face	The degree to which the domain/concept of consequence is sampled (Content)/looks as though is sampled (Face) by the items in the scale/outcome measure.	[2] - A clear description is provided of the measurement aim, the target population, the concepts that are being measured, and the item selection AND target population and (investigators OR experts) were involved in item selection; [1] - A clear description of above-mentioned aspects is lacking OR only target population involved OR doubtful design or method; [0] - No target population involvement, or no attempt to describe development.
Criterion (Concurrent)	The degree to which the scores of an HR-PRO instrument are an adequate reflection of a 'gold standard'	[2] - Convincing arguments that gold standard is "gold" AND correlation with gold standard ≥0.70; [1] – Describes correlation with any tool, but no convincing arguments that tool used is the "gold standard". [0] - Correlation with other outcomes not discussed
Criterion (Predictive)	The degree to which a scale/outcome measure can forecast a specific outcome at later time points.	Any appropriate mathematical method for demonstrating predictive relationship *[Not part of the quality assessment]*
Construct Validity (Hypothesis-testing)	Whether a scale performs as hypothesized by a priori defined relationships/constructs.	[2] - Specific hypotheses were formulated AND at least 75% of the results are in accordance with these hypotheses; [1] - Less than 75% of hypotheses were confirmed, despite adequate design and methods. [0] – Not attempted (e.g., no hypotheses, or exploration);
Construct Validity (Structural)	The extent to which a factor analysis supports the interrelationship between a set of items on a scale and the domains or the constructs theoretically measured by the scale or by subscale structure.	[2] - Factor analyses performed on adequate sample size (4-10 subjects per variable, and minimum 100 subjects in total); [1] – Factor analyses performed, but with inadequate sample size; [0] – Exploratory/confirmatory factor analysis not performed;
Construct Validity (Cross Cultural)	The degree to which the performance of the items on a translated or culturally adapted instrument are an adequate reflection of the performance of the items of the original version of the instrument	[2] – Confirmatory factor analysis of the translated tool, based on: (a) at least two forward translations from the source language that yields a pooled forward translation; (b) at least one, backward translations to the source language that results in another pooled translation; (c) a review of translated versions by lay and expert panels with revisions; [1] – At least one of the following is missing: (a) at least two forward translations from the source language that yields a pooled forward translation; (b) at least one, backward translations to the source language that results in another pooled translation; (c) a review of translated versions by lay and expert panels with revisions; (d) confirmatory factor analysis; [0] - No confirmatory factor analysis conducted.
Other		
Responsiveness	The extent to which a scale has the ability to assess clinically important change over time.	[2] – Standardized co-efficient of responsiveness reported and suggestive of moderate to high responsiveness (e.g. Cohens d (Effect size) and SRM[B] > 0.5. Or Guyatts Responsiveness Ratio > 1.96 OR AUC[C] ≥ 0.70 [1] – Standardized co-efficient of responsiveness reported and suggestive of low to moderate responsiveness (e.g. Cohens d (Effect size) or SRM < 0.5. Or RR< 1.96, or AUC< 0.70. [0] - Doubtful design/ method, or not attempted; Note - SRM * √ 2 * √(1−r) = Cohen's d[15] [B] The SRM (Standardized Response Mean) is the mean change divided by the standard deviation of the change scores [C] The AUC (Area under Curve) is the area under the curve can be used as a quantitative method for assessing a scale's ability to distinguish patients who have improved from those who have minimally or not changed based on the global rating of change.

Figure 4 – An illustration of the different areas captured by the work ability tools.

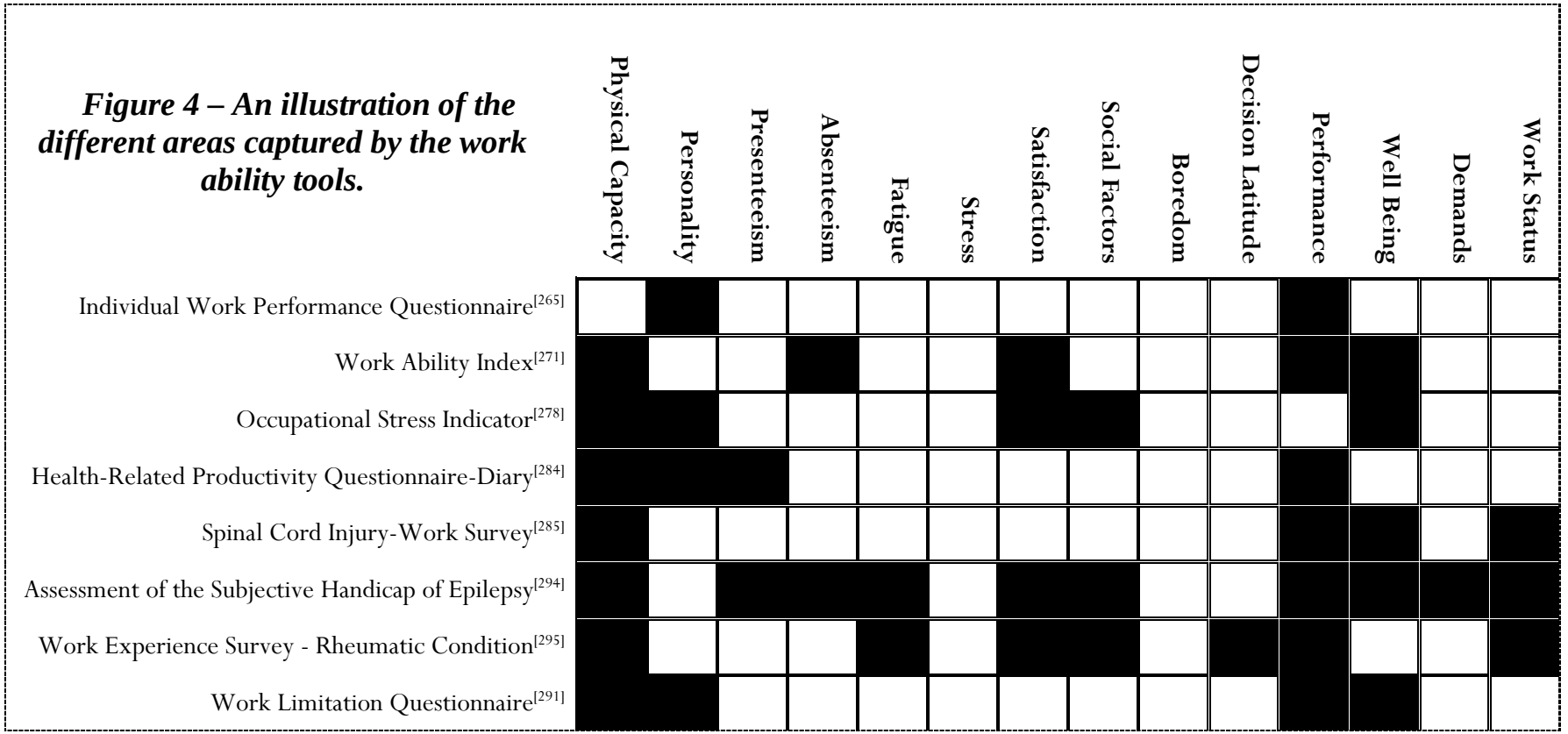


Fig. 2 - List of Codes relating to Work-related Skills

101 - Absenteeism	**102** - Burnout	**103** - Decision Latitude
104 - Presenteeism	**105** - Satisfaction	**106** - Fatigue
107 - Quality	**108** - Ergonomic	**109** - Job Demands
110 - Productivity	**111** - Social Factors	**112** - Well-being
113 - Performance	**114** - Work Ability	**115** - Personality
116 - Psychosocial	**117** - Self-efficacy	**118** - Work Instability
119 - Mood	**120** - Work Status	**121** - Physical Capacity
122 - Environment	**123** - Stress	**124** - Boredom

Abstract

Purpose - Selecting the most appropriate health-related work outcome to evaluate an intervention can be fraught with difficulty. To aid clinicians in navigating this problem we have developed a model which illustrates how pathology can affect specific measureable quantities, such as work instability.

Methods – Using a modified-Delphi procedure, a panel of experts met initially to analyze the content of 95 health-related work outcome measures and organize the identified areas of measurement into a coherent model, complemented by a narrative review of the literature. This initial model underwent two rounds of stakeholder-based feedback, the results of which were incorporated in the final expert panel meeting to produce the States-traits Work Instability Model (SWIM).

Results – The States-traits Work Instability Model (SWIM) illustrates how changes to an individual's physical and psychological states and traits might affect their work-related performance, well-being and self-efficacy. Moreover, each concept utilized in the model was specifically selected as it represents a measurable quantity, for which there are tools available.

Conclusion - The SWIM is arguably the first holistic model of work that is based on both the clinical realities of vocational rehabilitation, sociological research and is born from analysing the basis of practical measurements.

Word count - 196

Implications for Rehabilitation

Work Instability

- Work instability has multiple causes many of which many are amenable to intervention
- The model clarifies the measureable domains of vocational rehabilitation interventions, which is of particular benefit for services working with people with disability at work who are struggling to remain in work
- The model conceptualises how the potential areas for intervention may be related based on evidence available in the literature.

Introduction

'Health is no longer viewed as the absence of organ pathology, but rather as the possession of a repertoire of skills that enables people to achieve their goals' [1]. Therefore, individuals with disabilities are not precluded from 'being healthy' as they are still capable of "developing and using skills to achieve their goals" [2]. One example of these goals that plays a substantial role in the adult years of an individual's life, is the acquisition, and retention of employment. There is substantial evidence to suggest that work is a central component of an individual's identity, and that there are significant adverse effects of job insecurity and unemployment on health [3]. Conversely, employment has physical and mental health benefits [3].

Vocational Rehabilitation is 'whatever helps someone with a health problem to stay at, return to and remain in work' [4]. A key problem these services face is deciding how best to measure the effectiveness of the specific interventions they provide. For example, work status (the binary distinction of employed or unemployed) can be an effective measure for people who are out of work at the time of intervention. However, there are many who are in work and struggling, for whom this measure is inappropriate. Fortunately, when the definition of health [1] is contextualized to work, it suggests that there are numerous measureable quantities (e.g. physical capacity, self-efficacy, psychological well-being), which can be used to measure the effects of an intervention.

There are several examples in the literature of models for specific work-related concepts. Two of the more widely recognised models are the Job Demand-Control Model [5], and Job Demands-Resources (JD-R) Model [6]. The Job-Demand Control model [5] is a predictive tool which illustrates how the variation of two concepts, job demands and control, predict activity level and strain. The Job Demands-Resources (JD-R) Model [6] is a more comprehensive construct that describes the interaction of job demands, and the (resources) external constructs, such as autonomy and social factors that the individual has to support them in meeting the aforementioned demands. The methodology commonly utilized to create such models, as was done for both of the aforementioned models, is a theory-driven approach, coupled with post-hoc quantitative proofs of the theorized relationships. The result of which is the selection of concepts which are relevant to the theory, and exclusion of all others. This process has led to the under-appreciation of specific concepts such as personality, well-being, and consequently, work ability (the combination of physical and psychosocial concepts) in affecting an individual's ability to work; each of which appears to have dedicated tools that measure it [7]. As a consequence of this hypothesis-driven approach, the literature appears to lack a clear conceptual model capable of aggregating, and demonstrating the interactions between all of the commonly measured work-related concepts.

The relationships identified between work-related concepts are usually correlative in nature. As such, the results of any experiment can be interpreted in a number of way, depending on the assumptions and theories that inform the investigator's perspective. The models described above are an example of how similar work-related concepts can be conceptualized into two different models depending on the theoretical approach that informs the investigator's work. We sought to utilize a data driven approach, using the range of concepts identified in a recent systematic review [7], coupled with a modified-delphi approach [8], to amass as many of these theoretical perspectives as possible.

The purpose of this study was to create a consensus derived model that illustrated how pathology can eventually lead to work instability, and the associated impact on the intermediary concepts, such as work performance, self-efficacy, and well-being.

Box 1 (Background)

The States-Traits Theory

The sociological concepts of both states and traits have been discussed in the academic sociology literature for close to a century, however, the premise of the state-trait distinction can be traced as far back as 45 B.C., to the work of Cicero. The discussion presented below is based on the work of Chaplin John, and Goldberg [9].

Traits are commonly defined as internally generated attributes, with a significant degree of temporal stability (i.e. they are long-lasting, and are unlikely to be vary day by day) [10]. Chaplin and colleagues suggest that traits allow individuals to make predictions in present circumstances based on the past. States are considered to be more transient, and brief; a consequence of external circumstances interacting with the aforementioned traits. The practical relevance of states is that they identify behaviors or emotions/feelings which can be manipulated by altering one's environment/situation [9]. This can be neatly summarized by Zukerman's (1983) 'locus of causality' definition [11]. States change as a function of an individual's external condition, whereas traits are the intrinsic factors that interact with the external condition to produce the state.

Worked Example

'Physical capacity', can be thought of as the amount of energy an individual is able to expend, coupled with what that individual's body is capable of achieving using that energy (mathematically this could be expressed in a very simplistic form as the maximum amount of power an individual can generate). 'Physical Capacity' would therefore be considered a trait, using the Zukerman definition [11]. Both of its constituents would be stable across all situations, although the result may be different, because each task requires a different amount of power to produce one unit of output. 'Fatigue' on the other hand, would be a state. Fatigue is commonly understood to be the process of gradual exhaustion resulting from a physical or mental cause. Therefore, as the energy requirements of each task differ, consequently, different tasks will result in different levels of fatigue, making it condition-dependent.

An important point from the work conducted by Chaplin, John, and Goldberg, is that the distinction between states and traits is not always absolute, but rather, can be 'fuzzy' [9]. In fact, it was Rosch that first argued that the categorisation of concepts as states or traits does not need to be absolute, but instead "class membership can be a matter of degree" [9,10]. The label of state or trait applied to the work-related concepts in this study are based on the consensus achieved by groups of specialists, and are by no means perfect. Consequently, the organisation of the concepts in the figures are suggestions. The role of different concepts and factors will vary from one individual to another, and thus, it is possible that there may be no single concept, that can fulfil the states or traits portion of the model, or there may be several. The purpose of the model is to provide a common language with which to discuss the work-related challenges an individual may face, in manner conducive to identifying solutions.

Methods

A modified-Delphi procedure [8] consisting of four stages was used to develop the following model (described in fig. 1). Firstly, a systematic review of the literature identified 95 health-related work outcome measures. Five databases were searched: Pubmed, PMC, EMBASE (+ EMBASE Classic), CINAHL, and PsycINFO, from inception to 2016. This process returned over twenty-two thousand abstracts (not corrected for duplicates). All of the outcome measures utilised in these studies were extracted (597 outcome measures), and then assessed to determine whether they were work/occupation-specific, based on their content. Tools not explicitly relevant to work, non-English language tools, or those specific to the work-environment and not the individual's behaviour, were excluded. More specific details regarding the inclusion/exclusion criteria and the process of content-analysis can be found in the relevant citation [7].

In between stage 1 and 2, a list of key terms relating to health-related work-specific outcomes was produced on the basis of research and clinical expertise, which was complimented by a narrative review of the literature (the search strategy is detailed in fig. 1).

The second stage comprised of an expert panel of vocational rehabilitation specialists consisting of a researcher, a clinical psychologist, a consultant neurologist, and three Occupational Therapists (OTs), each individually classified, and subsequently grouped the 95 work-related outcome measures [7] using the key terms. The identified groups were used to formulate a conceptual model of the work-related skills each individual possess, initially on the basis of clinical expertise. Then, the narrative review of the literature was utilised to determine the accepted definitions for each of the terms utilized, and to identify any data relating to, and discussing the nature of, the proposed relationships. This resulted in the Work-Instability Model available in the Appendix.

Stakeholder feedback was elicited twice in stage 3, where author B.A.M. acted as the Delphi procedure facilitator. On each occasion approximately 30 physicians, occupational therapists and psychologists working in the area of vocational rehabilitation attended a workshop. The workshops consisted of a one-hour long presentation of the model, the process of its conception, and its current state. This was followed by an unstructured round-table where the participants were invited to provide feedback on any aspect of the model. The individual contributions were considered by the group and those receiving majority support were subsequently incorporated by the expert panel, based on the recommendations of the stakeholders.

Finally, the expert panel re-convened to conduct a similar process as that in stage 2, with two notable changes. A large number of outcome measures spanned several categories, but were previously (stage 2) classified only according to the dominant category, disregarding any secondary categories. Therefore, the first additional element, was to identify all the generic terms/work-related skills for each of the complex outcome measures. Secondly, all the additional information gathered was then incorporated into the model, and again complimented with the narrative review described in fig. 1, to produce the States-Traits Work Instability Model (S.W.I.M.)

Results

The first meeting of the expert panel to classify the outcome measures (stage 2) resulted in 5 classes of tools. The classes of the outcome measures were Performance scales (encapsulating the generic terms: 'Presenteeism', 'Absenteeism', 'Quality', and 'Productivity'), Psychosocial scales ('Stress', 'Satisfaction', 'Boredom', and 'Social Factors'), Self-efficacy scales ('Self-efficacy'), Work instability Scales ('Work Instability') and Job (Demand) Scales ('Job Demands').

The principal contribution from the stakeholder-based sessions was that the model might benefit from being adapted to include the States-Traits theory. The reasoning was that this theory might provide a rubric with which to organise the identified terms. The second key contribution of the feedback sessions was the identification of specific work-related outcome measures [e.g. 12], which were not sampled in by systematic review [7], and introduced previously overlooked work-related concepts such as personality.

The States-Traits Work Instability Model (SWIM)

The SWIM (figure 2) can be thought of as the combination of three modules, each of which explains some aspects of an individual's interaction with work. The three modules are physical, psychological and work instability (Illustrations of each module are available in the Appendix).

The Physical Module - Work-related Physical Traits, States, and Products

The physical concept that best fit Zukerman's 'Locus of Causality' definition was agreed to be that of the physical capacity [13] of the individual in question. Practically, this can be understood to be the amount of energy an individual is able to expend, coupled with what that individual's body is capable of achieving using that energy. States [9] are more transient states-of-body, for example, fatigue. The role of the states is that they regulate the effectiveness with which an individual's input (their psychical capacity) is converted into output, or in this particular circumstance their productivity, and quality of work (i.e. their work performance). The relevance of the perception moderator and work-related self-efficacy in both the physical and psychological modules is discussed later.

The unique aspect of the physical module is that there are only a couple of explanations for how pathology impacts physical output, i.e. work-performance. These circumstances are described in the literature as presenteeism and absenteeism. Presenteeism is "the problem of workers being on the job, but, because of illness or other medical conditions, not fully functioning" [14], and absenteeism is medically certified absence, where the absence from work is attributed to disease, medical condition or accident" [15]. Although they are not true states-of-body, and do not fit the states-traits model, they are a useful set of terms for summarizing the impact of pathology on the physical module. The relationships between the work-related physical states, and these circumstances have been explored elsewhere in the literature [16-17].

The Psychological Module (Figure 3) - Work-related Psychological Traits, States, and Products

Personality is the trait in this module, due it's relative stability, and because it was thought to be the most intrinsic of the psychological work-related concepts.

The psychological states identified originally included: satisfaction, stress, boredom, and work-specific social interactions. Later this was expanded to include decision latitude, a term thought to comprise the individuals' beliefs surrounding their "intellectual discretion and personal schedule freedom" at work [5]. The relationships between these psychological states were also based on experimental/clinical data [18-23].

The long term product of this module is the psychological well-being on the individual, which Ryff describes as having 6 distinct components: self-acceptance, positive relationship with others, autonomy, environmental mastery, purpose in life, and personal growth [24].

Self Perception

Central to vocational rehabilitation is the individual's understanding of his or her own capabilities to produce designated levels of performance, this is known as their self-efficacy. These self-efficacy beliefs determine how people feel, think, motivate themselves, and even behave [25]. Moreover, these beliefs are shaped and influenced by learning, experience and feedback [26], suggesting they are highly malleable. For these reasons it seems appropriate to suggest that self-efficacy is much more likely to reflect short term outcomes. Whereas, to produce a plausible estimate of productivity or quality (performance), and psychological well-bring, the time course of the measurement would need to be substantially longer. This is to account for periodic fluctuations in self-efficacy beliefs which may result in small variations in day-to-day performance and well-being. This was the justification for implying that time-course (short vs. long term) is the basis for distinguishing between self-efficacy and the other outcomes (well-being & performance).

The relevance of the perception moderators is that they reflect an important process commonly seen in clinical practise. When an inequality exists between the patient's self-perceived capability (self-efficacy) and their objective capabilities (their performance and well-being), i.e. a patient overestimates their abilities, the clinical implications of this incongruence manifest as a change in psychological or physical state (e.g. they can experience greater levels of stress). A process illustrated by Gist several times [27-28]. The perception moderators are the visual representations of this process.

Work Instability

The model proposes that there are four different mechanisms by which work instability can arise (fig. 4). The first two inequalities are more intuitive, and are a result of physical or psychological demands exceeding the performance or psychological well-being/resilience of the patient. These two are likely to be more academic explanations, as, both physical and psychological states are known to be highly interdependent [29-32]. Therefore, the more clinically relevant explanations include that of work ability [33-34] – the combination of all the work-related skills an individual

possess – being exceed by the total job demands. And additionally, even where an actual deficit in ability does not exist, a self-perceived deficit (Demands > Self efficacy) can be equally as detrimental.

Finally, the literature (e.g. 35-36) suggests that pathology manifests as an alteration to the physical/psychological states and/or traits, which eventually results in the aforementioned change to work instability. Therefore, the hypothesized States-Traits Work Instability Model (SWIM – fig. 2) proposes that all changes (due to pathology) to the work-related states and traits possessed by an individual, eventually manifests as a change in the levels of work instability faced by the individual.

Box 2 (Practical Application of the S.W.I.M.)

Case Study (John Doe)

Summary

Mr. Doe is 52-year-old gentlemen working in customer service. He is experiencing difficulty at work as a result of the left anterior cerebral artery stroke he suffered nine months ago. As Mr. Doe was incapacitated for several months and has residual right-sided lower limb monoparesis, as a result he has lost a substantial amount of muscle mass reducing his ability to mobilise independently, and he is much more prone to fatigue. Furthermore, Mr. Doe has been acting oddly, and has begun to ask for more overtime shifts due to accruing gambling debts. His line manager describes him as being more impulsive, which is out of character given that he was previously a very cautious and careful man, with no previous predilection for gambling. Mr. Doe usually takes significant pride in his work, however the current difficulties he is facing, and the emotional toll taken by the period of ill health appears to have left him experiencing symptoms consistent with depression (apathy, lack of motivation, etc.).

Physical States and Traits

Mr Doe's physical needs can be separated into reduced work capacity (trait) due to reduced muscle mass, and an increased propensity for fatigue (state) as a more transient result of this. Furthermore, there is a second explanation for his reduced work capacity; the monoparesis (trait), which likely contributes to the fatigue (state). The former state-trait set can be treated with long term physical rehabilitation to improve his muscle strength, and co-ordination. The introduction of short term specific work-strategies can be used to manage this problem, whilst the physical therapy takes place, for example: work-demands adjustment to prevent fatigue (i.e reducing the number of hours worked or reducing the number of physically strenuous events in his day-to-day work), or even work-environment modification to reduce the amount that Mr. Doe needs to mobilize. As the model (fig. 1) illustrates, the fatigue is a result of the reduced muscle mass, so the rehabilitation is likely to help improve this problem over the long term. The latter series of problems (monoparesis-fatigue) is more likely to have a ceiling of potential benefit, as therapy can improve mobility and teach the patient to compensate, but we cannot as of yet, reverse the damage that causes the problem. As such, the long term solution is work-adjustment. The model provides a framework to separate out what we would normally see as a single holistic picture, into a series of problem sets, each of which requires individual consideration. The value of the latter approach is that it demonstrates the practical differences between the management of a patient with Guillian Barre Syndrome, for whom we could realistically expect a full recovery [37] and phased return to work, as per their pre-morbid state, versus, a stroke patient where work-adjustment of some form is the norm.

Psychological States and Traits

It appears as though Mr. Doe's illness has led to both a change in his personality (increased impulsiveness – gambling), and separately, a series of negative states that can be encapsulated in a diagnosis of depression. There are several pharmacological [38] and cognitive approaches [39] that are available for the treatment of impulsive behaviour. The importance of separating the psychological states and traits is more evident when we consider recent studies that have demonstrated that personality traits, not the more transient states which appear to mask the trait variance, predict depression [40-42] and short term outcome [43-44]. Whilst the treatment for the depression will likely be a combination of talk therapy and medication [45], as per normal, the model encourages early detection a by engaging clinicians with the concept of trait (i.e. personality) variance. In this case, there does not appear to be any psychological states that need to be handled as individual entities, as they all appear to be intricately involved with the other difficulties being faced. It should be noted that some patients may not demonstrate changes to their personality (trait) and may suffer from stress, and reduced decision latitude (states), due to their physical limitations which will require its own individual intervention.

Measuring Outcome

Finally, the purpose of the model is to help distinguish the states and traits from their eventual impact. For example, the physical limitations will influence an individual's work performance, and depression will impact their overall well-being. It is important to measure not only the direct effect of the intervention, i.e. work-related physical function improvements as a result of physical therapy, but also the impact on performance, work ability or work instability to ascertain whether the end goal of vocational rehabilitation is being met. Tools for each of the aforementioned concepts can be found in the previously conducted systematic review [7].

Discussion

The rationale for producing the SWIM was to provide a language and tool for clinicians to use, to increase transparency around vocational rehabilitation, the associated interventions. Embracing a widened definition of what it means to be healthy [1], requires non-specialists to understand which health-related work outcome measures are most appropriate to use in their specific circumstances. Moreover, providing a common framework will hopefully allow more specialist clinicians (in vocational rehabilitation) to communicate their decisions regarding which interventions are necessary. Better communication with less specialist colleagues, should hopefully serve to better integrate care and encourage engagement with vocational rehabilitation specialists. The SWIM presents a road map describing the relationship between the quantifiable health-related work outcomes (each rectangular box contains measurable quantities for which there are several available measure [7]). As such, the action of a specific intervention in reducing pathology-associated work instability is easily deduced from the model, and thus selecting the most suitable outcome measure from the many that exist is made easier.

An incidental consequence of the hypothesized model is that it allows us to improve the current definition of work instability [46]. A preliminary revision would be; work instability is the mismatch between an individual's *physical and/or psychological capacity, perceived or otherwise,* and their job demands. The inclusion of "perceived or otherwise" is based on the moderation of work-related self-efficacy against job demands, potentially producing a perceived deficit in capability. Furthermore, the original "functional capacity" reference in the definition of work instability [35] is improved by explicitly distinguishing between physical and psychological capacity.

Limitations

A limitation of the model is that it is presented with discrete boundaries, in reality the network of connections is appreciably more complex. Job boredom, for example, is correlated with lower levels of self-rated health, and also poor work ability [20], thus connecting the psychological states to the physical traits, as well as with overall health and the impact of the pathology. Whilst the model is a simplification of the realities of individual-work interactions, this is necessary to produce a model with clinical utility. There is always a trade-off between accurate representation and over-complication. Since more intricate models and theories already exist for most of the individual concepts, the S.W.I.M. was designed to help interpret how these concepts coalesce in each individual.

Validity

The validity of the model is drawn from the quantitative studies that demonstrate each of the relationships that individually make up the SWIM. However, the state of the health-related work outcome literature is not sufficiently developed to be able to conduct a holistic quantitative validation study. The solution commonly used to circumvent this problem is to conduct a factor loading analysis on a dataset based on the relevant measurement tools, which would allow us to determine whether they are measuring different things. Subsequently, based on inferential content analysis, the tools clustering around each factor could be linked to one of the areas in the SWIM,

thereby demonstrating validity. Moreover, preliminary validity could be inferred by using outcome measures currently available demonstrating relationships consistent with the predictions of the SWIM, for example, a patient with documented and measured fatigue-related problems demonstrating higher levels of work-instability than a similar non-fatigued colleague. Whilst, the number of concepts the SWIM attempts to aggregate makes this task substantially more difficult than the more focused models for which these methods are commonly used, it is possible for individual sections of the SWIM to be validated using this method."

Implications

What this paper and model aim to do, is clarify the natural accumulation of factors related to work, starting with performance and psychological measures, and with increasing complexity, leading to changes in self-efficacy & work instability. This model should help clinicians achieve two main goals. Firstly, it should aid in the decision of which outcome measures are most appropriate for their service. And secondly, it allows clinicians to map where their intervention is targeted in relation to the different characteristics of an individual.

Conclusion

The S.W.I.M. is arguably the first holistic model of work that is based on both the clinical realities of vocational rehabilitation, sociological research and is born from analysing the basis of practical measurements. By bringing together research from the past seven decades, we've been able to create an up-to-date road map of how the principal work-related concepts interact, resulting in work instability. Finally, we have been able to contextualize these concepts to clinical practise where they are most needed, and thus have been able to present an update to the definition of work instability.

Declarations of Interest –

Dr Diane Playford was supported by the National Institute for Health Research University College London Hospitals Biomedical Research Centre. The authors have no conflict of interests to declare. This study received no funding

References

[1] – Yerxa E. Health and the Human Spirit for Occupation. American Journal of Occupational Therapy. 1998;52:412-418.

[2] – Pörn I. Health and adaptedness. Theor Med Bioeth. 1993;14:295-303.

[3] – Waddell G, Burton AK. Is work good for your health and well-being?. The Stationery Office; 2006 Sep 6.

[4] – Waddell G, Burton AK, Kendall NA. Vocational rehabilitation–what works, for whom, and when? (Report for the Vocational Rehabilitation Task Group). TSO; 2008.

[5] – Karasek R. Job Demands, Job Decision Latitude, and Mental Strain: Implications for Job Redesign. Administrative Science Quarterly. 1979;24:285.

[6] – Bakker A., Demerouti E. The Job Demands-Resources model: state of the art. Journal of Managerial Psychology. 2007;22:309-328.

[7] - Mateen B, Doogan C, Hayward K, Hourihan S, Hurford J, Playford E. A Systematic Review of Health-related Work Outcome Measures, and Quality Criteria based Evaluations of their Psychometric Properties. Archives of Physical Medicine and Rehabilitation. 2016 [September 1st]. DOI: 10.1016/j.apmr.2016.06.013

[8] - Hsu C.C., Sandford B.A. The Delphi technique: making sense of consensus. Practical assessment, research & evaluation. 2007 Aug;12(10):1-8

[9] – Chaplin W. E., John O. P., Goldberg L. R, Conceptions of States and Traits: Dimensional Attributes With Ideals as Prototypes, Journal of Personality and Social Psychology. 1988;54(4):541-557

[10] - Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), Cognition and categorization (pp. 27-48). Hillsdale, NJ: Erlbaum.

[11] - Zuckerman, M. (1983). The distinction between trait and state scales is not arbitrary: Comment on Allen and Potkay's "On the arbitrary distinction between traits and states." Journal of Personality and Social Psychology, 44, 1083-1086.

[12] - Tyreman, A., Tyreman, R., The Work Personality Profile, Occupational Medicine (Lond). 2009;59(5):361-362.

[13] - Shephard, R. Age and Physical Work Capacity. Experimental Aging Research. 1999; 25(4), pp.331-343.

[14] – Hemp P., Presenteeism: at work—but out of it., Harvard Business Review. 2004;82(10):49–58.

[15] – Luz J, Green MS., Sickness absenteeism from work--a critical review of the literature, Public Health Rev. 1997;25(2):89-122.

[16] - Schilling R.S. Industrial Health Research: The Work of the Industrial Health Research Board, 1918-44. British journal of industrial medicine. 1944 Jul;1(3):145.

[17] - Jeon S, Leem J, Park S, Heo Y, Lee B, Moon S, Jung D, Kim H. Association among Working Hours, Occupational Stress, and Presenteeism among Wage Workers: Results from the Second Korean Working Conditions Survey. Ann of Occup and Environ Med. 2014;26:6.

[18] – Richardsen A.M., Burke R.J., Occupational stress and job satisfaction among physicians: sex differences. Soc Sci Med. 1991;33(10):1179-87.

[19] – Fisher C.D., Boredom at Work: A Neglected Concept, Human Relations 1993;46:395

[20] – Kass S.J., Vodanovich S.J, Callender A. State-trait boredom: relationship to absenteeism, tenure, and job satisfaction. J Bus Psychol. 2001;16:317–327.

[21] – Harju L, Hakanen J, Schaufeli W. Job Boredom and Its Correlates in 87 Finnish Organizations. Journal of Occupational and Environmental Medicine. 2014;56:911-918.

[22] – MacKay C, Cousins R, Kelly P, Lee S, McCaig R. 'Management Standards' and work-related stress in the UK: policy background and science. Work & Stress . 2004;18(2):91-112.

[23] – Meyers, L., Social Relationships Matter in Job Satisfaction. Journal of Applied Psychology, 2007;38(4):14.v

[24] – Ryff, C. D., Happiness is everything, or is it? Explorations on the meaning of psychological well-being. Journal of personality and social psychology, 1989;57(6):1069-1081

[25] – Bandura, A. Self-efficacy. In V. S. Ramachaudran (Ed.), Encyclopedia of human behaviour. Vol. 4. New York: Academic Press; 1994. p 71-81.

[26] – Gist, M.E. and Mitchell, T.R., Self-Efficacy: A Theoretical Analysis of Its Determinants and Malleability, The Academy of Management Review, 1992;17(2):183-211

[27] – Gist, M. E., The influence of training method on self-efficacy and idea generation among managers. Personal Psychology, 1989;42:787-805.

[28] – Gist, M. E., Schwoerer, C., & Rosen, B., Effects of alternative training methods on self- efficacy and performance in computer software training. Journal of Applied Psychology, 1989;74:884-891.

[29] – Gosselin E, Lemyre L, Corneil W. Presenteeism and absenteeism: Differentiated understanding of related phenomena. Journal of Occupational Health Psychology. 2013;18:75-86.

[30] – Hansen, C. and Andersen, J. Going ill to work – What personal circumstances, attitudes and work-related factors are associated with sickness presenteeism?. Social Science & Medicine, 2008;67(6), pp.956-964.

[31] – Woo M, Yap AK, Oh TG, Long FY, The relationship between stress and absenteeism, Singapore Med J. 1999 Sep;40(9):590-5.

[32] – Pflanz SE, Ogle AD, Job stress, depression, work performance, and perceptions of supervisors in military personnel, Mil Med. 2006;171(9):861-5.

[33] – Ilmarinen J. The Work Ability Index (WAI). Occup Med. 2007;57(2):160

[34] – Tengland, P., The Concept of Work Ability. J Occup Rehabil, 2010; 21(2), pp.275-285.

[35] – Morris J., Effects of Right Hemisphere Strokes on Personality Functioning, Top Stroke Rehabilitation, 2009;16(6):425–430

[36] - Duncan PW. Stroke disability. Phys Ther. 1994;74:399-407.

[37] - González-Suárez I, Sanz-Gallego I, Rodríguez de Rivera F, Arpa J. Guillain-Barré Syndrome: Natural history and prognostic factors: a retrospective review of 106 cases. BMC Neurology. 2013;13.

[38] - Prado-Lima PA. Pharmacological treatment of impulsivity and aggressive behavior. Revista Brasileira de Psiquiatria. 2009 Oct;31:S58-65.

[39] - Hodgins DC, Peden N. Cognitive-behavioral treatment for impulse control disorders. Revista Brasileira de

Psiquiatria. 2008 May;30:S31-40.

[40] - Clark LA, Vittengl J, Kraft D, Jarrett RB. Separate personality traits from states to predict depression. Journal of personality disorders. 2003 Apr;17(2):152.

[41] - Vittengl JR, Clark LA, Thase ME, Jarrett RB. Replication and extension: Separate personality traits from states to predict depression. Journal of personality disorders. 2014 Apr;28(2):225.

[42] - Barnett JH, Huang J, Perlis RH, Young MM, Rosenbaum JF, Nierenberg AA, Sachs G, Nimgaonkar VL, Miklowitz DJ, Smoller JW. Personality and bipolar disorder: dissecting state and trait associations between mood and personality. Psychological medicine. 2011 Aug 1;41(08):1593-604.

[43] - Blom MB, Spinhoven P, Hoffman T, Jonker K, Hoencamp E, Haffmans PJ, van Dyck R. Severity and duration of depression, not personality factors, predict short term outcome in the treatment of major depression. Journal of affective disorders. 2007 Dec 31;104(1):119-26.

[44] - Merz EL, Roesch SC. Modeling trait and state variation using multilevel factor analysis with PANAS daily diary data. Journal of research in personality. 2011 Feb 28;45(1):2-9.

[45] - Hollon SD, Thase ME, Markowitz JC. Treatment and prevention of depression. Psychological Science in the public interest. 2002 Nov 1;3(2):39-77.

[46] – Gilworth G, Chamberlain MA, Harvey A, Woodhouse A, Smith J, Smyth MG, Tennant A: Development of a work instability scale for rheumatoid arthritis. Arthritis Rheum. 2003, 49:349-354.

[47] – Huppert, F. Psychological Well-being: Evidence Regarding its Causes and Consequences. Applied Psychology: Health and Well-Being, 2009; 1(2), pp.137-164.

Figure Legends

Figure 1

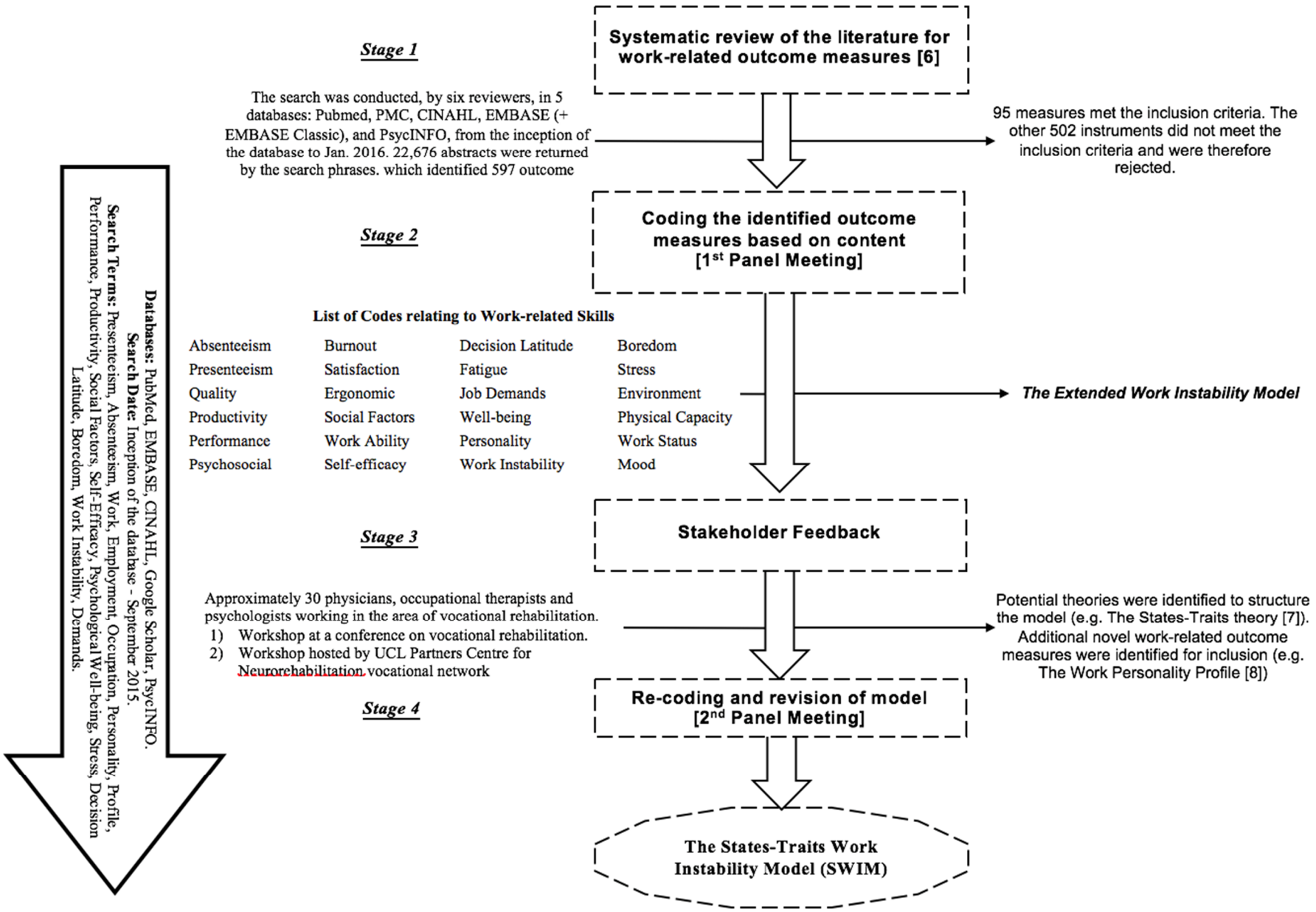
A modified-Delphi procedure consisting of four stages, coupled with a review of the literature to identify supporting evidence for the hypothesized models.

Figure 2

The States-Traits Work Instability Model (SWIM) proposes that all the work-related skills an individual possess can be described as being either physical or psychological in nature. Traumatic or progressive pathology are thought to directly impact both the physical and psychological states and traits [35-36]. These changes eventually alter the levels of work instability faced by the individual.

The physical and psychological states included in the model are examples of the many potential states that exist. Using the description of the States-Trait model in Box 1, it should be possible to discern whether a concept not discussed here belong in the states or traits category.

The model itself is hierarchically organized from most intrinsic/least external factor involvement, to least intrinsic/most external factor involvement. The core work-related traits are at the top (physical capacity and personality), and with each progressive movement along a pathway (arrow) the effect of external factors increases. For example, psychological well-being is in essence an individual's personality moderated by their potential psychological state(s) [47]. This has been represented as an arrow from personality to well-being transected by the psychological states.


Stage 1
Systematic review of the literature for work-related outcome measures [6]
The search was conducted, by six reviewers, in 5 databases: Pubmed, PMC, CINAHL, EMBASE (+ EMBASE Classic), and PsycINFO, from the inception of the database to Jan. 2016. 22,676 abstracts were returned by the search phrases. which identified 597 outcome
95 measures met the inclusion criteria. The other 502 instruments did not meet the inclusion criteria and were therefore rejected.
Stage 2
Coding the identified outcome measures based on content [1st Panel Meeting]
List of Codes relating to Work-related Skills
Absenteeism
Presenteeism
Quality
Productivity
Performance
Psychosocial
Burnout
Satisfaction
Ergonomic
Social Factors
Work Ability
Self-efficacy
Decision Latitude
Fatigue
Job Demands
Well-being
Personality
Work Instability
Boredom
Stress
Environment
Physical Capacity
Work Status
Mood
The Extended Work Instability Model
Stage 3
Stakeholder Feedback
Approximately 30 physicians, occupational therapists and psychologists working in the area of vocational rehabilitation.
1) Workshop at a conference on vocational rehabilitation.
2) Workshop hosted by UCL Partners Centre for Neurorehabilitation vocational network
Potential theories were identified to structure the model (e.g. The States-Traits theory [7]). Additional novel work-related outcome measures were identified for inclusion (e.g. The Work Personality Profile [8])
Stage 4
Re-coding and revision of model [2nd Panel Meeting]
The States-Traits Work Instability Model (SWIM)
Databases: PubMed, EMBASE, CINAHL, Google Scholar, PsycINFO.
Search Date: Inception of the database - September 2015.
Search Terms: Presenteeism, Absenteeism, Work, Employment, Occupation, Personality, Profile, Performance, Productivity, Social Factors, Self-Efficacy, Psychological Well-being, Stress, Decision Latitude, Boredom, Work Instability, Demands.

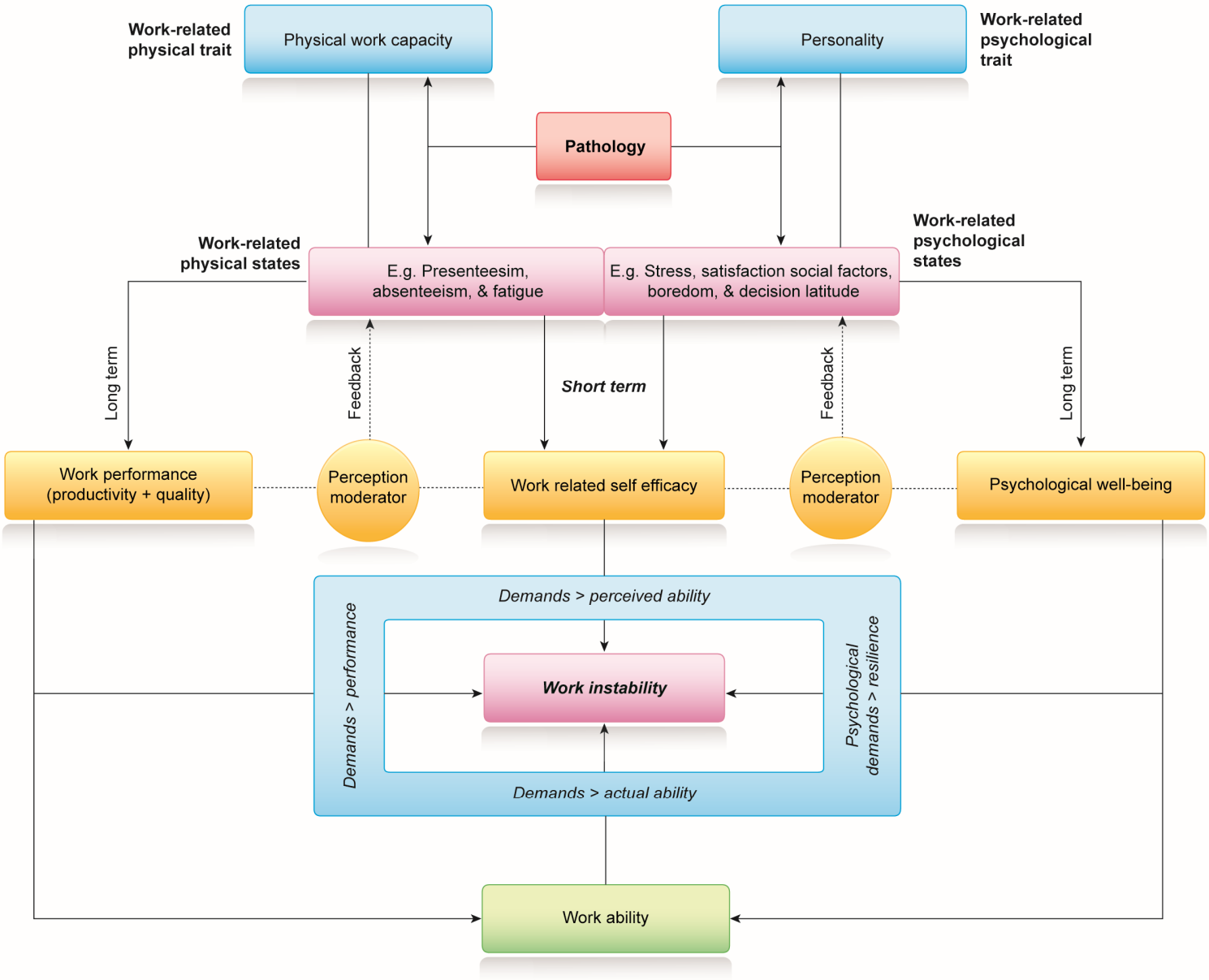

Work-related physical trait
Physical work capacity
Personality
Work-related psychological trait
Pathology
Work-related physical states
E.g. Presenteesim, absenteeism, & fatigue
E.g. Stress, satisfaction social factors, boredom, & decision latitude
Work-related psychological states
Long term
Feedback
Short term
Feedback
Long term
Work performance (productivity + quality)
Perception moderator
Work related self efficacy
Perception moderator
Psychological well-being
Demands > perceived ability
Demands > performance
Work instability
Psychological demands > resilience
Demands > actual ability
Work ability

Table 1 – An Evaluation of the Psychometric Properties of Outcome Measures related to the Psychological aspects of Work

Name of Outcome Measure/Tool	Psychometric Validation Study	Reliability			Validity					Responsiveness	Total Score
		Internal Consistency	Test-Retest,	Measurement Error	Content & Face	Criterion (Concurrent)	Construct (Structural)	Construct (Cross Cultural)	Construct (Hypothesis-testing)		
Personality Scales											
Dutch Workaholism Scale (DUWAS)[17]	Rantanen et al., 2015 [16]	1	0	0	1	0	2	0	1	0	5 (28%)
	Schaufeli et al., 2009 [17]	2	0	0	1	1	2	1	1	0	8 (44%)
Work Personality Profile[18]	Bolton et al., 1986 [18]	Unable to access data									-
	Siu et al., 1998 [19]	Unable to access data									-
	Law et al., 2006 [20]	0	0	0	1	0	2	N/A	0	0	3 (19%)
	Williams, 1997 [21]	2	1@	0	2	1	0	N/A	1	0	7 (44%)
Work Locus of Control Scale (WLOCS)[22]	Spector, 1988 [22]	Unable to access data									-
	Oliver et al., 2006 [23]	2	0	0	1	1	2	N/A	0	0	6 (38%)
Life Situation Questionnaire[24]	Crewe et al., 1991 [24]	Unable to access data									-
	Krause, 1998 [25]	2	0	0	1	1	2	N/A	1	0	7 (44%)
	Krause, 1992 [26]	2	2	0	1	1	2	N/A	0	0	8 (50%)
Stress Scales											
Work-related Stress Questionnaire[27]	Holmgren et al., 2009 [27]	0	!	0	2	0	0	N/A	0	0	2* (13%)
Work Stress Inventory - Dental Assistants[28]	Khader et al., 2009 [28]	2	2	0	2	0	2	N/A	0	0	8 (50%)
Nursing Stress Scale[31]	Rosnawati et al., 2010 [29]	2	2	0	1	0	0	0	0	0	5 (28%)
	Lee et al., 2007 [30]	2	2	0	1	0	2	2	0	0	9 (50%)
	Gray-Toft et al., 1981 [31]	2	2	0	2	1	2	N/A	2	0	11 (69%)
Expanded Nurse Stress Scale[32]	French et al., 2000 [32]	2	0	0	1	1	2	N/A	1	0	7 (44%)
US National Institute for Occupational Safety and Health Generic Job Stress Questionnaire[33]	Hurrell et al., 1988 [33]	2	0	0	1	1	0	N/A	0	0	4 (25%)
Satisfaction Scales											
Stamps and Piedmont's Index of Work Satisfaction[34]	Stamps et al., 1978 [34]	2	0	0	2	0	2	N/A	0	0	6 (38%)
McCloskey/Mueller Satisfaction Scale (MMSS)[35]	Mueller et al., 1990 [35]	2	1	0	1	1	2	N/A	0	0	7 (44%)
	Prosen et al., 2013 [36]	2	0	0	1	0	2	2	0	0	7 (39%)
	Tourangeau et al., 2006 [37]	1	0	0	2	1	2	N/A	2	0	8 (50%)
Minnesota Satisfaction Questionnaire [41]	Gillet et al., 1975 [38]	0	0	0	1	1	0	N/A	1	0	3 (19%)
	Asegid et al., 2014 [39]	1	0	0	1	1	0	N/A	1	0	4 (25%)
	Weng et al., 2010 [40]	2	0	0	1	1	2	N/A	1	0	7 (44%)

	Weiss et al. 1967 [41]	2	2	0	1	0	2	N/A	2	0	9 (56%)
Minnesota Satisfaction Questionnaire - SF [41]	Hirschfeld, 2000 [42]	0	0	0	1	0	2	N/A	1	0	4 (25%)
	Hancer et al., 2003 [43]	2	0	0	1	0	2	N/A	0	0	5 (31%)
	Weiss et al. 1967 [41]	2	2	0	1	1	0	N/A	2	0	8 (50%)
Boredom Scales											
Job Boredom Scale[44]	Lee, 1986 [44]	Unable to access data									-
Well-being Scales											
Oldenburg Burnout Inventory (OLBI)[49]	Tsutsumi et al., 2001 [45]	0	0	0	1	1	0	0	0	0	2 (11%)
	Lundkvist et al., 2014 [46]	0	0	0	1	1	2	N/A	0	0	4 (25%)
	Lekutle et al., 2014 [47]	1	0	0	1	1	2	N/A	0	0	5 (31%)
	Peterson et al., 2011 [48]	2	0	0	1	1	2	1	0	0	6 (33%)
	Demerouti et al., 2001 [49]	0	0	0	2	0	2	N/A	2	0	6 (38%)
	Qiao et al., 2010 [50]	2	0	0	2	1	2	N/A	2	0	9 (56%)
	Halbesleben et al., 2005 [51]	2	1	0	2	1	2	N/A	2	0	10 (63%)
Job-related Affective Well-being Scale (JAWS)[52]	Katwyk et al., 2000 [52]	Unable to access data									-
	Basinska et al., 2014 [53]	1	0	0	1	0	2	2	0	0	6 (33%)
	Makikangas et al., 2007 [54]	2	0	0	1	0	2	N/A	1	0	6 (38%)
Psychological Empowerment Scale[60]	Hancer, 2005 [55]	Unable to access data									-
	Hancer et al., 2005 [56]	Unable to access data									-
	Kraimer et al., 1999 [57]	0	0	0	1	0	2	N/A	1	0	4 (25%)
	Uner et al., 2010 [58]	2	0	0	1	0	2	1	0	0	6 (33%)
	Albar et al., 2012 [59]	2	0	0	1	0	2	2	0	0	7 (39%)
	Spreitzer, 1995 [60]	2	2	0	2	1	2	N/A	1	0	10 (63%)
Work & Well-being Survey (Utrecht Work Engagement Scale : 17 -UWES)[67]	Seppala et al., 2008 [61]	0	0	0	1	0	2	N/A	0	0	3 (19%)
	Nerstad et al., 2010 [62]	2	0	0	1	0	2	N/A	0	0	5 (31%)
	Lekutle et al., 2014 [47]	2	0	0	1	1	2	N/A	0	0	6 (38%)
	Simbula et al., 2013 [63]	2	0	0	1	0	2	N/A	0	1	6 (38%)
	Mills et al., 2011 [64]	2	0	0	1	1	2	N/A	0	0	6 (38%)
	Fong et al., 2011 [65]	2	0	0	1	1	2	N/A	0	1	7 (44%)
	Zecca et al., 2015 [66]	2	0	0	1	1	2	N/A	1	0	7 (44%)
	Schaufeli et al., 2003 [67]	2	1	0	2	1	2	N/A	0	0	8 (50%)
	Shimazu et al., 2008 [68]	2	0	0	1	1	2	2	0	0	8 (44%)
	Viljevac et al., 2012 [69]	2	0	0	1	1	2	N/A	2	0	8 (50%)
Utrecht Work Engagement Scale : 9-UWES[73]	Seppala et al., 2008 [61]	0	0	0	1	0	2	N/A	0	0	3 (19%)
	Nerstad et al., 2010 [62]	2	0	0	1	0	2	N/A	0	0	5 (31%)
	de Bruin et al., 2013 [70]	2	0	0	1	0	2	N/A	0	0	5 (31%)
	Mills et al., 2011 [64]	2	0	0	1	1	2	N/A	0	0	6 (38%)
	Simbula et al., 2013 [63]	2	0	0	1	0	2	N/A	0	1	6 (38%)

	Villotti et al., 2013 [71]	2	0	0	1	1	2	N/A	0	0	6 (38%)
	Zecca et al., 2015 [66]	2	0	0	1	1	2	N/A	1	0	7 (44%)
	Fong et al., 2011 [72]	2	0	0	1	1	2	N/A	0	1	7 (44%)
	Schaufeli, 2006 [73]	1	1	0	2	1	2	N/A	1	0	8 (50%)
	Wefald et al., 2011 [74]	2	0	0	1	1	2	N/A	2	0	8 (50%)
	Panthee et al., 2014 [75]	2	0	0	1	1	2	2	1	0	9 (50%)
	Klassen et al., 2012 [76]	2	0	0	1	1	2	2	1	0	9 (50%)
	Balducci et al., 2010 [77]	2	0	0	1	1	2	1	2	0	9 (50%)
	Littman-Ovadia et al., 2013 [78]	2	0	0	1	1	2	2	2	0	10 (56%)
Organizational Civility Scale (OCS)[79]	Clark et al., 2013 [79]	2	0	0	2	1	2	N/A	0	0	7 (44%)
Work and Organization Assessment Questionnaire (WOAQ)[80]	Wynne-Jones et al., 2009 [81]	2	0	0	1	0	2	N/A	0	0	5 (31%)
	Karimi et al., 2015 [82]	!	0	0	1	1	2	N/A	1	0	5 (31%)
	Griffiths et al., 2006 [83]	2	2	0	2	1	2	N/A	0	0	9 (56%)
Psychosocial Aspects of Work questionnaire[84]	Symonds et al., 1996 [84]	2	0	0	1	1	2	N/A	0	0	6 (38%)
Demand-Control-Support Questionnaire (DCSQ)[87]	Chungkham et al., 2013 [85]	!	0	0	1	0	2	N/A	0	0	3* (19%)
	Hokerburg et al., 2014 [86]	!	0	0	1	0	2	N/A	1	0	4* (25%)
	Sanne et al., 2005 [87]	1	0	0	1	0	2	N/A	0	0	4 (25%)
	Mase et al., 2012 [88]	2	1	0	1	1	2	2	0	0	9 (50%)
Capability Set for Work[89]	Abma et al., 2016 [89]	0	0	0	2	1	0	N/A	1	0	4 (25%)
Job Involvement Scale[90]	Lodahl et al., 1965 [90]	Unable to access data									-
	Corner et al., 1995 [91]	2	0	0	1	1	2	N/A	0	0	6 (38%)

! - Utilized statistics which are not included in our quality criteria, and therefore cannot be assessed. The total score is therefore followed by an "*" to illustrate that it does not account for all the psychometric measurements. @ - The authors have conducted an inter-rater reliability analysis.

Table 2 - ***An Evaluation of the Psychometric Properties of Outcome Measures related to the Physical aspects of Work***

Name of Outcome Measure/Tool	Psychometric Validation Study	Reliability			Validity					Responsiveness	Total Score
		Internal Consistency	Test-Retest,	Measurement Error	Content & Face	Criterion (Concurrent)	Criterion (Structural;)	Construct (Cross Cultural)	Construct (Hypothesis-testing)		
Physical Capacity Scales											
Fear-Avoidance Belief Questionnaire[107]	Holden et al., 2010 [92]	0	0	0	1	0	0	0	0	0	1 (6%)
	Ketenci et al., 2014 [93]	0	2	0	0	0	2	1	0	0	5 (28%)
	George et al., 2010 [94]	0	2	1	1	1	0	N/A	0	1	6 (38%)
	Inrig et al., 2011 [95]	2	1	0	1	1	0	N/A	1	0	6 (38%)
	S-Meewisse et al., 2003 [96]	2	2	0	1	1	0	N/A	0	0	6 (38%)
	Dedering et al., 2012 [97]	2	2	0	1	1	0	N/A	0	0	6 (38%)
	Kovacs et al., 2006 [98]	2	0	0	2	1	0	2	0	0	7 (39%)
	de Souza et al., 2008 [99]	2	2	0	1	1	0	1	0	1	8 (44%)
	Cleland et al., 2008 [100]	2	2	0	1	1	0	N/A	0	0	6 (50%)
	Mintken et al., 2010 [101]	0	2	0	1	1	2	N/A	0	2	8 (50%)
	Pfingsten et al., 2000 [102]	2	2	0	1	1	2	0	0	0	8 (44%)
	Laufer et al., 2012 [103]	2	2	0	1	1	0	1	0	0	8 (44%)
	A-Ashtiani et al., 2014 [104]	2	2	0	1	1	0	1	0	0	8 (44%)
	Terho et al., 2016 [105]	2	1	0	1	1	2	2	0	0	9 (50%)
	Chaory et al., 2005 [106]	0	2	0	1	1	2	2	0	1	9 (50%)
	Waddell et al., 1993 [107]	2	2	0	2	1	2	N/A	0	0	9 (56%)
	Matsudaira et al., 2014 [108]	2	0	0	1	1	2	2	2	0	10 (56%)
	Korkmaz et al., 2009 [109]	2	2	0	1	1	2	2	0	1	11 (61%)
	Pei et al., 2010 [110]	2	2	0	1	1	2	2	0	1	11 (61%)
	Rostami et al., 2014 [111]	2	2	0	2	1	2	2	0	0	11 (61%)
	Staerkle et al., 2004 [112]	2	2	0	1	1	2	2	0	1	11 (61%)
	Monticone et al., 2012 [113]	2	2	1	1	1	2	2	0	1	12 (67%)
	Grotle et al., 2006 [114]	2	2	1	1	1	2	2	0	2	13 (72%)
	Lee et al., 2006 [115]	2	2	0	2	1	2	2	2	1	14 (78%)
Dutch Musculoskeletal Questionnaire (DMQ)[116]	Hildebrandt et al., 2001 [116]	2	0	0	1	1	2	N/A	0	0	6 (38%)
Cornell Musculoskeletal Discomfort Questionnaire for Sedentary Workers[117]	Erdinc et al., 2011 [118]	2	1	0	2	1	0	1	0	0	7 (39%)
Absenteeism and Presenteeism Scales											
Rijeka Absenteeism Scale (RAS-6)[119]	Lalić et al., 2012 [119]	2	0	0	1	0	0	N/A	0	0	3 (19%)

Absenteeism Screening Questionnaire[120]	Truchon et al., 2012 [120]	2	1	0	2	1	0	N/A	0	2	8 (50%)
Work Productivity and Activity Impairment - General Health[123]	Gawlicki et al., 2006 [121]	0	0	0	1	0	0	1	0	0	2 (11%)
	Zhang et al., 2010 [122]	0	0	0	1	1	0	N/A	0	0	2 (13%)
	Reily et al., 1993 [123]	0	0	0	2	1	0	N/A	0	0	3 (19%)
	Ciconelli et al., 2006 [124]	2	2	0	1	1	0	1	0	0	7 (39%)
Work Productivity and Activity Impairment – Specific Health problem/Group[123]	Giovannetti et al., 2009 [125]	0	0	0	1	1	0	N/A	0	0	2 (13%)
	Wahlqvist et al., 2002 [126]	0	0	0	1	1	0	N/A	0	0	2 (13%)
	Lambert et al., 2014 [127]	0	0	0	2	0	0	1	0	0	3 (17%)
	Vergara et al., 2011 [128]	!	!	0	1	1	0	1	0	!	3* (17%)
	Ozcan et al., 2014 [129]	0	0	0	1	1	0	1	0	0	3 (17%)
	Wahlqvist et al., 2007 [130]	0	0	0	1	1	0	N/A	0	1	3 (19%)
	Reily et al., 2010 [131]	0	0	0	1	1	0	N/A	0	1	3 (19%)
	Reily et al., 2010 [132]	0	0	0	1	1	0	N/A	0	1	3 (19%)
	Vergara et al., 2009 [133]	0	!	0	2	1	0	1	0	0	4* (22%)
	Reily et al., 2004 [134]	0	2	0	1	1	0	N/A	0	0	4 (25%)
	Reily et al., 2008 [135]	0	0	0	1	1	0	N/A	2	1	5 (31%)
Single-item presenteeism question (SIPQ)[136]	Kigozi et al., 2014 [136]	0	0	0	1	1	0	N/A	2	1	5 (31%)
Lam Employment Absence and Productivity Scale (LEAPS)[139]	Pumpaisalchai et al., 2013 [137]	2	0	0	0	0	0	0	0	2	4 (22%)
	Lam et al., 2014 [138]	2	2	0	1	1	0	N/A	0	0	6 (38%)
	Lam, 2009 [139]	2	0	0	1	2	2	N/A	0	0	7 (44%)
Headache-Attributed Lost Time Index[140]	Steiner, 2007 [140]	0	0	0	2	0	0	N/A	0	0	2 (13%)
Performance (Productivity and Quality) Scales											
Quantity and Quality Instrument[141]	Meerding et al., 2005 [141]	0	0	0	1	1	0	N/A	0	0	2 (13%)
	Brouwer et al., 1999 [142]	0	0	0	1	0	0	N/A	1	0	2 (13%)
Finnish Institute for Occupational Health Questionnaire[143]	Puolakka et al., 2009 [143]	2	0@	0	2	1	0	N/A	0	0	7 (44%)
Work Role Functioning Questionnaire^	Durand et al., 2004 [144]	1	0	0	1	0	0	1	0	0	3 (17%)
	Abma et al., 2012 [145]	2	0	0	2	0	0	1	0	0	5 (28%)
	Gallasch et al., 2007 [146]	1	2	0	2	0	0	1	0	0	6 (33%)
	Ramada et al., 2013 [147]	2	2	0	2	0	0	1	0	0	7 (39%)
	Ramada et al., 2014 [148]	2	0	1	1	1	0	N/A	2	1	8 (50%)
	Ramada et al., 2014 [149]	2	0	0	1	1	2	N/A	2	0	8 (50%)
Work Role Functioning Questionnaire 2.0[151]	Abma et al., 2015 [150]	2	0	0	1	1	2	N/A	2	0	8 (50%)
	Abma et al., 2013 [151]	2	1	1	1	1	2	N/A	2	1	11 (69%)
Vocational Cognitive Rating Scale (VCRS)[152]	Greig et al., 2004 [152]	1	2@	0	2	1	0	N/A	0	0	6 (38%)
Occupational Functioning Scale (OFS)[153]	Hannula et al., 2006 [153]	0	0@	0	1	1	0	N/A	0	0	2 (13%)
The Work Assessment Rating Scale (WARS)[154]	Griffiths, 1973 [154]	0	2@	0	1	1	0	N/A	0	1	5 (31%)
Employment Expectation Questionnaire[155]	Millington et al., 2000 [155]	0	0	0	1	1	2	N/A	0	0	4 (25%)
Work Rehabilitation Questionnaire[156]	Finger et al., 2011 [156]	0	0	0	1	0	0	N/A	1	0	2 (13%)
Arthritis Specific Work Productivity Survey [157]	Osterhaus et al., 2009 [157]	0	0	0	1	0	0	N/A	0	1	2 (13%)
	Osterhaus et al., 2014 [158]	0	0	0	1	1	0	N/A	1	1	4 (25%)
	Osterhaus et al., 2014 [159]	0	0	0	1	1	0	N/A	2	1	5 (31%)
Sheehan Disability Scale[160]	Hambrick et al., 2004 [161]	1	0	0	1	1	0	N/A	0	0	3 (19%)
	Esmaeili et al., 2014 [162]	2	1	0	1	1	0	1	0	0	6 (33%)

	Hodgins et al., 2013 [163]	2	0	0	1	1	2	N/A	0	1	7 (44%)
	Cole et al., 2014 [164]	2	2	0	1	1	0	N/A	1	1	8 (50%)
	Luciano et al., 2010 [165]	2	0	0	1	1	2	1	1	2	10 (56%)
	Leu et al., 2015 [166]	2	2	0	1	1	2	2	1	2	13 (72%)
	Arbuckle et al., 2009 [167]	2	2	2	1	2	2	N/A	1	1	13 (81%)
Health and Labour Questionnaire (SF)[168]	No CTT psychometric data identified										-
Occupational Impact of Sleep Disorder Questionnaire[169]	David et al., 2007 [169]	*Unable to access data*									-
Health and Work Questionnaire[170]	Halpern et al., 2001 [170]	0	0	0	1	1	0	N/A	1	0	3 (19%)
	Schwarz et al., 2014 [171]	0	0	0	1	1	2	2	0	0	6 (33%)
	Shikiar et al., 2001 [172]	2	0	0	1	1	2	N/A	1	0	6 (38%)
Vocational Independence Scale (VIS)[173]	No CTT psychometric data identified										-
The Prolo Economic and Functional Rating Scale[174]	Vanti et al., 2013 [174]	*Unable to access data*									-
WHO Health and Work Performance Questionnaire[175]	No CTT psychometric data identified										-
Multiple Sclerosis Work Difficulties Questionnaire[177]	Ellenberger et al., 2015 [176]	1	0	0	1	1	2	2	0	0	7 (39%)
	Honan et al., 2012 [177]	2	0	0	2	1	2	N/A	1	0	8 (50%)
Multiple Sclerosis Work Difficulties Questionnaire – SF[178]	Honan et al., 2014 [178]	2	0	0	1	1	2	N/A	1	0	7 (44%)
Productivity Cost Questionnaire[179]	No CTT psychometric data identified										-

! - Utilized statistics which are not included in our quality criteria, and therefore cannot be assessed. The total score is therefore followed by an "*" to illustrate that it does not account for all the psychometric measurements. @ - The authors have conducted an inter-rater reliability analysis (the resulting ICCs were reported to be > 0.7). ^ - indicates there is no original citation for the tool [69].

Table 3 - ***An Evaluation of the Psychometric Properties of Outcome Measures related to Work Status, and Self Efficacy***

Name of Outcome Measure/Tool	Psychometric Validation Study	Reliability			Validity					Responsiveness	Total Score
		Internal Consistency	Test-Retest,	Measurement Error	Content & Face	Criterion (Criterion)	Criterion (Structural)	Construct (Cross Cultural)	Construct (Hypothesis-testing)		
Self-Efficacy Scales											
Rijeka Presenteeism Scale (RPS-6)[119]	Malec et al., 1993 [173]	*Unable to access data*									-
Lam Assessment of Stages of Employment Readiness (LASER)[181]	Lam et al., 2010 [180]	0	0	0	2	0	2	N/A	0	0	4 (25%)
	Chan et al., 2006 [181]	1	1	0	2	0	2	1	1	0	8 (44%)
Readiness For Return to Work[182]	Franche et al., 2007 [182]	1	0	0	1	1	2	N/A	2	0	7 (44%)
	Braathen et al., 2012 [183]	2	0	0	1	1	2	N/A	2	0	8 (50%)
Stanford Presenteeism Scale 6 (SPS-6)[186]	Laranjeira et al., 2013 [184]	2	0	0	1	1	2	0	0	0	6 (33%)
	Frauendorf et al., [185]	2	1@	0	2	1	0	1	0	0	7 (39%)
	Koopman et al., 2002 [186]	2	0	0	1	1	2	N/A	1	0	7 (44%)
	Hutting et al., 2013 [187]	2	2	0	1	1	2	2	0	0	10 (56%)
Stanford Presenteeism Scale 13 (SPS-13)[188]	Turpin et al., 2004 [188]	2	0	0	1	1	0	N/A	1	0	5 (31%)
Obstacles to Return-to-Work Questionnaire (ORQ)[189]	Marhold et al., 2002 [189]	2	2	0	1	1	2	N/A	0	0	8 (50%)
The Occupational Self-Efficacy Scale[190]	Schyns et al., 2002 [190]	2	0	0	1	1	2	N/A	1	0	7 (44%)
Endicott Work Productivity Scale[191]	Endicott et al., 1997 [191]	2	2	0	1	1	0	N/A	0	0	6 (38%)
Nurses Work Functioning Questionnaire[194]	Gartner et al., 2012 [192]	0	0	1	1	0	0	N/A	0	0	2 (13%)
	Gartner et al., 2011 [193]	0	2	0	1	1	0	N/A	0	1	5 (31%)
	Gartner et al., 2011 [194]	2	0	0	2	0	2	N/A	0	0	6 (38%)
The Migraine Work and Productivity Loss Questionnaire[195]	No CTT psychometric data identified										-
Work Status Scales											
Vocational Outcome Scale (VOS)[196]	No CTT psychometric data identified										-
Valuation of Lost Productivity (VOLP)[197]	Zhang et al., 2011 [198]	0	0	0	1	1	0	N/A	0	0	2 (13%)
Sydney Psychosocial Reintegration Scale[200]	Kulpers et al., 2004 [199]	*Unable to access data*									-
	Tate et al., 1999 [200]	2	2@	0	2	1	0	N/A	0	0	7 (44%)
	De Wold et al., 2010 [201]	2	0	2	1	1	0	N/A	1	0	7 (44%)
Work function score (WFS)[202]	No CTT psychometric data identified										-
Occupational Outcome Questionnaire[203]	No CTT psychometric data identified										-
Job Demands											
Demand-Induced Strain Compensation Questionnaire[204]	Bova et al., 2013 [204]	2	0	0	1	1	2	N/A	2	0	8 (50%)

@ - The authors have conducted an inter-rater reliability analysis (the resulting ICCs were reported to be > 0.7).

Table 4 - ***An Evaluation of the Psychometric Properties of Outcome Measures related to Work Instability***

Name of Outcome Measure/Tool			Psychometric Validation Study	Reliability			Validity					Responsiveness	Total Score
				Internal Consistency	Test-Retest	Measurement Error	Content & Face	Criterion (Concurrent)	Construct (Structural)	Construct (Cross Cultural)	Construct (Hypothesis-testing)		
Rheumatoid Arthritis - Work Instability Scale[205]			Beaton et al., 2010 [206]	2	0	0	1	1	0	N/A	1	1	6 (38%)
			Roy et al., 2009 [207]	2	0	1	1	1	0	N/A	0	2	7 (44%)
			Revicki et al., 2015 [208]	2	2	0	1	1	0	N/A	1	0	7 (44%)
Ankylosing Spondylitis - Work Instability Scale[209]			Frauendorf et al., 2014 [185]	2	0@	0	1	1	0	1	0	0	5 (28%)
Nurse -Work Instability Scale[210]			Harling et al., 2014 [211]	0	0	0	1	0	0	N/A	0	2	3 (19%)
			Harling et al., 2013 [212]	2	0	0	1	1	0	0	2	0	6 (33%)
			Harling et al., 2013 [213]	2	0	0	1	1	0	N/A	2	0	6 (38%)
Traumatic Brain Injury - Work Instability Scale[214]			No CTT psychometric data identified										-
Multiple Sclerosis - Work Instability Scale[215]			No CTT psychometric data identified										-
Epilepsy - Work Instability Scale[216]			No CTT psychometric data identified										-
Manual Work Instability Scale[217]			No CTT psychometric data identified										
Workplace Stress Scale[218]			No CTT psychometric data identified										-
Job Content Questionnaire[220]	# of Items	53	Tabatabaee et al., 2013 [219]	1	1	0	1	0	2	1	0	0	6 (33%)
		49	Karasek et al., 1998 [220]	2	0	0	2	1	2	N/A	0	0	7 (44%)
		49	de Araujo et al., 2008 [221]	2	0	0	1	1	2	2	0	0	8 (50%)
		45/22	Phakthongsuk et al., 2008 [222]	1	0	0	1	0	2	1	0	0	5 (28%)
		39	Poanta et al., 2006 [223]	*Unable to access data*									-
		39	Choobineh et al., 2011 [224]	1	0	0	1	0	1	1	0	0	4 (22%)
		35	Hoang et al., 2013 [225]	1	0	0	1	1	2	2	0	0	7 (39%)
		34	Amin et al., 2015 [226]	2	0	0	1	1	2	N/A	0	0	6 (38%)
		32	Eum et al., 2006 [227]	1	1	0	1	1	2	1	0	0	7 (39%)
		31	Kawakami et al., 1996 [228]	1	0	0	1	1	2	1	0	0	6 (33%)
		31	Brisson et al., 1998 [229]	2	1	0	1	0	2	2	0	0	8 (44%)
		31	Niedhammer, 2002 [230]	1	0	0	1	1	2	0	0	0	5 (28%)
		30	Li et al., 2007 [231]	*Unable to access data*									-
		29	Zreda et al., 2014 [232]	2	0	0	1	1	2	2	0	0	8 (44%)
		22	Alexopoulos et al., 2015 [233]	1	0	0	1	0	2	2	0	0	6 (33%)
		22	Li et al., 2004 [234]	1	0	0	1	0	2	1	0	0	5 (28%)

		22	Cheng et al., 2003 [235]	1	1	0	1	0	2	1	0	0	6 (33%)
		21	Maizura et al., 2009 [236]	1	0	0	1	0	0	N/A	0	0	2 (11%)
		21	Hadi et al., 2006 [237]	1	0	0	1	0	2	1	0	0	5 (28%)
		21	Nehzat et al., 2014 [238]	2	0	0	1	0	2	N/A	0	0	5 (31%)
		17	Choi et al., 2014 [239]	0	1	0	1	0	0	0	0	0	2 (11%)
		14	d'Errico et al., 2008 [240]	1	1	0	1	0	0	N/A	0	0	3 (19%)
		14	Sale et al., 2002 [241]	1	1	0	1	0	2	2	1	0	8 (44%)
		10	Choi et al., 2012 [242]	1	0	0	1	0	2	1	1	0	6 (33%)
Copenhagen psychosocial questionnaire - II (COPSOQ-II)[243]			Pejtersen et al., 2009 [243]	2	0	0	1	0	0	N/A	1	0	4 (25%)
			Moncada et al., 2013 [244]	2	0	0	1	1	2	2	0	0	8 (44%)
Effort-Reward Imbalance Questionnaire[252]	# of Items	47	Hanson et al., 2000 [245]	2	0	0	1	1	2	2	0	0	8 (44%)
		46	Tsutsumi et al., 2001 [246]	2	0	0	1	1	2	1	0	0	7 (39%)
		23	Almadi et al., 2013 [247]	2	0	0	1	0	2	2	0	0	7 (39%)
		23	Yadegarfar et al., 2012 [248]	1	0	0	2	1	2	2	0	0	8 (44%)
		23	Msaouel et al., 2012 [249]	2	0	0	1	1	2	2	1	0	9 (50%)
		23	Aboa-Éboulé et al., 2011 [250]	2	0	0	1	0	2	N/A	0	0	5 (31%)
		23	Zurlo et al., 2010 [251]	2	0	0	1	1	2	2	1	0	9 (50%)
		23	Chor et al., 2008 [252]	2	2	0	1	0	2	2	0	0	9 (50%)
		23	Griep et al., 2009 [253]	2	0	0	1	1	2	N/A	0	0	6 (38%)
		23	Buapetch et al., 2008 [254]	1	1	0	1	1	2	1	0	0	7 (39%)
		23	Siegrist et al., 2004 [255]	2	0	0	2	0	2	N/A	0	0	6 (38%)
		23	Juarez-Garcia et al., 2015 [256]	1	0	0	1	1	2	2	0	0	7 (39%)
		23	Eum et al.,2007 [257]	2	0	0	1	1	2	2	0	0	8 (44%)
		17	Gomez Ortiz, 2010 [258]	1	0	0	1	1	2	2	0	0	7 (39%)
		16	Magnavita et al., 2012 [259]	1	0	0	1	1	2	0	0	0	5 (28%)
Nordic Questionnaire for Psychological and Social Factors at Work[260]			Dallner et al., 2000 [260]	2	0	0	1	1	0	N/A	0	0	4 (25%)
			Wännström et al., 2009 [261]	2	0	0	1	1	2	N/A	0	0	6 (38%)
Workplace Activity Limitations Scale (WALS)[262]			Beaton et al., 2010 [206]	2	0	0	1	1	0	N/A	1	2	7 (44%)
Areas of Worklife Scale[263]			Gascon et al., 2013 [264]	2	2	0	1	1	2	2	0	0	10 (56%

@ - The authors have conducted an inter-rater reliability analysis (the resulting Interclass Correlation Coefficients (ICCs) were reported to be > 0.7).

Table 5 - ***An Evaluation of the Psychometric Properties of Outcome Measures related to Work Ability***

Name of Outcome Measure/Tool	Psychometric Validation Study	Reliability			Validity					Responsiveness	Total Score
		Internal Consistency	Test-Retest,	Measurement Error	Content & Face	Criterion (Criterion)	Criterion (Structural)	Construct (Cross Cultural)	Construct (Hypothesis-testing)		
Individual Work Performance Questionnaire (IWPQ)[265]	Koopmans et al., 2014 [266]	0	0	0	1	1	0	N/A	2	0	4 (25%)
Work Ability Index[271]	Leggett et al., 2015 [279]	*Unable to access data*									-
	Yang et al., 2013 [280]	0	2	0	1	0	0	0	1	0	4 (22%)
	Martus et al., 2010 [281]	1	0	0	1	0	2	N/A	0	0	4 (25%)
	Martinez et al., 2009 [275]	2	0	0	1	1	2	N/A	1	0	7 (44%)
	Kaewboonchoo et al., 2015 [276]	1	2	0	1	1	2	2	0	0	9 (50%)
	Abdolalizadeh et al., 2012 [277]	2	2	0	1	1	2	2	1	0	11 (61%)
Occupational Stress Indicator (OSI)[278]	Russinova et al., 1997 [279]	*Unable to access data*									-
	Davis, 1996 [280]	1	0	0	1	1	0	N/A	0	0	3 (19%)
	Swan et al., 1993 [281]	1	0	0	1	1	0	1	0	0	4 (22%)
	Lyne et al., 2000 [282]	2	0	0	1	0	2	N/A	0	0	5 (31%)
	Steiler et al., 2009 [283]	2	0	0	1	0	2	2	0	0	7 (39%)
Health-Related Productivity Questionnaire-Diary[284]	Kumar et al., 2003 [284]	0	0	0	2	1	0	N/A	0	0	3 (19%)
Spinal Cord Injury-Work Survey[285]	No CTT psychometric data identified										-
Work Limitation Questionnaire (25-items)[291]	Tang et al., 2012 [286]	0	0	0	1	0	2	N/A	0	0	3 (19%)
	Walker et al., 2005 [287]	0	0	0	1	0	2	N/A	0	0	3 (19%)
	Lerner et al., 2002 [288]	2	0	0	1	1	0	N/A	2	0	6 (38%)
	Beaton et al., 2010 [206]	2	0	0	1	1	0	N/A	1	1	6 (38%)
	Kono et al., 2014 [289]	2	1	0	1	1	2	N/A	0	0	7 (44%)
	Verhoef et al., 2012 [290]	2	1	0	1	1	0	N/A	2	0	7 (44%)
	Lerner et al., 2001 [291]	2	0	0	2	1	1	N/A	1	0	7 (44%)
	Takegami et al., 2014 [292]	2	0	0	2	1	2	1	0	0	8 (44%)
	Tamminga et al., 2013 [293]	2	1	2	1	1	0	N/A	2	1	10 (63%)
Assessment of the Subjective Handicap of Epilepsy[294]	O'Donoghue et al., 1998 [294]	2	2	1	2	1	2	N/A	2	0	12 (75%)
Work Experience Survey - Rheumatic Condition Manual[295]	Hammond et al., 2011 [295]	*Unable to access data*									-

@ - The authors have conducted an inter-rater reliability analysis (the resulting Interclass Correlation Coefficients (ICCs) were reported to be > 0.7).

Sink or S.W.I.M.; A Practical Model of Work for Vocational Rehabilitation

Supplementary Material

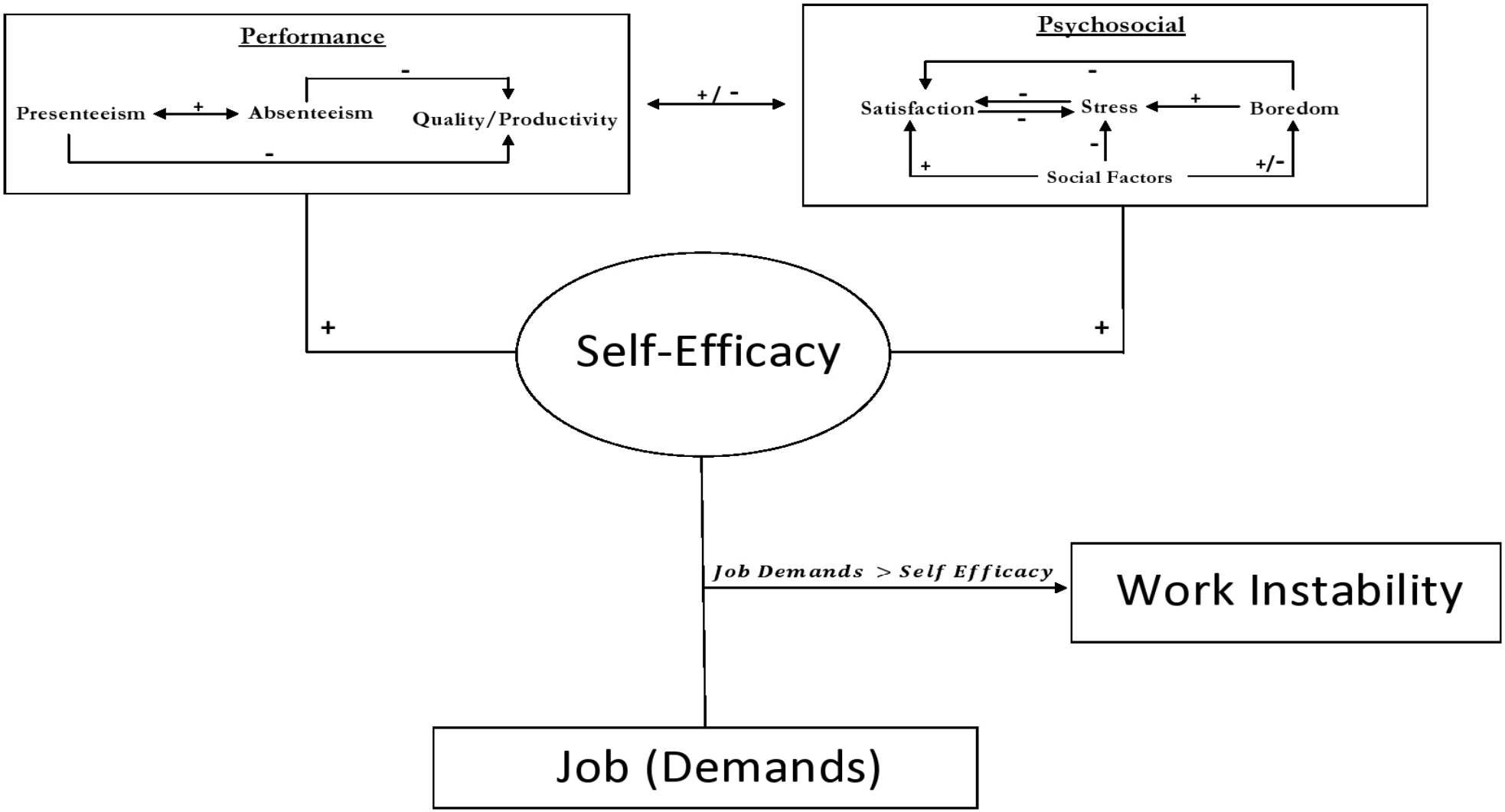

Performance
Presenteeism
+
Absenteeism
-
Quality/Productivity
-
+/ -
Psychosocial
-
Satisfaction
-
-
Stress
+
Boredom
-
-
+
Social Factors
+/-
+
+
Self-Efficacy
Job Demands > Self Efficacy
Work Instability
Job (Demands)

Figure S1 Legend - The Extended Work Instability Model

Initially the identified work related concepts were divided into two groups, relating to work-performance and the psychosocial aspects of work. The relationships between each of these concepts were derived from the literature.

Within performance there are three important terms that were identified during the 1st MDT exercise. The first is 'Presenteeism;, which is defined as "the problem of workers being on the job, but, because of illness or other medical conditions, not fully functioning" [12]. Next, there is Absenteeism; a "medically certified absence, where the absence from work is attributed to disease, medical condition or accident" [13]. The relationship of both of these concepts to productivity/quality can be found in their definition. For presenteeism the relationship is made explicit in the definition, whereas in absenteeism it can be inferred that the individual would be unable to perform at the level of an uninhibited worker, as they are physically absent from work, captured again by decreased productivity/quality of work.

The relationships between the four terms identified as psychosocial factors related to work, are largely based on experimental/clinical data rather than the definitional relationships used for the performance-related concepts. The satisfaction-stress relationship has been investigated several times, and there is broad ranging agreement that the high work-stress leads to lower levels of satisfaction, and vice versa, suggesting an inverse correlation [e.g. 15]. Boredom has a peculiar relationship with stress and social factors; depending on the individual with whom the social interaction occurs, boredom can be increased or decreased. For example, a co-workers description of a task, can have an impact on the workers attitude towards that particular task, including their perception of how boring it is, this effect is called 'social influence [16]. Boredom has also been linked to higher levels of dissatisfaction [17], and has even been linked with increasing the likelihood of stress symptoms [18]. And finally the relationship between social factors, stress, and satisfaction are the last remaining to be defined within the psychosocial taxa. Social support has been demonstrated to be protective against work stress [19]. And, quite surprisingly, one of the most crucial factors in determining work-related satisfaction are the work-related social factors [20].

These two modules (performance and psychosocial factors at work) are not completely independent. For example, the literature suggests that absenteeism and presenteeism are the end points of the same 'decision-making process' [y,z]. Taking time off work can lead to increased work-related stress and thus is more likely to drive the decision to attend work whilst sick (presenteeism), since absence is only more likely to further increase the stress burden. Therefore, negative psychological states mean that individuals are more likely to attend work, thereby decreasing absenteeism, but increasing presenteeism [25]. Moreover, on several occasions the effect of increased work-stress on lowering productivity has been illustrated [e.g. 24].

The definition of self-efficacy utilised is, "people's beliefs about their capabilities to produce designated levels of performance". These self-efficacy beliefs determine how people feel, think, motivate themselves and even behave [26]. Gist and Mitchell's review succinctly bring together the evidence for the relationship between self-efficacy, work-related performance, and psychological factors [27]. At this juncture it seems relevant to re-iterate that to as great a degree as possible, all of the concepts included in the model. are clinically measurable quantities, with outcome measures associated with them. The notable exceptions to this are job demand outcome measures. The parenthetical demands portion is representative of the fact that the associated tools measure a subjective variant of job demands, where the demand was contextualised to the individual's ability, akin to self-efficacy [26]. Although the authors of this paper disagree about the extent to which a measure of objective job demands is a viable measurement, we agree that none were identified in the original review [7]. The inclusion of job demands is therefore justified by its necessity in reaching the endpoint of work instability: a mismatch between an individual's functional capabilities and their job demands [30].

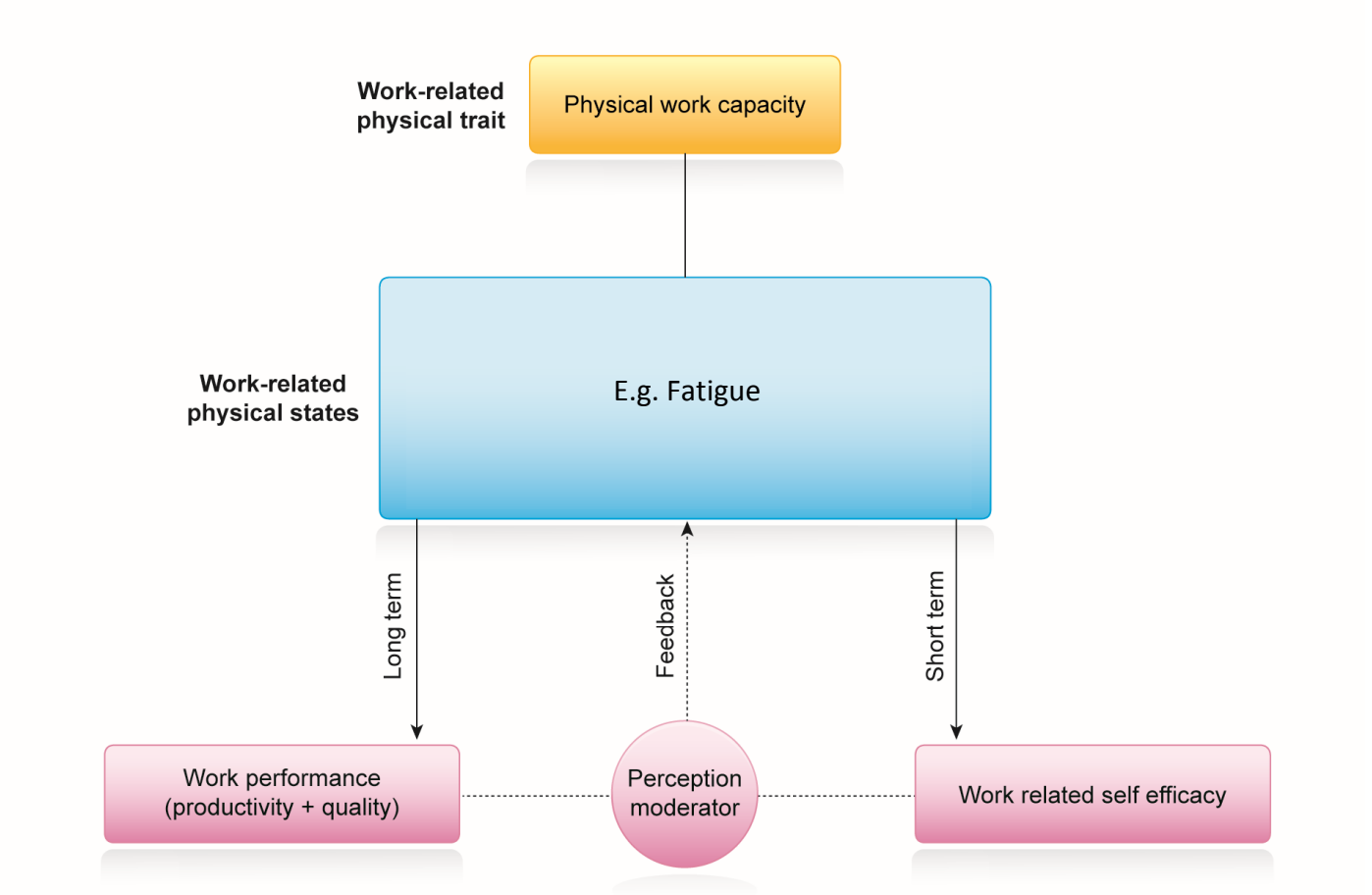


Figure S2
The physical module of the SWIM describes how physical work capacity moderated by different physical states (e.g. fatigue) reflects an individual's work performance, and their physical self-efficacy. The interceding perception moderator illustrates an important clinical effect, where mismatch between one's belief (self-efficacy) and objective measurements (performance) can further disable the individual.

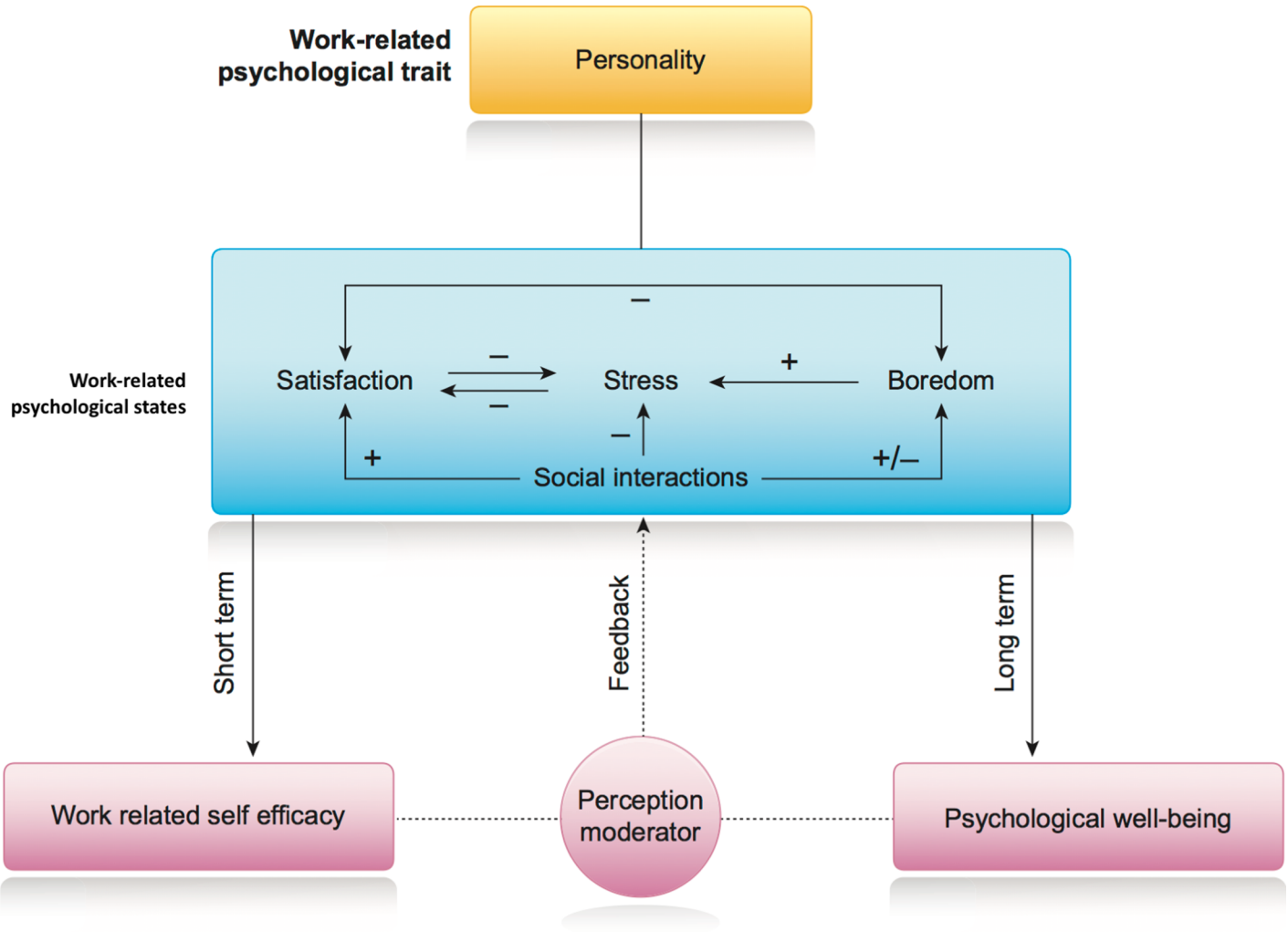


Figure S3
The psychological module of the SWIM describes how personality moderated by different psychological states (e.g. stress, satisfaction, boredom, and social interactions) reflects an individual's well-being, and their psychological self-efficacy. The interceding perception moderator illustrates an important clinical effect, where mismatch between one's belief (self-efficacy) and objective measurements (well-being) can further disable the individual.

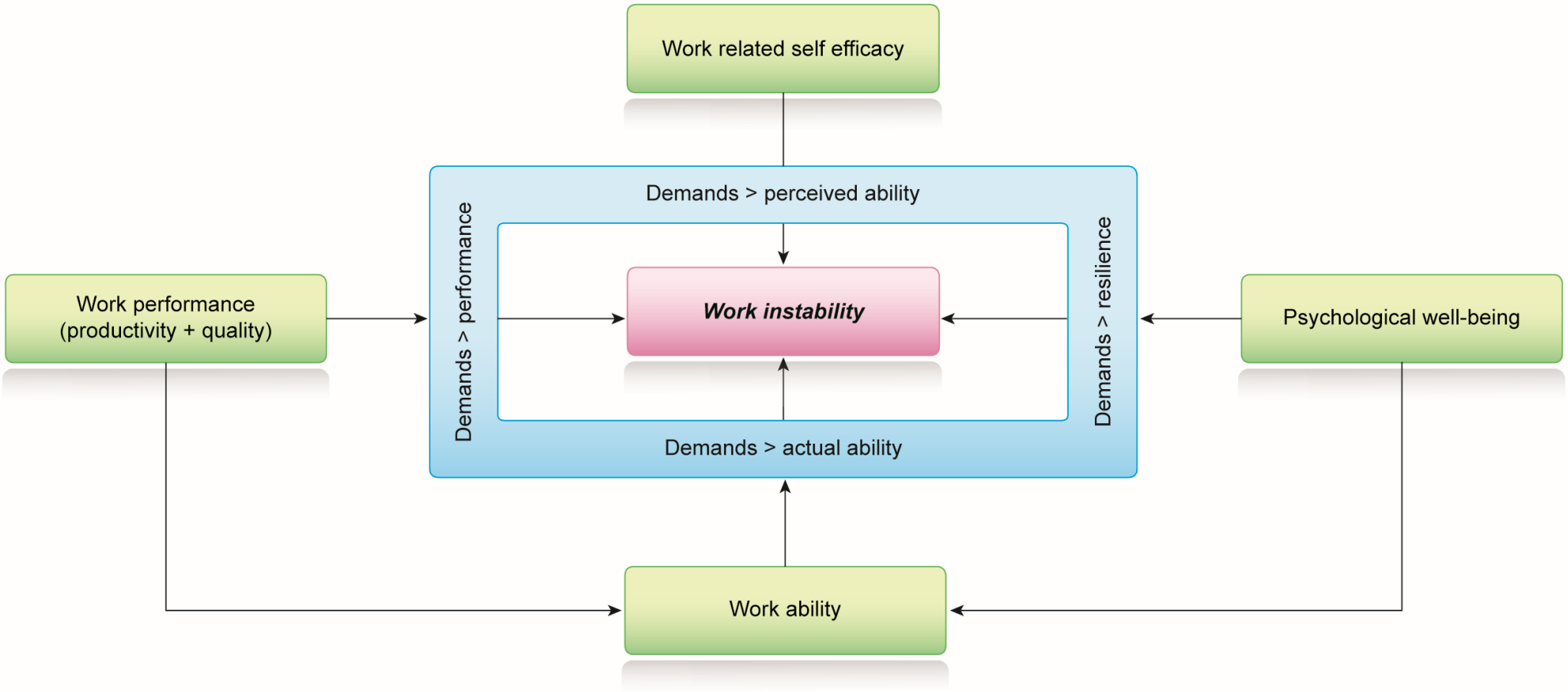


Figure S4
Work instability can occur as a result of any four potential inequalities between an individual's capabilities or perceived ability, and the demands associated with their work.