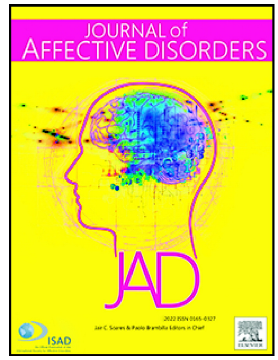
Journal Pre-proof



Measuring distress in older population: Rasch analysis of the Kessler Psychological Distress Scale

Cailen J. Calkin, Katya Numbers, Perminder S. Sachdev, Henry Brodaty, Oleg N. Medvedev

PII: S0165-0327(23)00284-7

DOI: https://doi.org/10.1016/j.jad.2023.02.116

Reference: JAD 15815

To appear in:

Received date: 7 September 2022

Revised date: 11 February 2023

Accepted date: 22 February 2023

Please cite this article as: C.J. Calkin, K. Numbers, P.S. Sachdev, et al., Measuring distress in older population: Rasch analysis of the Kessler Psychological Distress Scale, (2023), https://doi.org/10.1016/j.jad.2023.02.116

Measuring Distress in Older Population: Rasch Analysis of the Kessler Psychological Distress Scale

Cailen J. Calkin[1], Katya Numbers[2], Perminder S. Sachdev[2], Henry Brodaty[2], and Oleg N. Medvedev[1*]

[1]University of Waikato, School of Psychology, Hamilton, New Zealand.

[2]Centre for Healthy Brain Ageing (CHeBA), Discipline of Psychiatry & Mental Health, School of Clinical Medicine, University of New South Wales, Sydney, Australia

***Corresponding author:** **Oleg N. Medvedev, PhD**

School of Psychology, University of Waikato Private Bag 3105, Hamilton 3240, New Zealand. Phone: + 64 7 837 9212; Email: oleg.medvedev@waikato.ac.nz

Cailen J. Calkin, Email: cailencalkin@icloud.com

Katya Numbers, Email: k.numbers@unsw.edu.au

Perminder S. Sachdev, Email: p.sachdev@unsw.edu.au

Henry Brodaty, Email: h.brodaty@unsw.edu.au

Abstract

Objectives: The 10-item Kessler Psychological Distress Scale (K-10) is a widely applied distress measure; however, its psychometric properties were not established with older populations using advanced methodology. The aim of this study was to examine psychometric properties of the K-10 through application of Rasch methodology and if possible, develop an ordinal-to-interval conversion to improve its reliability in older populations.

Method: The Partial Credit Rasch Model was applied to analyse K-10 scores of the sample including 490 participants (56.3% females) aged 70 to 90 years and without dementia from the Sydney Memory and Aging Study (MAS).

Results: The initial analysis of the K-10 showed poor reliability and significant deviation from the expectations of the Rasch model. The best model fit was evident after correcting disordered thresholds and creating two testlet models to address local dependency between items ($\chi^2(35)=29.87$, $p=0.71$). The modified K-10 demonstrated strict unidimensionality, enhanced reliability and scale invariance across personal factors, such as sex, age, and education and permitted development of ordinal-to-interval transformation algorithms.

Limitations: Ordinal-to-interval conversion can only be applied for older adults with complete data.

Conclusions: The K-10 satisfied principles of fundamental measurement defined by Rasch model after minor modifications. Clinicians and researchers can transform K-10 raw scores into interval-level data using converging algorithms published here without altering the original scale response format, which increases reliability of the K-10.

Keywords: Rasch analysis; psychometrics; Kessler psychological distress scale; K-10; reliability; assessment

Measuring Distress in Older Population: Rasch Analysis of the Kessler Psychological Distress Scale

Distress is frequently referred to as an overall subjective stress reaction including affective symptoms of depression and anxiety (Matthews, 2016). Risks of cardiovascular disease, chronic obstructive pulmonary disease, and incident arthritis are all increased with elevated distress levels in older populations (McLachlan and Gale, 2018). Internationally, persons aged 65 years and older account for 9% of the population (703 million) (United Nations et al., 2020). Psychological distress has been found to be a risk factor for cognitive decline and dementia in older people. A number of studies found that older adults who experience symptoms of depression and anxiety, for example, may be at an increased risk for cognitive impairment and the development of dementia (Najjar & Lengenfelder, 2017; Rohleder & and Kirsh, 2011). Older populations are continuing to grow exponentially and are anticipated to triple by 2050 (World Health Organisation [WHO], 2011). Old age is associated with distress, decline in functionality, and cognition, which can impact older adults' mental well-being; therefore, the demand is evermore present for reliable and valid screening tools to measure distress in older populations (Vasiliadis et al., 2015). Accurately measuring distress in older people is imperative to devise optimal treatment plans that enhance quality of life and lessen chances of developing serious mental illness (Atkins et al., 2013; Kessler et al., 2003). These premises accentuate the significance of determining the most reliable and valid measurement of distress as predictor of psychopathological onset in older populations, which is useful to prevent harmful physiological outcomes, such as dementia (Vasiliadis et al., 2015).

The 10-item Kessler Psychological Distress Scale (K-10) is a widely applied measure of psychological distress suitable for both general and clinical populations (Peixoto et al., 2021). As the K-10 is a brief measure of distress, it is an appealing tool for screening of serious mental illness (Kessler et al., 2003). The K-10 has shown good construct and factorial validity, as well as worthy predictability (Easton et al., 2017; Kessler et al., 2002, 2003). K-10 validity has also been established through cross-cultural application, including Arabic/Palestinian cultures, Portuguese cultures, and Australian Aboriginal cultures (Bougie et al., 2016; Easton et al., 2017; Pereira et al., 2019). The original K-10 English version had a high level of internal consistency (Cronbach's α=0.93) (Kessler et al., 2002), and has maintained internal consistency in other applications of the English version (α=0.88) (Sampasa-Kanyinga et al., 2018). Moreover, acceptable internal consistency has been found in other countries, such as Argentina (Brenlla and Aranguren, 2010), Canada (Sampasa-Kanyinga et al., 2018), South Africa (Andersen et al., 2011), Japan (Sakurai et al., 2011), France (Arnaud et al., 2010), the Netherlands (Donker et al., 2010), and Portugal (Pereira et al., 2019). Although the internal validity of the K-10 is well established, there is lack of research with older populations.

The K-10 has previously been applied as a screening instrument in older populations to examine the correspondence between psychological distress and mood and anxiety disorders, generalised anxiety disorder (GAD), post-traumatic stress disorder (PTSD), and major depressive disorder (MDD) (Anderson et al., 2013; Vasiliadis et al., 2015). However, applying the K-10 to older populations should be taken with precaution as there are notable differences in physical functionality and illness among older adults that can increase psychological distress (Ormel et al., 1997). Consequently, these differences can cause issues with cut-off scores which may invalidate the application of the K-10 in older adult populations (Anderson et al., 2013). Only a few studies have validated the K-10 in older populations using traditional psychometric methods (Jong Won and Sun Hae, 2015; Lins et al., 2021). Although both studies found that the K-10 has acceptable reliability for measuring distress in older populations, more rigorous methods are needed to validate the use of the K-10 in older populations (Jong Won & Sun Hae, 2015; Lin et al., 2021). For example, Rasch analysis is more advantageous compared to other measurement models because it offers possibility for conversion of ordinal scores into interval level data, which ultimately improves the reliability of the assessment (Tennant & Conaghan, 2007).

The K-10 is an ordinal scale and ordinal-level data only differentiates response options by rank order. Moreover, traditional analyses do not account for item difficulty and personal ability, therefore limiting precision of assessment (Hobart & Cano, 2009). The ordinal K-10 scale assumes all items contribute equally to the overall distress score, however, some symptoms are more severe than others. For example, item 9 ‘so sad nothing could cheer you up’ indicates a higher severity of distress over Item 1 ‘tired for no good reason’. Therefore, it is necessary to acknowledge differences in item contribution to the overall distress score. Rasch analysis can increase the precision of measurement of the K-10 as it considers personal distress levels and the differential contribution of each item to the overall distress score (also known as item difficulty/location) (Medvedev and Krägeloh, 2022). Rasch model defines an interval scale that satisfies the fundamental principles of measurement. This can ultimately improve precision of distress assessment if satisfactory fit to the Rasch model is achieved (Bond et al., 2007).

The Rasch measurement model (Rasch, 1960; Tennant & Conaghan, 2007) is a statistical psychometric method applied to assess compliance of individual items to the overall scale with the fundamental principles of measurement, including invariance across personal factors, such as sex, age, education, and culture; unidimensionality, and uniformity of measurement units across the scale continuum (Tennant and Conaghan, 2007; Thurstone, 1931). As there are differences in item difficulty across items, distances between ordinal-level data cannot be attained, creating diagnostic bias when items are interpreted as the same value, which invalidates the total score (Marateb et al., 2014; Merkin et al., 2020). According to the Rasch model, only two parameters influence individual item responses: personal ability (e.g. personal distress level) and item difficulty (e.g. severity of a symptom reflected by an item) (Medvedev and Krägeloh, 2022). Rasch analysis converts ordinal scale scores into interval level data using both personal ability and item-difficulty, after the fundamental principles of measurement are satisfied (Medvedev & Krägeloh, 2022). Through the inclusion of both parameters, an authentic contribution of items is accounted for in the overall distress score, thereby eliminating assessment bias. Rasch analysis can examine fundamental measurement parameters of both the scale and individual items, which helps to identify issues (e.g., misfit, disordered thresholds) and address them, if possible, to increase accuracy and functionality of a scale (Christensen et al., 2012).

To date, Rasch analysis has been applied to the Bangladesh and Chinese versions of the K-10 using interview format (Uddin et al., 2018; Zhu et al., 2021) but there have been no accounts of Rasch analysis evaluating the original K-10 in older populations. Therefore, the present study aimed to evaluate and if possible, improve psychometric properties of the original K-10 in an older population sample by applying Rasch methodology. To do this, if an acceptable fit to the Rasch model is achieved, a conversion algorithm for ordinal-to-interval data transformation would be created to enhance accuracy of measurement. This study is essential to ensure the application of the K-10 in older populations is reliable and valid, and will allow researchers and clinicians to accurately determine distress levels and assess risks and effectiveness of interventions to lessen the chance of onset of a serious mental illness (Atkins et al., 2013; Kessler et al., 2003).

Method

Participants

Participants (*n*=1037) were extracted from the Sydney Memory and Aging Study (MAS study) based out of the University of New South Wales (UNSW), which is a community-based longitudinal cohort study of older adults aged 70 to 90 years old without dementia (Sachdev et al., 2010). The authors' institutional ethics committee approved the study. In 2005, 8914 older adults from the Eastern suburbs of Sydney were invited to participate. After applying exclusion criteria, the baseline MAS sample comprised 1037 older Australians. (Sachdev et al., 2010). MAS participants were followed-up bi-annually – called a Wave – over six waves (12 years).

It is important to have people with different levels of distress and ages when conducting Rasch analysis of the distress measure (K-10) in an aging population for several reasons. First, the validity of the K-10 measure in an aging population depends on having a sample that represents the full range of distress levels in this population. This includes individuals with high levels of distress as well as those with low levels of distress (Meade et al., 2018). Second, having individuals of different ages in the sample is important because the level and type of distress can vary with age. Third, having individuals with different levels of distress and ages can also help to address measurement invariance issues that may arise in the Rasch analysis (Wright & Masters, 1982).

MAS participants were followed-up bi-annually – called a Wave – over six waves (12 years). The data in the present study was extracted from the first and last assessment of the MAS study (baseline/Wave 1 and Wave 6). Figure 1 displays the selection method for the participants in the present study (*n*=490). An appropriate sample size for Rasch analysis is between 250 and 500 participants as it minimises Type I and Type II errors (Hagell and Westergren, 2016). To represent the full spectrum of distress levels in the sample and to account for temporal difference in data collection, we extracted an upper K-10 scores minority (M=14.00, SD=4.18) of the sample from Wave 6 (n=240, 57.5% females), because participants with higher distress level were mainly identified in this wave. To achieve an adequate sample for the current Rasch analysis (n=490) we then randomly selected 250 participants from the independent sample at Wave 1 (n=784), who were not included in the study at Wave 6. We then combined the two independent extractions from Wave 1 (n=250) and Wave 6 (n=240) to achieve the final overall sample (n= 490) for Rasch analysis. Additionally, 29.2% of participants from Wave 1 sample competed high school education and 29.2% completed primary education. The Wave 6 sample included 32.5% participants who competed high school education and 67.5% primary education. There were no significant disproportions across sex and education levels in these subsamples.

<Insert Figure 1 Here>

Measures

The K-10 (Kessler & Mroczek, 1994) is a 10-item measure of psychological distress that has been applied as a screening tool for mental illness (Kessler et al., 2003). Distress levels are measured from the past month on a 5-point Likert-type scale. For example, “In the past 4 weeks, about how often did you feel nervous?” is rated “All of the time (5)”, “Most of the time (4)”, “Some of the time (3)”, “A little of the time (2)”, or “None of the time (1)”. After completing the 10 items, scores range from 10 – 50, with higher scores representing higher distress levels (Kessler et al., 2003). Cronbach’s alpha and McDonald’s omega reliability estimates were 0.82 with the current dataset.

Data Analyses

Descriptive statistics of the sample and reliability of the K-10 were computed using IBM SPSS v.28. The Rasch model RUMM2030 software package was used to conduct Rasch analysis (Andrich, 1978). Rasch analysis involves two polytomous models that both assume variability throughout item-response category thresholds: the Rating Scale model (Andrich, 1978) and the Partial Credit model (Masters, 1982). The Partial Credit model allows threshold distances across items to vary, allowing one to disregard the item uniformity assumption (Masters, 1982). In contrast, the Rating Scale model assumes uniform threshold distribution across items (Andrich, 1978). To determine which model should be more suitable, the likelihood-ratio test compares thresholds distances across individual items (Tennant & Conaghan, 2007). A significant test result suggests differences across item thresholds and, therefore, rejects the Rating Scale model and supports appropriateness of the unrestricted Partial Credit model.

Rasch analysis began by evaluating individual item fit and overall data fit to the Rasch criteria. Overall data fit was assessed by an item-trait interaction using chi-square ($p > 0.05$, Bonferroni adjusted); this statistic should be non-significant for acceptable overall data fit. Additionally, person-item fit residual means should be near 0.00, with standard deviations of 1.00. Individual item fit residuals should fall between ±2.50. Differential item functioning (DIF), items working substantially different for separate groups (e.g., age groups, sex, education), should not be significantly different. This is tested through ANOVA between relevant sample groups. Items that exhibit poor fit or DIF are commonly removed, however, this can offset construct validity (Medvedev & Krägeloh, 2022). Therefore, item removal would only be used as a last resort. Instead, locally dependent items would be combined into testlets to attempt to improve overall model fit. This involved analysing residual correlations between items as this can influence the fit to the Rasch model. Local dependency is determined by residual correlation above 0.20 compared to the mean of all residual correlations (Christensen et al., 2017). After each modification, overall model fit is tested (Tennant & Conaghan, 2007) until an acceptable model fit is obtained. Person separation index (PSI) is the reliability measure used in Rasch analysis. Once individual items and the overall fit to the model are acceptable, and unidimensionality is confirmed, Rasch analysis is completed.

Dimensionality in Rasch analysis is tested through evaluating the first principal component after removing the latent trait (e.g., distress; Smith Jr., 2002). Subsequently, independent *t*-tests are applied with binominal confidence intervals (± 95%) to compare person-locations for two groups of items. The first group consists of items with the highest positive loadings on the first principal component of the residuals, whereas the second includes items with the lowest negative loadings. Therefore, to deem the scale unidimensional, the lower bound of the confidence intervals must be less than 5% in the person-estimates *t*-tests (Tennant and Pallant, 2006). After the data meets Rasch model expectations, conversion of ordinal-to-interval level data can be produced through a transformation table based on person estimates. Converting ordinal-level data into interval-level data was done to improve precision of assessment. The cut-off point for statistical significance was 0.05.

Results

A likelihood-ratio test for the K-10 sample data was computed to determine whether the Rating Scale (Andrich, 1978) or the unrestricted Partial Credit Model (Masters, 1982) would be used based on the uniformity of item thresholds. The results indicated statistically significant chi square ($\chi^2(26)=88.36$, $p<0.001$); therefore, variation of individual item thresholds violated assumptions of the Rating Scale in favour of the Partial Credit Model.

Analysis of the initial K-10 items suggested an overall poor fit to the Rasch model as indicated by a significant interaction between latent traits and items ($\chi^2(40)=77.81$, $p<0.001$) and poor reliability (PSI=0.61), but unidimensionality was clearly evident (see Table 1, A1). Item-fit residuals were checked to identify any misfitting items falling outside the ±2.50

standard deviations (SD) (Tennant & Conaghan, 2007). All individual item fit residuals fell within ±2.50 SDs except item 3, which was marginally misfitting, as displayed in Table 2. However, disordered thresholds were identified in all items of the K-10 suggesting that threshold parameters associated with an item lack sequential increase on ordinal scale response categories (Adams et al., 2012). For example, Figure S1 (upper panel) displays item 1's category probability curves, which depict disordered thresholds. As shown, the probability of a participant choosing response option 4 after 2 is greater than 3. This implies that the response categories lack consistency in reporting sequential progression in the latent trait (e.g., distress). To correct disordered thresholds, we uniformly rescored items by collapsing response options 3 "A little of the time" and 4 "None of the time" in RUMM2030, which corresponds to ordinal scale response options 4 and 5, as these categories' thresholds were disordered across most K-10 items. After this modification was implemented, all thresholds became ordered (see Figure S1, bottom panel), except item 10, which showed only a marginal disordering. Thresholds of other items before and after rescoring are presented in Supplementary Figures S1-S11. As expected, threshold ordering improved the overall fit ($\chi^2(40)=77.18$, $p<0.001$) and reliability of the scale, however, the model fit was not yet acceptable to permit ordinal-to-interval transformation (Table 1, A2).

<Insert Table 1 Here>

The residual correlation matrix was examined for local dependency, which is a strong association of two or more items beyond the latent variable that can impact on reliability, infringe the local independency assumption, and jeopardise estimation of model parameters (Medvedev and Krägeloh, 2022; Wright, 1996). Previous studies have found that residual correlations exceeding 0.20 reflect local dependency (Christensen et al., 2017). Residual correlations were identified between two groups of items in our study (i.e., 1, 5, 6, 7) and (i.e., 2, 8, 10). Removal of locally dependent or misfitting items is a common method of modification to improve the Rasch model, however, this can offset the construct validity of a measure (Medvedev & Krägeloh, 2022); therefore, it should only be done if there is evident redundancy or semantic and/or conceptual unsuitability of an item(s) (Medvedev & Krägeloh, 2022).

To resolve local dependency issue, testlets were created as an alternative strategy following a previously well-established methodology (Lundgren-Nilsson et al. 2013; Medvedev & Krägeloh, 2022). Note, that this testlet methodology is only applicable in scales reflecting one overarching latent trait because multidimensional scales with testlets fail to achieve Rasch model fit (Mitchell-Parker et al., 2018). Testlets were constructed using locally dependent items as follows: Testlet 1 (Items: 1, 5, 6 and 7); and Testlet 2 (Items: 2, 8 and 10). Items 3, 4, and 9 were excluded from the testlets as they did not significantly contribute to local dependency. Subsequently, the best model fit was achieved with non-significant item-trait interactions ($\chi^2(25)=29.87$, $p=0.71$) as displayed in Table 1, A3. These modifications improved psychometric properties of the scale resulting in strict unidimensionality, invariance (no DIF) by age, sex, and education factors, enhanced reliability (PSI=0.72) and appropriately ordered thresholds as shown in Figure S11.

We also compared individual scores (locations) of the original scale before modifications (analysis A1) with the modified final K-10 version (analysis A3) using paired *t*-test, which indicated a significant difference between person estimates with a large effect size ($t(489)=12.11$, $p<0.001$, Cohen's $d=0.55$). This may reflect a significant gain in accuracy because Rasch transformation considers the unique contribution of each test item to the overarching construct of psychological distress as measured by the K-10. Accuracy in this context refers to compliance of the K-10 to the criteria of an interval level scale defined by the Rasch model and is in line with the principles of the fundamental measurement

(Medvedev & Krägeloh, 2022). This is important given that a main goal of this study was to achieve parameters of an interval scale (e.g., temperature, height, etc.) for the K-10.

<Insert Table 2 Here>

After the best Rasch model fit was confirmed, a person-item threshold distribution plot was used to evaluate how adequate the item thresholds cover the distress levels across the sample. Figure 2 displays the person-item threshold distribution, demonstrating an acceptable coverage of sample distress scores by item-thresholds of the scale with no evidence of ceiling or floor effects. Additionally, the sample mean is below the item mean, suggesting the overall sample had low levels of distress, which justifies inclusion of subsample from wave 6 with the highest distress levels to permit evaluation of the scale across different levels of distress. This also implies that the newly modified scale has potential to measures higher distress scores, which is especially useful for clinical populations.

<Insert Figure 2 Here>

As the modified K-10 had an acceptable Rasch model fit and demonstrated excellent coverage across all distress levels of the sample, an ordinal-to-interval conversion table was created based on Rasch model estimates (Table 3). Adjusting raw data scores to interval-level scores enhances accuracy of the K-10 and allows for parametric statistical analysis of the transformed data (Hopkins et al., 2021). Table 3 provides the interval-level scores in both logit units and the rescored scale range and can be used in future studies for transforming K-10 raw scores into interval-level data. As shown in Table 3, the conversion table is easy to apply using the following instructions: rescore all items (e.g., 1=1, 2=2, 3=3, 4=4, 5=5), compute total scores by adding individual item scores together: corresponding interval-level scores are found on the right-hand side in logits and the rescored scale range in Table 3.

<Insert Table 3 Here>

Discussion

The present study provides compelling evidence for the psychometric validity of the K-10 as an interval scale, after making some minor modifications to meet the expectations of the unidimensional Rasch model. The original version of the K-10 suffered from disordered thresholds and poor reliability, however, these issues were successfully addressed by using modern Rasch methodology without removing any items from the scale. As a result, the modified K-10 showed strict unidimensionality, improved reliability (PSI=0.71), and scale invariance across different demographic factors such as gender, age, and education. The person-item threshold distribution indicated that the modified K-10 is capable of capturing all levels of distress without any ceiling or floor effects, making it a valuable tool for identifying distress in older populations. To enhance the accuracy of the K-10, an easy-to-use ordinal-to-interval table was developed based on the best Rasch model fit. This study has important implications for both research and clinical practice as it demonstrates the reliability and validity of the K-10 as a measure of distress.

Disordered thresholds were identified in all items within the original K-10 scale. Disordered thresholds refer to the lack of sequential progression in response options within the scale. For example, for item 1 “tired for no reason”, our older sample was most likely to choose the response option 2 “A little of the time”, followed by response option 4 “None of the time”, while likely missing the response option 3 “A little of the time”. Therefore, a linear increase of distress was not progressively reflected by response options. Tennant & Conaghan (2007) proposed that rescoring response options is a robust method of correcting disordered thresholds. Therefore, response option 3 “A little of the time” was collapsed with response option 4 “None of the time” across all items to achieve a progressive response option probability (Figure S1). After rescoring response options, a progressive increase of distress levels was perfectly reflected by the increase of response options achieved using testlet models. This highlights the statistical benefits of rescoring disordered thresholds, which has been successfully applied in previous studies (Medvedev et al., 2016, 2017; Sierra-Siegert & David, 2007).

One specific advantage of Rasch analysis is identification of local dependency between items, which is an association between two or more items that are not related to a latent trait being measured (e.g., distress) and can occur due to different reasons including item wording (e.g., negatively worded items; Medvedev & Krägeloh, 2022). Local dependency affects reliability, violates unidimensionality criteria, and jeopardises estimation of model parameters (Medvedev & Krägeloh, 2022; Wright & Stone, 1979). All items in the present study exhibited local dependency except item 5 and item 3, which were only marginally misfitting. The issue of local dependency was eliminated by applying two testlets models, designed based on residual correlations and aimed to reduce individual item errors and spurious correlations. Testlets (or super-items) have been successfully applied in various Rasch analyses to improve psychometric properties of measures without removing any items (An and Yu, 2021; Finaulahi et al., 2021; Hopkins et al., 2021; Merkin et al., 2020; Truong et al., 2021).

After satisfactory Rasch model fit was established, an ordinal-to-interval conversion table was created so that researchers could apply an algorithm that enhances precision of the K-10 up to an interval scale, which is useful to ensure reliability and validity for future older populations studies that may wish to include the K-10. This is a significant benefit of Rasch analysis that other more traditional validation methods cannot offer. Previous studies applying Rasch analysis to the K-10 have not implemented an ordinal-to-conversion table, which limits the practical applicability of their findings for future research. Converting ordinal-level data into interval-level data enhances measurement through disclosing exact intervals between total scale scores. For convenience of users, we have presented interval scores using the same range as an ordinal scale after rescoring disordered categories. For example, a total distress score of 20 corresponds to an interval-level score of 25.65 suggesting a higher level of distress compared to the ordinal scale estimate. As Rasch transformed scores represent interval-level data, parametric analyses can now be applied to generate more in-depth information that ordinal-level data cannot achieve as authentic values are unknown (Mircioiu and Atkinson, 2017).

The K-10 has been demonstrated to be a reliable and valid measure in older populations (Min & Lee, 2015), which may impose the question as to why convert ordinal-level data into interval-level data. Ordinal-level data does not account for item difficulty and personal ability (Tennant & Conaghan, 2007); therefore, true distinctions cannot be made about the differences in distress levels based on overall ordinal K-10 scores. This also limits our ability to distinguish whether interventions are appropriately working (Barber et al., 2022). For example, in a hypothetical scenario, Person X scores 23 in the K-10, whilst Person Y scores 33. Both individuals undergo an intervention to mitigate their distress and subsequently score lower; Person X scores 13 and Person Y scores 23. Though - both Person X and Y have scored 10 points lower, we cannot assume that distress has reduced equally as we have not accounted for differences in measurement units across the scale continuum (Barber et al., 2022). However, using the Rasch conversion table (Table 3), Person X has reduced their distress by 9.23 units, whereas Person Y has only reduced their score by merely 4.44 units. This demonstrates the substantial differences between true reductions in distress levels of these hypothetical persons when genuine interval scale is applied. Therefore, Rasch transformation is important to accurately evaluate and monitor distress levels in older populations (Barber et al., 2022).

Our modified K-10 has been shown to encompass all levels of distress in an older population. The positive skew in the K-10 scores from the current sample is expected as the current sample is non-clinical and scored low for symptoms of distress, where a clinical sample is expected to be negatively skewed. The bottom graph displays the item and item-option severity distributions, which appears to be negatively skewed. This implies that the modified K-10 can accurately measure high distress scores, despite this not being strongly reflected in the current sample. Therefore, converting raw scale scores from the K-10 into interval-level data using the Rasch conversion algorithm may be useful for assessing distress in clinical and non-clinical older adult samples more accurately.

Overall, this analysis indicated that the K-10 was inadequate for screening prior to modification. As the K-10 is a widely applied tool for measuring distress, it is imperative that it accurately screens distress in older populations to allow appropriate treatment plans to be constructed to prevent the onset of severe mental illness. Older populations are expected to triple by 2050 (World Health Organisation, 2011) and distress is prevalent in this population (Chittleborough et al., 2011), therefore having accurate tools to measure distress is important to mitigate distress symptoms prior to the progression of more severe mental illnesses. The conversion algorithm can be applied in future studies to enhance the precision of measurement in older populations to allow researchers and clinicians to formulate comprehensive treatment plans based on distress levels, emphasising the relevance of the conversion algorithm created by the current analysis. The conversion algorithm offers a linear expression of distress scores through accounting for unequal difficulties across scale items, which enhances the ability to accurately diagnose and address distress symptoms without any ambiguity. Therefore, this tool is evermore salient as older populations continue to grow, which increases the demand for accurate screening tools for mental illness; highlight the novelty of the conversion algorithm.

The present study is novel in applying Rasch analysis to the K-10 in older populations and using advanced testlet methodology to improve the reliability of the scale without modifying its response format, thereby, reducing measurement error and method effects. Previous studies have applied Rasch methodology to the K-10, however, have removed items and not created a ordinal-to-interval conversion table, which potentially affected construct validity and reduced its applicability in future research (Uddin et al., 2018). The present study was the first to create an ordinal-to-interval conversion table to enhance the accuracy of distress scores in older populations which allows future researchers to apply the algorithm to improve precision of distress measurement. It is imperative that psychometric assessments are accurate to provide true representation of well-being especially with the growing demand for mental health supports for older adults. Accurate distress measurement will allow us to use the K-10 more accurately as a screening tool for psychological distress, which may help reduce or even prevent progression to serious mental illness (Kessler et al., 2003). Preventing significant distress enhances quality of life among older adults (Atkins et al., 2013). Therefore, the novelty of this study is emphasised in the improvements of the K-10, which may benefit older adults by accurately capturing psychological distress allowing for early detection and interventions and more tailored mental health plans for older individuals.

Limitations and Directions for Future Research

This study has limitations that must be acknowledged. The sample used in the present study was collected from a small catchment area (the Eastern suburbs) in Sydney, Australia. Therefore, the sample is predominantly White European, affluent and well-educated. In this sense, it could be problematic to assume the conversion table is valid across different cultures considering the differences between in results of the Bangladeshi version, despite validation of the K-10 use in various cultures (Peixoto et al., 2021). There are limitations in the MAS dataset as well. Participants who responded through the six waves are not representative of the entire sample at wave 1 (non-random attrition) – therefore, participants from wave 6 are likely to not only be younger, more educated, and have higher cognition function, but also less distressed, as this is associated with withdrawal. Although Rasch analysis is generally considered less sample dependent compared to other methods (Tennant and Conaghan, 2007), and the sample size in the present study was adequate to minimise Type I and Type II errors, future studies must focus on validating this methodology using the K-10 in other ethnicities, especially those in low-and-middle-income countries.

Conclusions

Overall, the findings of this study indicate that the modified K-10 is reliable and valid screening tool for assessing distress in the current older adult sample. Rasch analysis resolved issues of local dependency, misfits, and disordered thresholds which contributed to the inadequacy of the scale prior modifications. Transforming ordinal-to-interval level scores using the transformation algorithm developed by this study enhances precision of the K-10 in older population. This allows the K-10 to be used as a screening instrument to accurately detect distress in older populations, which assists with application of preventive intervention to diminish risk of progression to more serious mental disorders and to improve quality of life among older adults.

References

Adams, R.J., Wu, M.L., Wilson, M., 2012. The Rasch rating model and the disordered threshold controversy. Educ. Psychol. Meas. 72, 547–573. https://doi.org/10.1177/0013164411432166

An, M., Yu, X., 2021. A Rasch analysis of emerging adults’ health motivation questionnaire in higher education context. PloS One 16, e0248389. https://doi.org/10.1371/journal.pone.0248389

Andersen, L.S., Grimsrud, A., Myer, L., Williams, D.R., Stein, D.J., Seedat, S., 2011. The psychometric properties of the K10 and K6 scales in screening for mood and anxiety disorders in the South African Stress and Health study. Int. J. Methods Psychiatr. Res. 20, 215–223. https://doi.org/10.1002/mpr.351

Anderson, T.M., Sunderland, M., Andrews, G., Titov, N., Dear, B.F., Sachdev, P.S., 2013. The 10-item Kessler psychological distress scale (K10) as a screening instrument in older individuals. Am. J. Geriatr. Psychiatry Off. J. Am. Assoc. Geriatr. Psychiatry 21, 596–606. https://doi.org/10.1016/j.jagp.2013.01.009

Andrich, D., 1978. Application of a Psychometric Rating Model to Ordered Categories Which Are Scored with Successive Integers [WWW Document]. URL https://journals.sagepub.com/doi/10.1177/014662167800200413 (accessed 4.15.22).

Arnaud, B., Malet, L., Teissedre, F., Izaute, M., Moustafa, F., Geneste, J., Schmidt, J., Llorca, P.-M., Brousse, G., 2010. Validity study of Kessler’s Psychological Distress scales conducted among patients admitted to French emergency department for alcohol consumption-related disorders: Validity study of Kessler’s Psychological Distress scales. Alcohol. Clin. Exp. Res. no-no. https://doi.org/10.1111/j.1530-0277.2010.01201.x

Atkins, J., Naismith, S.L., Luscombe, G.M., Hickie, I.B., 2013. Psychological distress and quality of life in older persons: relative contributions of fixed and modifiable risk factors. BMC Psychiatry 13, 249. https://doi.org/10.1186/1471-244X-13-249

Barber, C.C., Middlemiss, W., Medvedev, O.N., 2022. Applying Rasch methodology to examine and enhance precision of the Edinburgh Postnatal Depression Scale. J. Affect. Disord. 308, 391–397. https://doi.org/10.1016/j.jad.2022.04.009

Bond, T.G., Fox, C.M., Kathryn, C., Lacey, H., 2007. Applying the Rasch model: Fundamental measurement, in: In the Social Sciences (2nd.

Boone, W.J., 2016. Rasch analysis for instrument development: Why, when, and how? CBE Life Sci. Educ. 15, rm4. https://doi.org/10.1187/cbe.16-04-0148

Bougie, E., Arim, R.G., Kohen, D.E., Findlay, L.C., 2016. Validation of the 10-item Kessler Psychological Distress Scale (K10) in the 2012 Aboriginal Peoples Survey. Health Rep. 27, 3–10.

Brenlla, M., Aranguren, M., 2010. Adaptación argentina de la Escala de Malestar Psicológico de Kessler (K10). Rev. Psicol. 28, 308–340.

Chittleborough, C.R., Winefield, H., Gill, T.K., Koster, C., Taylor, A.W., 2011. Age differences in associations between psychological distress and chronic conditions. Int. J. Public Health 56, 71–80. https://doi.org/10.1007/s00038-010-0197-5

Christensen, K.B., Kreiner, S., Mesbah, M., 2012. Rasch models in health. John Wiley & Sons.

Christensen, K.B., Makransky, G., Horton, M., 2017. Critical values for Yen's Q3: Identification of local dependence in the Rasch model using residual correlations. Appl. Psychol. Meas. 41, 178–194. https://doi.org/10.1177/0146621616677520 de Morton, N.A., Keating, J.L., Davidson, M., 2008. Rasch analysis of the Barthel Index in the assessment of hospitalized older patients after admission for an acute medical condition. Arch. Phys. Med. Rehabil. 89, 641–647. https://doi.org/10.1016/j.apmr.2007.10.021

Donker, T., Comijs, H., Cuijpers, P., Terluin, B., Nolen, W., Zitman, F., Penninx, B., 2010. The validity of the Dutch K10 and extended K10 screening scales for depressive and anxiety disorders. Psychiatry Res. 176, 45–50. https://doi.org/10.1016/j.psychres.2009.01.012

Easton, S.D., Safadi, N.S., Wang, Y., Hasson, R.G., 2017. The Kessler psychological distress scale: translation and validation of an Arabic version. Health Qual. Life Outcomes 15, 215. https://doi.org/10.1186/s12955-017-0783-9

Embretson, S.E., 1996. The new rules of measurement. Psychol. Assess. 8, 341–349. https://doi.org/10.1037/1040-3590.8.4.341

Finaulahi, K.P., Sumich, A., Heym, N., Medvedev, O.N., 2021. Investigating psychometric properties of the Self-Compassion scale using Rasch methodology. Mindfulness 12, 730–740. https://doi.org/10.1007/s12671-020-01539-8

Hagell, P., Westergren, A., 2016. Sample size and statistical conclusions from tests of fit to the Rasch model according to the Rasch Unidimensional Measurement model (Rumm) program in health outcome measurement. J. Appl. Meas. 17, 416–431.

Hobart, J., Cano, S., 2009. Improving the evaluation of therapeutic interventions in multiple sclerosis: The role of new psychometric methods. Health Technol. Assess. 13, 1–200. https://doi.org/10.3310/hta13120

Hopkins, E.G., Lyndon, M.P., Henning, M.A., Medvedev, O.N., 2021. Applying Rasch analysis to evaluate and enhance the Academic Motivation Scale. Aust. J. Psychol. 73, 348–356. https://doi.org/10.1080/00049530.2021.1904794

Jong Won, M., Sun Hae, L., 2015. Validation of the K6/K10 Scales of psychological distress and their optimal cutoff scores for older Koreans. Int. J. Aging Hum. Dev. 80, 264–282. https://doi.org/10.1177/0091415015590316

Kessler, R., Mroczek, D., 1994. Final versions of our non-specific psychological distress scale. undefined.

Kessler, R.C., Andrews, G., Colpe, L.J., Hiripi, E., Mroczek, D.K., Normand, S.-L.T., Walters, E.E., Zaslavsky, A.M., 2002. Short screening scales to monitor population prevalences and trends in non-specific psychological distress. Psychol. Med. 32, 959–976. https://doi.org/10.1017/S0033291702006074

Kessler, R.C., Barker, P.R., Colpe, L.J., Epstein, J.F., Gfroerer, J.C., Hiripi, E., Howes, M.J., Normand, S.-L.T., Manderscheid, R.W., Walters, E.E., Zaslavsky, A.M., 2003. Screening for serious mental illness in the general population. Arch. Gen. Psychiatry 60, 184–189. https://doi.org/10.1001/archpsyc.60.2.184

Kreiner, S., 2007. Validity and objectivity: Reflections on the role and nature of Rasch models. Nord. Psychol. 59, 268–298. https://doi.org/10.1027/1901-2276.59.3.268

Lins, G.O. de A., Lima, N.A. da S., Sousa, G.S. de, Guimarães, F.J., Frazão, I. da S., Perrelli, J.G.A., 2021. Validity and reliability of Kessler Psychological Distress Scale for Brazilian elderly: A cross-sectional study. Rev. Bras. Enferm. 74. https://doi.org/10.1590/0034-7167-2020-0365

Marateb, H.R., Mansourian, M., Adibi, P., Farina, D., 2014. Manipulating measurement scales in medical statistical analysis and data mining: A review of methodologies. J. Res. Med. Sci. Off. J. Isfahan Univ. Med. Sci. 19, 47–56.

Masters, G.N., 1982. A Rasch model for partial credit scoring. Psychometrika 47, 149–174. https://doi.org/10.1007/BF02296272

Matthews, G., 2016. Chapter 26 - Distress, in: Fink, G. (Ed.), Stress: Concepts, Cognition, Emotion, and Behavior. Academic Press, San Diego, pp. 219–226. https://doi.org/10.1016/B978-0-12-800951-2.00026-1

Meade, A., Diamantopoulos, A., & Fernández-Feito, A. (2018). Best practices in conducting Rasch measurement analysis. In A. Meade & A. Diamantopoulos (Eds.), Handbook of partial least squares: Concepts, methods and applications (pp. 45-68). Springer.

McLachlan, K.J.J., Gale, C.R., 2018. The effects of psychological distress and its interaction with socioeconomic position on risk of developing four chronic diseases. J. Psychosom. Res. 109, 79–85. https://doi.org/10.1016/j.jpsychores.2018.04.004

Medvedev, O.N., Krägeloh, C.U., 2022. Rasch Measurement Model, in: Medvedev, O.N., Krägeloh, C.U., Siegert, R.J., Singh, N.N. (Eds.), Handbook of Assessment in Mindfulness Research. Springer International Publishing, Cham, pp. 1–18. https://doi.org/10.1007/978-3-030-77644-2_4-1

Medvedev, O.N., Siegert, R.J., Kersten, P., Krägeloh, C.U., 2017. Improving the precision of the Five Facet Mindfulness Questionnaire using a Rasch approach. Mindfulness 8, 995–1008. https://doi.org/10.1007/s12671-016-0676-8

Medvedev, O.N., Siegert, R.J., Kersten, P., Krägeloh, C.U., 2016. Rasch Analysis of the Kentucky Inventory of Mindfulness Skills. Mindfulness 7, 466–478. https://doi.org/10.1007/s12671-015-0475-7

Merkin, A.G., Medvedev, O.N., Sachdev, P.S., Tippett, L., Krishnamurthi, R., Mahon, S., Kasabov, N., Parmar, P., Crawford, J., Doborjeh, Z.G., Doborjeh, M.G., Kang, K., Kochan, N.A., Bahrami, H., Brodaty, H., Feigin, V.L., 2020. New avenue for the geriatric depression scale: Rasch transformation enhances reliability of assessment. J. Affect. Disord. 264, 7–14. https://doi.org/10.1016/j.jad.2019.11.100

Mircioiu, C., Atkinson, J., 2017. A comparison of parametric and non-parametric methods applied to a Likert scale. Pharm. Basel Switz. 5, E26. https://doi.org/10.3390/pharmacy5020026

Mitchell-Parker, K., Medvedev, O.N., Krägeloh, C.U., Siegert, R.J., 2018. Rasch analysis of the Frost Multidimensional Perfectionism Scale. Aust. J. Psychol. 70, 258–268. https://doi.org/10.1111/ajpy.12192

Ormel, J., Kempen, G.I.J.M., Penninx, B.W.J.H., Brilman, E.I., Beekman, A.T.F., Van Sonderen, E., 1997. Chronic medical conditions and mental health in older people: Disability and psychosocial resources mediate specific mental health effects. Psychol. Med. 27, 1065–1077. https://doi.org/10.1017/S0033291797005321

Peixoto, E.M., Zanini, D.S., de Andrade, J.M., 2021. Cross-cultural adaptation and psychometric properties of the Kessler Distress Scale (K10): An application of the rating scale model. Psicol. Reflex. E Crítica Rev. Semest. Dep. Psicol. UFRGS 34, 21. https://doi.org/10.1186/s41155-021-00186-9

Pereira, A., Oliveira, C.A., Bártolo, A., Monteiro, S., Vagos, P., Jardim, J., 2019. Reliability and factor structure of the 10-item Kessler Psychological Distress Scale (K10) among Portuguese adults. Ciênc. Amp Saúde Coletiva 24, 729.

Rasch, G., 1961. On general laws and the meaning of measurement in psychology. Proc. Fourth Berkeley Symp. Math. Stat. Probab. Vol. 4 Contrib. Biol. Probl. Med. 4.4, 321–334.

Rasch, G., 1960. Studies in mathematical psychology: I. Probabilistic models for some intelligence and attainment tests, Studies in mathematical psychology: I. Probabilistic models for some intelligence and attainment tests. Nielsen & Lydiche, Oxford, England.

Sachdev, P.S., Brodaty, H., Reppermund, S., Kochan, N.A., Trollor, J.N., Draper, B., Slavin, M.J., Crawford, J., Kang, K., Broe, G.A., Mather, K.A., Lux, O., the Memory and Ageing Study Team, 2010. The Sydney Memory and Ageing Study (MAS): methodology and baseline medical and neuropsychiatric characteristics of an elderly epidemiological non-demented cohort of Australians aged 70–90 years. Int. Psychogeriatr. 22, 1248–1264. https://doi.org/10.1017/S1041610210001067

Sakurai, K., Nishi, A., Kondo, K., Yanagida, K., Kawakami, N., 2011. Screening performance of K6/K10 and other screening instruments for mood and anxiety disorders in Japan: Screening performance of K6/K10. Psychiatry Clin. Neurosci. 65, 434–441. https://doi.org/10.1111/j.1440-1819.2011.02236.x

Sampasa-Kanyinga, H., Zamorski, M.A., Colman, I., 2018. The psychometric properties of the 10-item Kessler Psychological Distress Scale (K10) in Canadian military personnel. PLoS ONE 13, e0196562. https://doi.org/10.1371/journal.pone.0196562

Sierra-Siegert, M., David, A.S., 2007. Depersonalization and Individualism: The Effect of Culture on Symptom Profiles in Panic Disorder. J. Nerv. Ment. Dis. 195, 989–995. https://doi.org/10.1097/NMD.0b013e31815c19f7

Smith Jr., E.V., 2002. Understanding Rasch measurement: Detecting and evaluating the impact of multidimenstionality using item fit statistics and principal component analysis of residuals. J. Appl. Meas. 3, 205–231.

Tennant, A., Conaghan, P.G., 2007. The Rasch measurement model in rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch paper? Arthritis Rheum. 57, 1358–1362. https://doi.org/10.1002/art.23108

Tennant, A., Pallant, J.F., 2006. Unidimensionality matters! (a tale of two Smiths?). Rasch Meas. Trans. 20, 1048–1051.

Thurstone, L.L., 1931. Influence of motion pictures on children's attitudes. J. Soc. Psychol. 2, 291–305. https://doi.org/10.1080/00224545.1931.9918977

Truong, Q.C., Choo, C., Numbers, K., Merkin, A.G., Sachdev, P.S., Feigin, V.L., Brodaty, H., Kochan, N.A., Medvedev, O.N., 2021. Enhancing precision of the 16-item Informant questionnaire on cognitive decline in the elderly (IQCODE-16) using Rasch methodology. Int. Psychogeriatr. 1–11. https://doi.org/10.1017/S1041610221002568

Uddin, M.N., Islam, F.M.A., Al Mahmud, A., 2018. Psychometric evaluation of an interview-administered version of the Kessler 10-item questionnaire (K10) for measuring psychological distress in rural Bangladesh. BMJ Open 8, e022967. https://doi.org/10.1136/bmjopen-2018-022967

United Nations, Department of Economic and Social Affairs, Population Division, 2020. World population ageing 2019 highlights.

Vasiliadis, H.-M., Chudzinski, V., Gontijo-Guerra, S., Préville, M., 2015. Screening instruments for a population of older adults: The 10-item Kessler Psychological Distress Scale (K10) and the 7-item Generalized Anxiety Disorder Scale (GAD-7). Psychiatry Res. 228, 89–94. https://doi.org/10.1016/j.psychres.2015.04.019

World Health Organisation, 2011. Global health and aging.

Wright, B.D., 1996. Local dependency, correlations and principal components. Rasch Meas. Trans. 10, 509–511.

Wright, B.D., Stone, M.H., 1979. Best test design. MESA Press.

Wright, B. D., & Masters, G. N. (1982). Rating scale analysis: Rasch measurement. Sage Publications.

Zhu, G., Wang, Z., Guan, P., Zhu, Y., Li, J., Liu, Y., Shi, F., Wang, S., 2021. A Rasch Analysis of the 10-item Kessler Psychological Distress Scale (K10) in the Urban-rural fFinge of China (preprint). In Review. https://doi.org/10.21203/rs.3.rs-441385/v1

Table 1

Summary of fit statistics for the initial and the final Rasch analyses of the K-10.

	Item fit residual		Person fit residual		Goodness of fit		*PSI*	Significant *t*-tests (Unidimensionality)	
Analyses	Mean / SD		Mean / SD		$\chi2$ (df)	*p*			
A1 Initial	-1.53	1.04	-0.46	0.85	77.81(40)	<0.001	0.61	2.0	(YES)
A2 Rescoring	-1.23	1.30	-0.42	0.79	77.18(40)	<0.001	0.62	1.0	(YES)
A3 Final	-1.62	1.53	-0.46	0.70	29.87(35)	0.71	0.72	2.5	(YES)

Table 2

Overall Rasch model fit statistics for K-10 items

No.	Item Content	Location	FitResid	ChiSq
1	Tired for no reason	-1.59	0.55	8.85
2	Nervous	-0.50	-1.25	7.71
3	So nervous nothing could calm you	0.74	-2.52*	9.72
4	Hopeless	0.70	-1.90	6.51
5	Restless or fidgety	-0.06	-0.01	4.98
6	So restless you could not sit still	1.34	-2.19	6.92
7	Depressed	-0.35	-2.12	11.03
8	Everything was an effort	-0.94	-0.25	8.13
9	So sad nothing could cheer you up	0.43	-2.56	7.89
10	Worthless	0.23	-1.94	6.08

Note. Items in the table are paraphrased from the original K-10 scale. Significant values are notified with * ($p > 0.05$)

Table 3

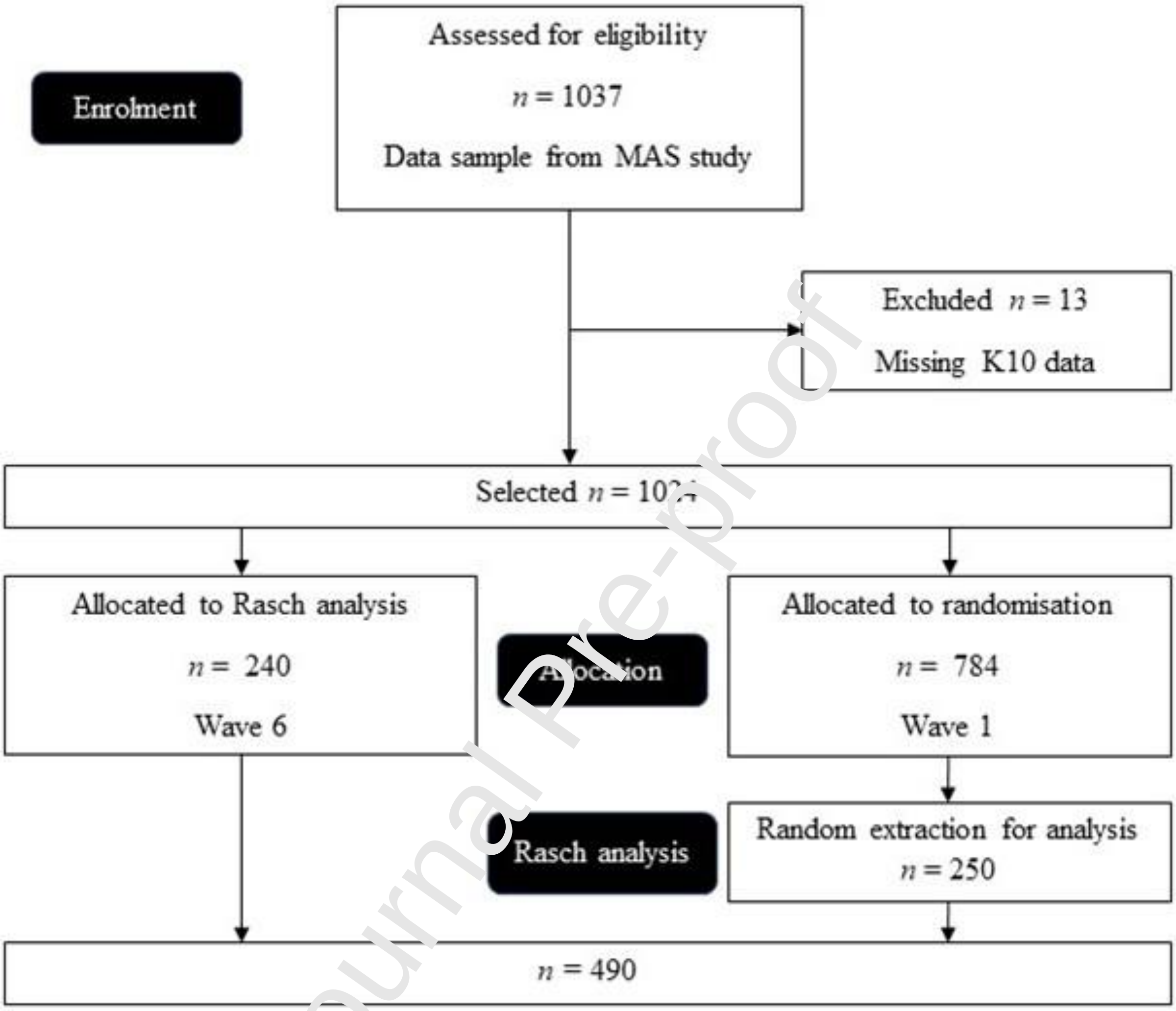
Ordinal-to-interval conversion table of the K-10 scores

Ordinal	Interval Scores	
Scale	Logits	Scale
10	-6.22	10.00
11	-5.10	13.28
12	-4.20	15.91
13	-3.48	18.03
14	-2.88	19.78
15	-2.39	21.22
16	-1.99	22.40
17	-1.65	23.40
18	-1.36	24.25
19	-1.11	24.99
20	-0.88	25.65
21	-0.68	26.23
22	-0.50	26.77
23	-0.33	27.26
24	-0.18	27.72
25	-0.03	28.16
26	0.12	28.58
27	0.26	28.99
28	0.40	29.41
29	0.54	29.82
30	0.69	30.25
31	0.84	30.70
32	1.01	31.18
33	1.18	31.70
34	1.38	32.27
35	1.60	32.92
36	1.86	33.67
37	2.17	34.59
38	2.57	35.78
39	3.16	37.50
40	4.02	40.00

Note. To transform ordinal K-10 scores first recode ordinal category 5 as 4 for all ten items; then compute total scores ranging from 10 to 40 and find corresponding interval-level scores on the right-hand side of the table. Interval transformation scores are presented in logit units in the middle column of the table. This table cannot be used for participants with missing item data.

Figure 1

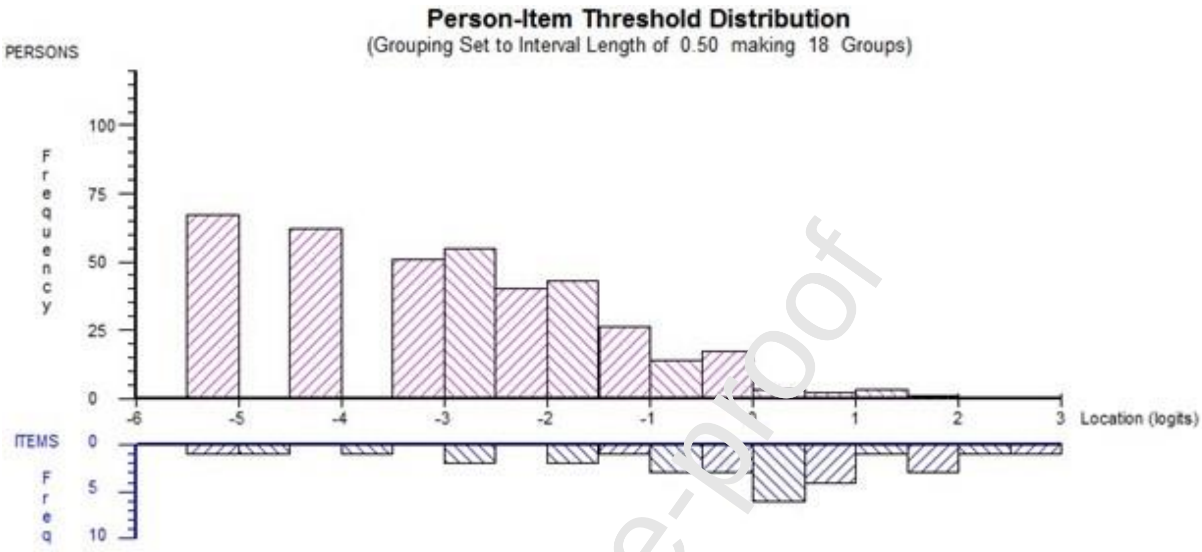
Flow Chart of the Study Population Extraction Method for Rasch Analysis of the K-10



Note. Participant K-10 scores ≥ 10 were taken from MAS Wave 6 and scores < 10 were taken from MAS Wave 1

Figure 2

Person-item threshold distribution of best model fit analysis (A3 Final) of the K-10 (Person mean in logits =3.71; SD=1.86)



Highlights

Psychological distress is a risk factor for cognitive decline and dementia in older people

Accurate measurement of distress is important for diagnosis and assessment of risks

Rasch analysis is a powerful methodology to examine and improve measurement

Rasch analysis was applied to the Kessler Psychological Distress Scale (K-10)

This study examined and improved accuracy of K-10 assessment scores

Ordinal-to-interval conversion tables were developed to enhance precision of K-10