

A Rasch Analysis of the 10-item Kessler Psychological Distress Scale (K10) in the Urban-rural fFinge of China

Gaopei Zhu

Weifang Medical University <https://orcid.org/0000-0002-2085-9336>

Zhongli Wang

Shandong University Cheeloo College of Medicine

Peixia Guan

Weifang Medical University

Yuhang Zhu

Weifang Medical University

Juan Li

Weifang Medical University

Yujie Liu

Weifang Medical University

Fuyan Shi

Weifang Medical University

Suzhen Wang (✉ wangsz@wfmc.edu.cn)

Weifang Medical University <https://orcid.org/0000-0003-2076-529X>

Research

Keywords: urban-rural fringe, K10, psychometry characteristics, Rasch analysis, Differential item functioning

Posted Date: May 4th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-441385/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background

The impact of urbanization on the physical and mental health of the rural-urban population has been ignored. The primary objective of this investigation was to demonstrate the reliability, validity, and responsiveness of the 10-item Kessler Psychological Distress Scale (K10) in the measurement of psychological distress in the rural-urban fringe population.

Methods

Data were obtained from the mental health section of chronic disease survey in Longzihu District, urban-rural fringe area, of Bengbu City, with 3354 participants. The Mandarin version of K10 was used for face-to-face interviews. The Rasch model was used to analyze the psychometry characteristics and differential item functioning (DIF) of K10.

Results

Rasch analysis results revealed that the K10 scale showed ordered response categories. The results of principal component analysis (PCA) and information-weighted fit statistic (infit) mean square (MNSQ) indicated that the K10 scale conforms to unidimensionality. The Cronbach's alpha coefficient of K10 was 0.916 (95% confidence interval (CI): 0.907,0.924), which had good reliability, but the Cronbach's alpha coefficient would be increased if the fifth item was removed. The results of the Rasch model showed that all the 10 items in the K10 scale had a good fitting effect (Infit MNSQ value, 0.928-1.072). A non-significant differential item functioning (DIF) was found on K10 of age and gender. Overall, the K10 scale was more difficult, and the psychological distress score of the subjects was generally low.

Conclusion

Rasch analysis showed that the Mandarin version of K10 was an effective and reliable scale for measuring and screening mental distress of residents in the urban-rural fringe. However, it was still recommended that further research should be conducted to solve inappropriate and difficult items.

Introduction

Mental health and well-being have become a key issue in modern public health. Around one billion people worldwide meet the criteria for common mental disorders [1], and about one-third of them have disabilities [2]. Mental disorders are the second largest contributor to the global burden of disease, accounting for 7–13% of disability-adjusted life expectancy [2]. Mental illness has evolved into one of the main drivers of the global burden of disease [3]. It is very necessary and important to monitor, screen and manage the mental health symptoms of residents.

With the continuous expansion of global cities, the urbanization rate has been rising, and developing countries have contributed 90% of the urbanization process [4]. From 1992 to 2015, China experienced rapid and large-scale urban expansion. Urban land use has increased nearly five times, with an average annual growth rate of 8.10% [5]. Previous study has shown that urban expansion is squeezing the activity space of residents in the urban-rural fringe, causing air, noise, and water pollution, reducing green space per capita and increasing social exclusion [6]. Recent study has shown that air pollution is associated with deteriorating mental health. In addition, noise and water pollution increase the risk of psychological stress and disease [7]. Many studies have linked green spaces to mental health [8, 9]. A cluster randomized trial found a 62.8% reduction in mental distress among people living near green Spaces [10]. The urban-rural fringe, as a direct contact area in the process of urbanization, has brought residents more opportunities to

exposed the harm caused by urbanization, and the urban-rural floating population has experienced social exclusion, which may have a significant impact on mental health [6]. When we discuss health and well-being within the framework of "health in all policy" [11] and sustainable development [12], the psychological distress of residents in the urban-rural fringe should be an important social and health issue worthy of attention.

At present, there are many tools for measuring mental health, among which 10-item Kessler Psychological Distress Scale (K10) is widely favored by researchers because of its good screening characteristics and ease of use [13–17]. Kessler mental distress scale has been validated in Canada [18], the United States [19], Australia [13], etc. In addition, K10 has been tested in the military personnel [14], diabetes [20], adolescents [21], the elderly [22]. The tool has also been validated in different language versions, including Mandarin [23, 24], Arabic [15], Korean [25], Danish [26], etc. Although Kessler mental distress scale has been verified in different populations and languages, our literature review has not found an empirical study on the measurement attribute of mental distress (K10) of residents in the urban-rural fringe. It is surprising that research on the psychological distress of rural-urban populations has been ignored.

In conclusion, the purpose of this study is to evaluate the psychometric characteristics of the K10 Mandarin version [27] based on the epidemiological survey data in the urban-rural fringe, and to evaluate the different item functioning of K10 in different characteristics of urban-rural integration in China.

Method

Study design and participants

The data were collected from the survey on health status of residents in Longzihu District, Bengbu City, China in 2015. In this study, multi-stage stratified random sampling was used to select the residents of 7 community service centers, a total of 3591 people were investigated and 3354 samples were effectively recovered. A total of 56 community health educators were trained to conduct health check-up and questionnaire surveys.

Measures and instrument

The indicators were obtained by questionnaire survey, including basic information of residents, two weeks of illness, mental health, and health-related literacy and so on.

The 10-item Kessler Psychological Distress Scale (K10) was the non-specific psychological distress scale, which consisted of 10 items about anxiety and depression symptoms experienced by the respondents in the past 4 weeks [28]. K10 was developed for the redesigned national health interview survey (NHIS). Kessler's research showed that screening for severe mental disorders is very effective. The measure had five response categories (P1 = "none of the time"; P2 = "a little of the time"; P3 = "some of the time"; P4 = "most of the time"; P5 = "all of the time"), ranging from 0 (none of the time) to 4 (all of the time). The total scores for these items ranged from 0 to 40, with higher scores indicating higher levels of psychological distress [14]. The Mandarin version of K10 was used for the survey of mental health, which was measured by asking subjects how they felt about items of psychological distress over the past four weeks [23].

Factor variables of differential item functioning

Age and gender are important factors variables. As a continuous variable, age was transformed into categorical variable. Participants were divided into two groups, including younger group (aged 18-59 years) and older group (aged

60-90 years old).

Statistical Analysis

Rasch model was a parametric latent trait model based on Item Response Theory (IRT). Rasch analysis, developed by Danish mathematician Georg Rasch, was a formal test of the results of a scale using mathematical models [29]. Rasch analysis provided an estimate of the level of a person's potential variables and allowed for subsequent statistical analysis. In addition, the model was a robust model for measuring potential traits and had been proved to solve some of the weaknesses of CTT [30, 31]. Rasch analysis was widely used to evaluate the psychometric attributes of questionnaires and scales. It could measure human ability and project difficulty simultaneously on the same scale [29].

In the current research, Rasch analysis was used to test the quality of the items. To validate the Chinese version of the K10 in the urban-rural fringe, six key indicators were used, which included the following: (1) Information-weighted mean-square statistics (Infit MNSQ). The infit MNSQ values between 0.7 and 1.3 were considered unidimensionality [32]. Infit MNSQ value greater than 1.3 indicated that the variance of measurement was too large, and infit MNSQ value less than 0.7 indicated that the variance of measurement was too small [33]. (2) Principal component analysis (PCA). The PCA of the residuals was performed to test the unidimensionality of the K10 [34]. Two criteria were used in PCA. The first was that the variance explained by the first principal component should be sufficient (> 50%). The second was that if the ratio of the first factor eigenvalue to the second factor eigenvalue was close to or greater than 3.0, then the scale was unidimensionality [32, 35]. (3) Category threshold order. The order of category threshold reflected by the category probability curve was an important parameter to demonstrate the usage of response categories, and it was the basis to distinguish personnel ability and project difficulty. When it was difficult to distinguish the orderly response options, the disorder threshold appeared [36]. (4) Differential item functioning (DIF). DIF is an indispensable part of psychometric analyses aiming to measure invariance across sample groups, e.g., for male and female [37]. (5) Cronbach's alpha coefficient. Cronbach's alpha coefficient was a reliability index, which indicated how reliable the consistency of measurement was. The value range was 0-1. If Cronbach's alpha coefficient was greater than 0.9, it meant good consistency [38]. Person-item map. Rasch model users called a "Wright Map". In Rasch model, the difficulty parameter of project was on the same scale as human ability θ , so the difficulty of project could be explained by the same scale as human ability. The person item map was a useful tool for understanding how projects and people are distributed along θ [39]. Ideally, the difficulty of the project would be distributed within the range of meaningful potential characteristics or capabilities, so that the project could measure capability across the entire range.

All statistical analysis and graphic plotting were performed using R version 4.0.3 R software (The R Foundation for Statistical Computing, Vienna, Austria). And the software packages "mirt" and "ltm" were used to build the Rasch model. The "lordif" function was used to analysis the DIF.

Results

Characteristics of the patients

A total of 3354 patients were included in the study. Table 1 showed the distribution characteristics of scores of K10. The K10 had an average score of 3.34 (standard deviation was 5.28), with a median score of 1.00 and the scale's score range was 0-40. The study comprised 1885 females (56.2%) and 1469 males (43.8%), with a mean age of 56.0

years (range 18-91 years). The Mann Whitney U test of K10 scale showed that there were significant differences in the total scores of gender and age. The average score of female psychological distress was higher than that of male, and the average score of psychological distress of the elderly was higher than that of young people. In addition, the test of single item showed that the average scores of other items except item 10 were higher in women than those in men. The average scores of K10 of the elderly in items 5 and 6 were higher than that of young people, and there was no significant difference in the average score between the elderly and young in the remaining items.

Unidimensionality

Principal component analysis showed that the residual explained 60.22% (greater than 50%) of the original variance. Moreover, the ratio of the characteristic root of the first factor to that of the second factor was 7.642 (the ratio was far greater than 3). The Infit MNSQ value for each item (see Table 2) was acceptable (0.928-1.072). Both results indicated that there was no evidence that K10 was multidimensional.

The goodness of fit of the Rasch model

In the Rasch model, infit MNSQ values were usually used to evaluate the goodness of fit of the items. As Wright and Linacre's (1994) suggestion, when the infit MNSQ value of an item was greater than 1.5 or less than 0.5, it could be considered that the fitting effect of the item was poor. Rasch model analysis showed that 10 items in K10 scale had a good fitting effect (infit MNSQ values, 0.928-1.072), as shown in Table 2.

Threshold order and Category probability curves

There was no evidence of disorder threshold in the category probability curve, and the category threshold increased in an orderly manner (as shown in Figure 1 and table 2). All projects had five response categories, and it was recommended that each item had four thresholds. The items 1, 3, 5, and 10 in K10 were slightly lack of discrimination, which showed that P2 (a little of the time) was indistinguishable from P1 (none of the time) and P3 (some of the time), and P2 was poorly differentiated. Since the performance of these items were not good, the subsequent screening indicators could be deleted.

Reliability

The Cronbach' alpha coefficient of K10 was 0.916 (0.907, 0.924). Removing the item 5 would increase Cronbach' alpha value to 0.919 (0.911, 0.927), while deleting any other items would decrease Cronbach' alpha value (see Table 2).

Differential item functioning (DIF)

The items of K10 were assessed for DIF across gender and age (table 3). We used likelihood ratio (LR) Chi-square value to test DIF, and set the detection level α to 0.01. Significant DIF was found on item 3 when assessing age. However, DIF between ages was not significant when α was adjusted by Bonferroni ($0.01/10 = 0.001$). There was no DIF difference between genders.

The block diagram on the left of figure 2a showed the dispersion of K10 scores between ages. The shaded boxes in the block diagram were inter-quartile range, representing a 50% difference in the middle (the boundary is at the bottom and top of the shadow box), ranging from - 0.002 logit to + 0.003 logit, with a median value of about 0.001 logit. The left of figure 2b showed a box plot of the differences of gender. The interquartile range, representing the middle 50% of the differences (bound between the bottom and top of the shaded box), ranged from - 0.002 to 0.003, with a median value of about 0.001.

In the graph on the right the same difference scores were plotted against the initial scores ignoring DIF ("initial theta"), separately for younger and older individuals. Guidelines were placed at 0.0 (solid line), i.e., no difference, and the mean of the differences (dotted line). The positive value on the left side of the graph showed that for older, if DIF was considered, the score would decrease (i.e., ignoring the true score of DIF minus the score of DIF greater than 0, so the calculated DIF score was less than the true score), which was the opposite for younger. Figure 2b showed that the median 50% difference in gender DIF ranges from - 0.010 to 0.008. The solid line (standard line) on the left coincided with the dashed line (the average of the differences) indicating that DIF did not differ between genders. The influence of DIF on K10 score of male and female could be ignored.

Person-item map

In the person-item map, the histogram of human ability estimation $\hat{\theta}$ was shown at the top. The black dots on the lines at the bottom of Figure 3 were the threshold of item difficulty. In the K10 scale, the individual's ability $\hat{\theta}$ was between - 3.00 logit and - 2.00 logit, and individuals were more inclined to choose items with low scores of "none of the time" and "a little of the time". In other words, participants were more likely to choose items with difficulty θ higher than 1.00 logit. Figure 3 showed that the difficulty of items 1, 3, and 5 were around 0.00 logit, while the difficulty of other items was generally around 1.00 logit, and the overall difficulty of K10 was relatively high.

Discussion

The purpose of this study was to evaluate the applicability of K10 in measuring the psychological stress of residents in the urban-rural fringe of China. The Mandarin version of K10 showed high internal consistency and reliability. The results of PCA and Rasch models supported the unidimensional structure of K10, which was consistent with previous study [28]. Rasch analysis showed that the response categories of all items were ranked in order of severity. Although the DIF of the item3 (Did you feel so nervous that nothing could calm you down?) was statistically different between ages, there was no significant DIF in gender and age after adjustment. Our research supported that the K10 distress scale was an effective measurement tool [13], which could be used to quantify the non-specific psychological distress of residents in the urban-rural fringe of China.

Rasch analysis was a powerful tool that allowed the identification of items that were not sensitive to potential traits. The analysis of K10 showed that although not all items were in full compliance with the Rasch model, all items were effective for measurement, that was, the infit MNSQ values were less than 1.3 [32]. These results showed that K10 was a good measure of a potential structure. Principal component analysis of residuals further supported the unidimensional of K10 [34]. Although some studies have found that K10 was two-dimensional and had different two-factor structure [24, 40, 41]. However, this study supported the unidimensional hypothesis of K10 [28]. The Cronbach' alpha coefficient of K10 in this study was very satisfactory, which was similar to the reliability of Mandarin and other versions [23, 24].

There were differences in gender and age of psychological distress. It was reported that female respondents showed significantly higher anxiety than males [18, 42, 43]. In line with that, we found that with the exception of the "feel worthless" item, the average scores in female of other items were higher than that of male. Moreover, our study revealed that the level of psychological distress of the elderly was higher than that of the young, in the items of "feel restless or fitness" and "feel so restless that you could not sit still", there was no difference between the older and the younger in the remaining items. However, Evelynne bougie [18] found that people aged 55 or older suffered significantly less than younger people, which was contrary to our conclusion.

Our study demonstrated the ordered threshold in the K10 category probability curve, which meant that patients with more mental distress for an item had more anxiety than those who claimed to have less. In other words, the K10 scale could accurately distinguish subjects with low and high levels of mental distress [36]. Our results were consistent with previous studies [13-17], which suggesting that K10 was a good measure of patients with mental distress. However, the discrimination ability of items 1, 3, 5, and 10 in K10 scale was poor. Specifically, the distinction between "none of the time", "a little of the time" and "some of the time" is not good. This is consistent with was a rural study in Bangladesh [44], which found that there was a disorder threshold in the category probability curve and suggested merge the disordered parts of the sorting.

K10 scale was used to measure the psychological distress of residents in urban-rural fringe, and the DIF of age and gender was not shown. This was consistent with the results of K10 validation in rural areas of Bangladesh [44]. K10 had consistent DIF evidence in measuring different genders and ages [37].

Person-item map represented the location of item difficulty and the distribution of person along the logit scale [45]. In the K10 scale, the ability of individuals was between - 3.00 logit and - 2.00 logit, and individuals were more inclined to choose items with low scores of "none of the time" and "a little of the time". In other words, the K10 scale was more difficult, and it was not easy to distinguish the subjects with low levels of psychological anxiety [39]. A similar phenomenon was also found in rural Bangladesh [44]. This difference might be partly due to the psychologically distressed patients in the evaluation group were more satisfied with the current psychological anxiety status, who were generally lower than the expected level of psychological anxiety level [44]. In the future research, data on the floating population in the urban-rural fringe should be considered, which can better verify these measurement attributes of the K10 scale.

Some limitations should be considered in our findings: Firstly, all patients were recruited from one district, and there might be selection bias, which needed to be verified in random samples from other urban-rural fringe areas, so that the results could truly represent the psychological distress of urban-rural fringe areas in China. Secondly, the subjects of this study were adults, and there is a lack of measurement data for children and adolescents in the urban-rural fringe, which had an impact on the extrapolation of the model. In addition, although our study analyzed the DIF between different characteristics, we only validated the DIF of age and gender, and more DIF between features needed to be explored later. Finally, due to the large population mobility in the urban-rural fringe, and remedial measures had been taken during the study period, this might affect the universality of the results.

Conclusions

In conclusion, the Mandarin version of K10 had high reliability and structural validity in the measurement and identification of anxiety patients in urban-rural fringe. The K10 scale satisfied most of the hypotheses of Rasch model, which did not show DIF for age and gender. According to the ordered threshold of category probability curve, the K10 scale could accurately distinguish the subjects with low level and high level of psychological distress. This study

conducted a more comprehensive analysis of the measurement attributes of the Mandarin version of K10, but further research is still recommended to solve the problems of inappropriate items and high difficulty.

Abbreviations

K10: The 10-item Kessler Psychological Distress Scale; PCA: principal component analysis; DIF: Differential item functioning; mean: The mean of Item score; SD: Standard deviation; Infit: Information-weighted fit statistic; MNSQ: Mean square; CTT: Classical Test Theory; CI: Confidence interval.

Declarations

Acknowledgments

We would like to acknowledge and thank the study participants for being involved in this project. Special thanks to Professor Xuesen Wu of Bengbu Medical University for providing the data and details of the design and implementation of the study in Longzihu District, Bengbu City.

Funding

This study was partially supported by the National Natural Science Foundation of China (Project approval No.: 81872719, 81803337), the National Bureau of Statistics Foundation Project (Project approval No.: 2018LY79), the Natural Science Foundation of Shandong Province (Project approval No.: ZR2019MH034). The funders did not play any role in the study design, data collection and interpretation of data, or in writing the manuscript.

Availability of data and materials

All authors had full access to the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Authors' contributions

G. Z., contributed to conceptualization, methodology, data curation, software, and original manuscript writing; Z. W., contributed to the review and editing of writing; P. G. and Y. Z., contributed to data curation and the review and editing of writing; J. L., contributed to supervision, software, and validation; Y. L., contributed to supervision and formal analysis; S. W. and F. S., contributed to methodology and the review and editing of writing. All authors gave final approval and agreed to be accountable for all aspects of the work.

Ethics approval and consent to participate

The Ethics Committee of Weifang Medical University (WFMU), Weifang, China approved the study protocol. All participants were informed research purpose and gave their written informed consent prior to the commencement of the study.

Consent for publication

Not applicable.

Competing interests

The author declares that he has no competing interests.

References

1. Steel Z, Marnane C, Iranpour C, et al. The global prevalence of common mental disorders: a systematic review and meta-analysis 1980-2013. *Int J Epidemiol*. 2014; 43(2):476-93.
2. Vigo D, Thornicroft G, Atun R. Estimating the true global burden of mental illness. *Lancet Psychiat*. 2016; 3(2):171-8.
3. Murray CJL, Barber RM, Foreman KJ, et al. Global, regional, and national disability-adjusted life years (DALYs) for 306 diseases and injuries and healthy life expectancy (HALE) for 188 countries, 1990–2013: quantifying the epidemiological transition. *Lancet*. 2015; 386(10009):2145-91.
4. UN DESA, Population Division. World urbanization prospects: the 2018 Revision, Online Edition. 2018. <https://www.un.org/development/desa/pd/content/world-urbanization-prospects-2018-revision>. Accessed 18 April 2021.
5. Xu M, He C, Liu Z, et al. How Did Urban Land Expand in China between 1992 and 2015? A Multi-Scale Landscape Analysis. *PloS one*. 2016; 11(5):e0154839.
6. Li J, Rose N. Urban social exclusion and mental health of China's rural-urban migrants - A review and call for research. *Health Place*. 2017; 48:20-30.
7. Klompaker JO, Hoek G, Bloemsma LD, et al. Associations of combined exposures to surrounding green, air pollution and traffic noise on mental health. *Environ int*. 2019; 129:525-37.
8. Feng X, Astell-Burt T. Residential green space quantity and quality and symptoms of psychological distress: a 15-year longitudinal study of 3897 women in postpartum. *BMC Psychiatry*. 2018; 18(1):348.
9. Dzhambov AM, Markevych I, Hartig T, et al. Multiple pathways link urban green- and bluespace to mental health in young adults. *Environ res*. 2018; 166:223-33.
10. South EC, Hohl BC, Kondo MC, et al. Effect of Greening Vacant Land on Mental Health of Community-Dwelling Adults: A Cluster Randomized Trial. *JAMA Netw Open*. 2018; 1(3):e180298.
11. WHO. Health in All Policies: Framework for Country Action. 2013. <https://www.who.int/healthpromotion/frameworkforcountryaction/en/>. Accessed 18 April 2021.
12. WHO. Sustainable Development Goals (SDGs). 2000. <https://www.who.int/health-topics/sustainable-development-goals>. Accessed 18 April 2021.
13. Furukawa TA, Kessler RC, Slade T, et al. The performance of the K6 and K10 screening scales for psychological distress in the Australian National Survey of Mental Health and Well-Being. *Psychol Med*. 2003; 33(2):357–62.
14. Sampasa-Kanyinga H, Zamorski MA, Colman I. The psychometric properties of the 10-item Kessler Psychological Distress Scale (K10) in Canadian military personnel. *PloS one*. 2018; 13(4):e0196562.
15. Easton SD, Safadi NS, Wang Y, et al. The Kessler psychological distress scale: translation and validation of an Arabic version. *Health Qual Life Out*. 2017; 15(1):215.

16. Tran TD, Kaligis F, Wiguna T, et al. Screening for depressive and anxiety disorders among adolescents in Indonesia: Formal validation of the centre for epidemiologic studies depression scale - revised and the Kessler psychological distress scale. *J Affect Disorders*. 2019; 246:189-94.
17. Fassaert T, De Wit MA, Tuinebreijer WC, et al. Psychometric properties of an interviewer-administered version of the Kessler Psychological Distress scale (K10) among Dutch, Moroccan and Turkish respondents. *Int J Meth Psych Res*. 2009; 18(3):159-68.
18. Bougie E, Arim RG, Kohen DE, et al. Validation of the 10-item Kessler psychological distress scale (K10) in the 2012 Aboriginal Peoples Survey. *Health Rep*. 2016; 27(1):3-10.
19. Kessler RC, Barker PR, Colpe LJ, et al. Screening for serious mental illness in the general population. *Arch Gen Psychiat*. 2003; 60(2):184-9.
20. Tan LS, Khoo EY, Tan CS, et al. Sensitivity of three widely used questionnaires for measuring psychological distress among patients with type 2 diabetes mellitus. *Qual Life Res*. 2015; 24(1):153-62.
21. Ferro MA. The Psychometric Properties of the Kessler Psychological Distress Scale (K6) in an Epidemiological Sample of Canadian Youth. *Can J Psychiat*. 2019; 64(9):647-57.
22. Anderson TM, Sunderland M, Andrews G, et al. The 10-item Kessler psychological distress scale (K10) as a screening instrument in older individuals. *Am J Geriatr Psychiat*. 2013; 21(7):596-606.
23. Zhou C, Chu J, Wang T, et al. Reliability and validity of 10-item Kessler Scale (K10) Chinese version in evaluation of mental health status of Chinese population. *Chin J Clin Psychol*. 2008; 16:627–9.
24. Bu XQ, You LM, Li Y, et al. Psychometric Properties of the Kessler 10 Scale in Chinese Parents of Children With Cancer. *Cancer Nurs*. 2017; 40(4):297-304.
25. Min JW, Lee SH. Validation of the K6/K10 Scales of Psychological Distress and Their Optimal Cutoff Scores for Older Koreans. *Int J Aging Hum Dev*. 2015; 80(3):264-282.
26. Thelin C, Mikkelsen B, Laier G, et al. Danish translation and validation of Kessler's 10-item psychological distress scale - K10. *Nord J Psychiat*. 2017; 71(6):411-416.
27. Xu L, Wang J, Sun H, et al. Researching the application of Kessler 10 for the first time and it' s important significance. *Soft Sci Health*. 2005; 6:410–412, 421.
28. KESSLER RC, ANDREWS G, COLPE LJ, et al. Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychol Med*. 2002; 32(2): 959-976.
29. Bang. CK, Kreiner S, Svend., Mesbah M. *Rasch Models in Health*. Wiley. 2013.
30. Tennant A, Penta M, Tesio L, et al. Assessing and adjusting for cross-cultural validity of impairment and activity limitation scales through differential item functioning within the framework of the Rasch model: the PRO-ESOR project. *Med Care*. 2004; 42(1 Suppl):137-48.
31. Tennant A, Conaghan PG. The Rasch measurement model in rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis Rheum-US*. 2007; 57(8):1358-62.
32. Lin X, Li M, Wang M, et al. Validation of Catquest-9SF questionnaire in a Chinese cataract population. *PloS one*. 2014; 9(8):e103860.
33. Visser MS, Dieleman M, Klijn S, et al. Validation, test-retest reliability and norm scores for the Dutch Catquest-9SF. *Acta Ophthalmol*. 2017; 95(3):312-9.
34. Gerbing DW, Anderson JC. An Updated Paradigm for Scale Development Incorporating Unidimensionality and Its Assessment. *J Marketing Res*. 1988; 25(2):186-92.
35. Hambleton RK, Swaminathan H. *Item response theory: Principles and applications*. Boston, MA: Kluwer-Nijhoff. 1985.

36. Robinson M, Johnson AM, Walton DM, et al. A comparison of the polytomous Rasch analysis output of RUMM2030 and R (Irtm/eRm/TAM/lordif). *BMC Med Res Methodol.* 2019; 19(1):36.
37. Hagquist C. Explaining differential item functioning focusing on the crucial role of external information - an example from the measurement of adolescent mental health. *BMC Med Res Methodol.* 2019; 19(1):185.
38. Xu Z, Wu S, Li W, et al. The Chinese Catquest-9SF: validation and application in community screenings. *BMC Ophthalmol.* 2018; 18(1):77.
39. Stelmack J, Szlyk JP, Stelmack T, et al. Use of Rasch person-item map in exploratory data analysis: a clinical perspective. *J Rehabil Res Dev.* 2004; 41(2):233-41.
40. Sunderland M, Mahoney A, Andrews G. Investigating the Factor Structure of the Kessler Psychological Distress Scale in Community and Clinical Samples of the Australian Population. *J Psychopathol Behav.* 2012; 34(2):253-9.
41. O'Connor SS, Beebe TJ, Lineberry TW, et al. The association between the Kessler 10 and suicidality: a cross-sectional analysis. *Compr Psychiat.* 2012; 53(1):48-53.
42. Oakley Browne MA, Wells JE, Scott KM, et al. The Kessler Psychological Distress Scale in Te Rau Hinengaro: the New Zealand Mental Health Survey. *Aust Nz J Psychiat.* 2010; 44(4):314-22.
43. Slade T, Grove R, Burgess P. Kessler Psychological Distress Scale: normative data from the 2007 Australian National Survey of Mental Health and Wellbeing. *Aust Nz J Psychiat.* 2011; 45(4):308-16.
44. Uddin MN, Islam FMA, Al Mahmud A. Psychometric evaluation of an interview-administered version of the Kessler 10-item questionnaire (K10) for measuring psychological distress in rural Bangladesh. *BMJ open.* 2018; 8(6):e022967.
45. Norhayati MN, Fatin Imtithal A, Nor Akma MJ. Psychometric properties of the Malay version of the Women's Views of Birth Labour Satisfaction Questionnaire using the Rasch measurement model: a cross sectional study. *BMC Pregnancy Childb.* 2020; 20(1):295.

Tables

Table 1

Score of K10 on age (younger and older) and gender (male and female)

Items	Score (mean ± SD)	Score by age			Score by gender		
		Older (mean ± SD)	Younger (mean ± SD)	P value	Male (mean ± SD)	Female (mean ± SD)	P value
Item1	0.59 ± 0.90	0.60 ± 0.92	0.56 ± 0.87	0.306	0.50 ± 0.84	0.65 ± 0.94	<0.001
Item2	0.29 ± 0.65	0.29 ± 0.65	0.29 ± 0.64	0.820	0.24 ± 0.59	0.34 ± 0.68	<0.001
Item3	0.23 ± 0.59	0.24 ± 0.60	0.22 ± 0.57	0.337	0.20 ± 0.54	0.26 ± 0.63	0.002
Item4	0.25 ± 0.61	0.25 ± 0.61	0.24 ± 0.61	0.225	0.20 ± 0.55	0.29 ± 0.65	<0.001
Item5	0.56 ± 0.92	0.59 ± 0.93	0.50 ± 0.88	0.003	0.46 ± 0.85	0.64 ± 0.95	<0.001
Item6	0.26 ± 0.63	0.29 ± 0.65	0.22 ± 0.57	0.004	0.22 ± 0.57	0.30 ± 0.67	<0.001
Item7	0.28 ± 0.64	0.29 ± 0.64	0.27 ± 0.64	0.079	0.23 ± 0.58	0.32 ± 0.69	<0.001
Item8	0.27 ± 0.65	0.28 ± 0.67	0.24 ± 0.60	0.131	0.24 ± 0.62	0.29 ± 0.67	0.026
Item9	0.25 ± 0.62	0.27 ± 0.64	0.23 ± 0.57	0.459	0.22 ± 0.58	0.28 ± 0.65	0.009
Item10	0.35 ± 0.71	0.35 ± 0.72	0.33 ± 0.69	0.312	0.32 ± 0.69	0.36 ± 0.72	0.113
K10	3.34 ± 5.28	3.46 ± 5.39	3.10 ± 5.05	0.023	2.82 ± 4.86	3.74 ± 5.55	<0.001

Note: *mean* the mean of Item score; *SD* Standard deviation; *Older* 60-90 years old; *Younger* 18-59 years old; *K10* the 10-item Kessler Psychological Distress Scale.

Table 2

Individuals' item fit statistics of K10

Items	Location	Threshold 1	Threshold 2	Threshold 3	Threshold 4	Infit ^a	Infit.z	Cronbach's alpha after removing the item(95%CI)
Item1	0.90130	-0.97961	-0.42755	1.19561	3.81674	1.037	1.181	0.915(0.906,0.924)
Item2	2.01364	0.13273	0.68479	2.30795	4.92908	1.018	0.446	0.905(0.894,0.914)
Item3	2.32681	0.44590	0.99796	2.62112	5.24225	1.013	0.310	0.905(0.895,0.913)
Item4	2.24282	0.36192	0.91397	2.53714	5.15826	0.928	-1.615	0.903(0.892,0.913)
Item5	0.98525	-0.89566	-0.34360	1.27956	3.90069	1.025	0.821	0.919(0.911,0.927)
Item6	2.16744	0.28653	0.83859	2.46175	5.08288	1.007	0.179	0.904(0.894,0.913)
Item7	2.05996	0.17905	0.73111	2.35427	4.97540	0.961	-0.906	0.903(0.891,0.912)
Item8	2.13090	0.25000	0.80205	2.42522	5.04634	1.029	0.665	0.904(0.894,0.913)
Item9	2.21130	0.33039	0.88245	2.50561	5.12674	1.022	0.495	0.904(0.894,0.913)
Item10	1.76965	-0.11126	0.44080	2.06396	4.68509	1.072	1.844	0.910(0.901,0.919)

^aOutfit statistics are not reported but revealed similar results

K10 Kessler 10-item questionnaire; *CI* confidence interval; *Infit* Information-weighted mean-square.

Table 3						
Differential item functioning (DIF) on age (younger and older) and gender (male and female)						
Items	DIF on age			DIF on gender		
	Wald χ^2	DF	p	Wald χ^2	DF	p
Item1	2.385	1	0.123	0.811	1	0.368
Item2	0.922	1	0.337	0.019	1	0.890
Item3	8.339	1	0.004	0.874	1	0.350
Item4	0.010	1	0.920	0.779	1	0.377
Item5	0.338	1	0.561	1.070	1	0.301
Item6	0.174	1	0.677	0.173	1	0.678
Item7	0.086	1	0.770	0.130	1	0.719
Item8	3.597	1	0.058	0.177	1	0.674
Item9	0.225	1	0.636	0.226	1	0.634
Item10	0.757	1	0.384	2.332	1	0.127

Note: *DIF* differential item functioning

Figures

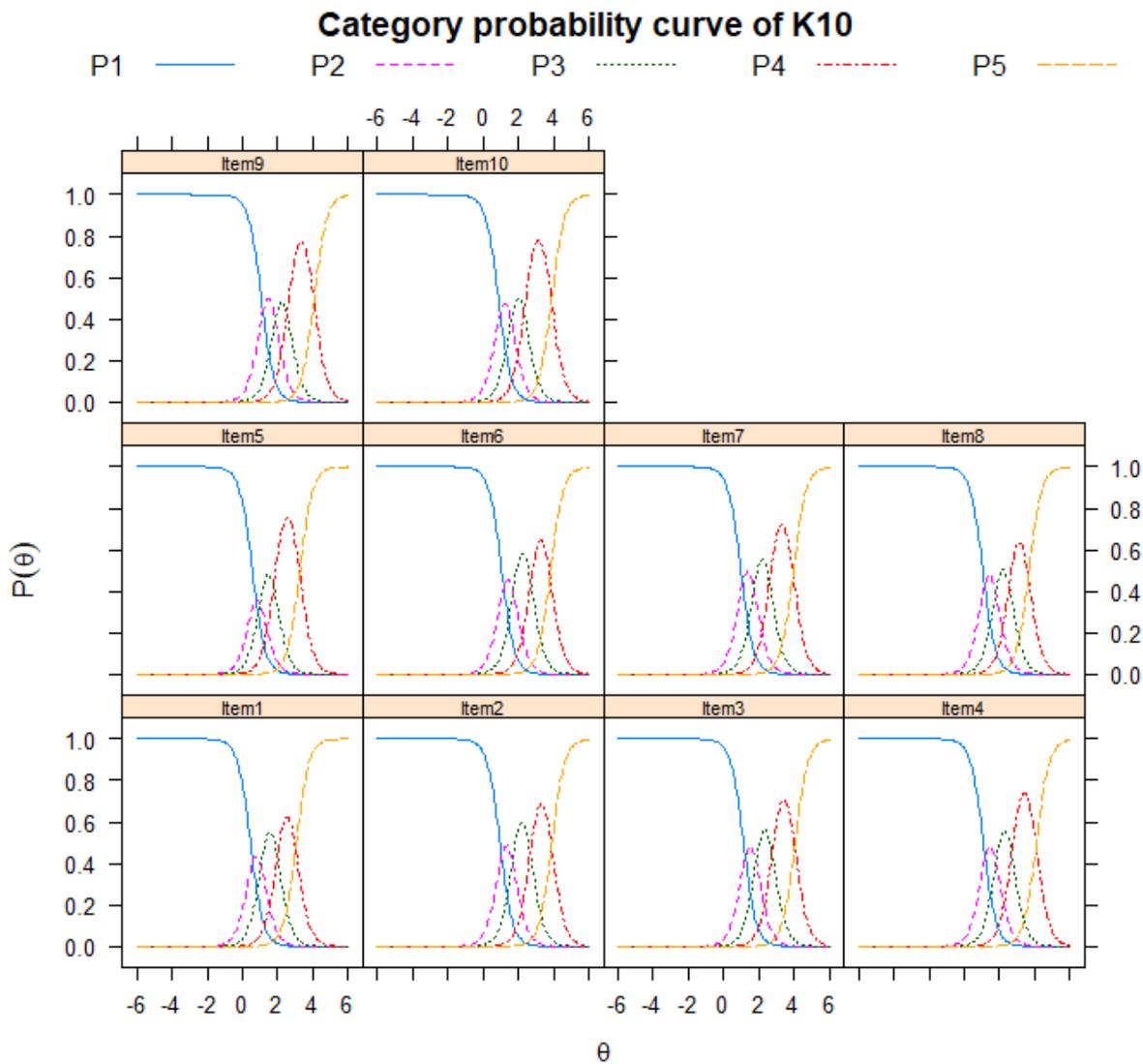


Figure 1

Category probability curves. Category probability curves for the 10-item Kessler Psychological Distress Scale (K10), which illustrated ordered threshold. The five curves from left to right represented 5 response categories (P1 = “none of the time”; P2 = “a little of the time”; P3 = “some of the time”; P4 = “most of the time”; P5 = “all of the time”).

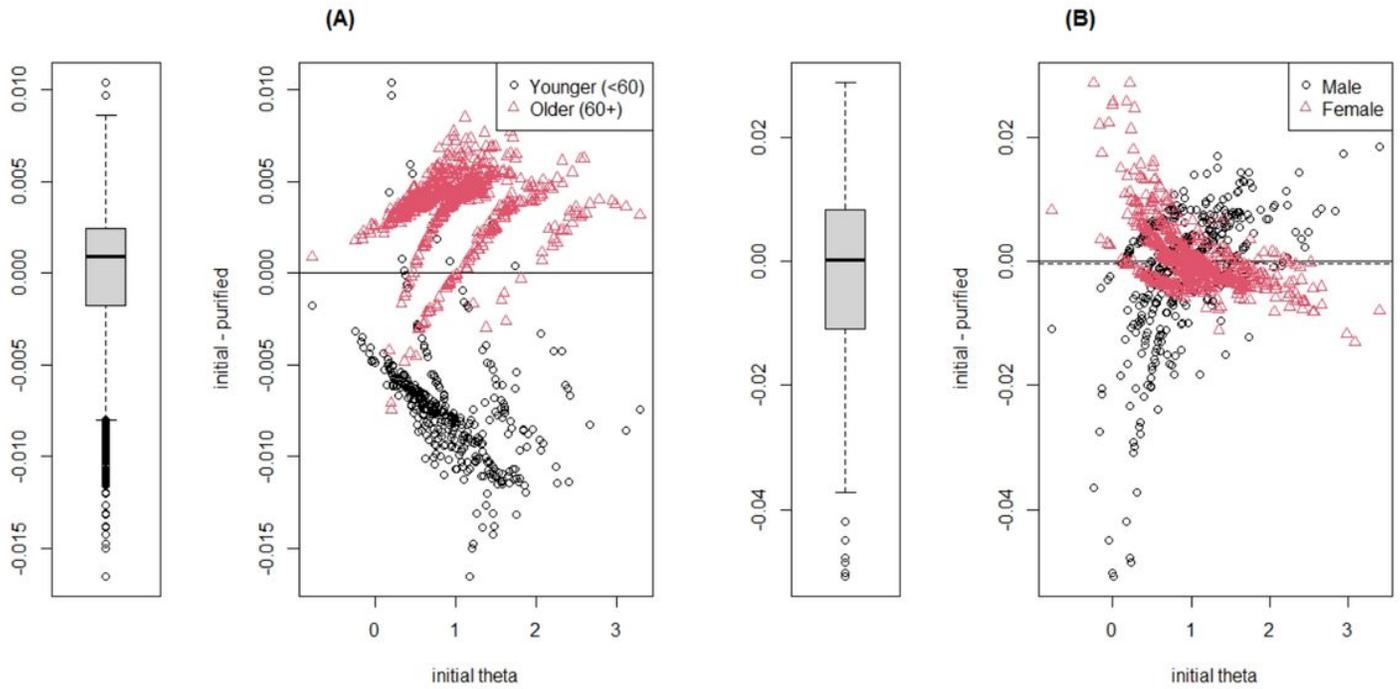


Figure 2

Individual-level DIF impact. (A) Individual-level DIF impact of age; (B) Individual-level DIF impact of gender. Note: These graphs showed the difference in score between using scores that ignored DIF and those that accounted for DIF. The graph on the left showed a box plot of these differences. The interquartile range represented the middle 50% of the differences (bound between the bottom and top of the shaded box). In the graph on the right the same difference scores were plotted against the initial scores ignoring DIF ("initial theta"), separately for younger and older individuals. Guidelines were placed at 0.0 (solid line), i.e., no difference, and the mean of the differences (dotted line).

Person-Item Map

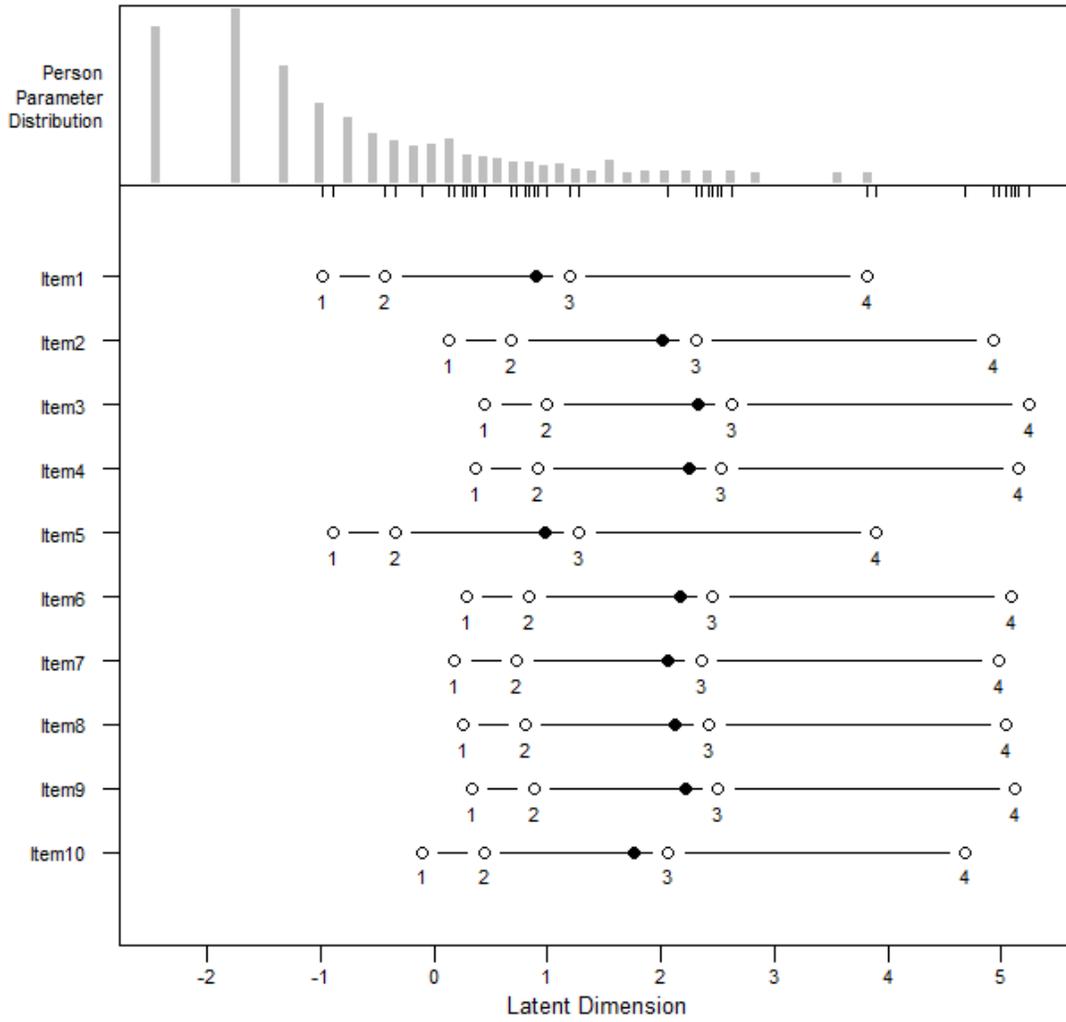


Figure 3

Person-item differential map of K10