



ORIGINAL ARTICLE

A comparative study of confirmatory factor analysis and Rasch Analysis as item reduction strategies for SAMHSA recovery inventory for Chinese (SAMHSA-RIC)



M.Y.L. Chiu^{a,*}, H.T. Wong^b, W.W.N. Ho^c

^a Department of Social and Behavioural Sciences, City University of Hong Kong, Hong Kong

^b Institute of Health Care Management, Department of Business Management, National Sun Yat-sen University, Kaohsiung, Taiwan

^c Beat Drugs Fund Association, Hong Kong

Received 1 July 2019; accepted 4 February 2020

KEYWORDS

Recovery;
SAMHSA-RIC;
Scale reduction;
Rasch Analysis

Abstract

Background and objectives: Factor analysis has been widely used as an item-reduction method, while Rasch Analysis is also beginning to gain some popularity in scale development, with a different perspective and assumptions. In view of the lack of a comparative study, this study reports the comparative use of both strategies in reducing a newly developed inventory based on the conceptual framework of the Substance Abuse and Mental Health Services Administration (SAMHSA) Consensus Statement on the recovery of people with psychosis.

Methods: The effectiveness of confirmatory factor analysis (CFA) and Rasch Analysis (RA) is assessed against the criteria of the number of items reduced, and the percentage of variance is explained for health-related quality of life measures (WHOQOL-BREF).

Results: The SAMHSA Recovery Inventory for Chinese (SAMHSA-RIC) was shortened by CFA and RA from 111 to 72 and 41 items respectively. The percentage of variance explained by the RA shortened SAMHSA-RIC is higher than the CFA shortened SAMHSA-RIC (81.3 % vs 78.4 %).

Conclusion: Evidence suggests that RA appears to be a viable option, in addition to, if not in replacement of, CFA.

© 2020 Asociación Universitaria de Zaragoza para el Progreso de la Psiquiatría y la Salud Mental. Published by Elsevier España, S.L.U. All rights reserved.

Introduction

Reliability and validity are two essential considerations for a good scale. On the one hand, high reliability ensures consistent results, and on the other, good validity ensures that the scale is measuring what it supposed to assess.¹ However,

* Corresponding author.

E-mail addresses: mylchiu@cityu.edu.hk (M.Y.L. Chiu), frankwong@connect.hku.hk (H.T. Wong), wnho@sb.gov.hk (W.W.N. Ho).

a scale with high reliability and validity is a necessary condition but not sufficient condition for clinical use. A lengthy scale can be theoretically useful but practically inconvenient and even troublesome for clinicians due to the high administrative cost and higher chance of respondent and administration error during the course of administration. Patients who have problems with memory and concentration may not be able to complete the scale, and this removes the opportunity to test the scale on low functioning patients. Moreover, even for patients without the above issues of cognitive limitations, the lengthy completion time required may increase the chance of outright rejection, or in practice reduce their motive for cooperation and lower their attentiveness during the course of completion. As such, there is both a practical and research need to shorten these existing high quality but lengthy scales, with a specific aim of maintaining the validity and reliability of the scales.

Currently, the most popular item reduction strategies are based on either classical test theory (CTT) (e.g. factor analysis¹) or item response theory (IRT) (e.g. Rasch Analysis (RA)²). The difference is mainly in the modeling approach. CTT focuses on test-level modeling, whereas IRT is a more sophisticated approach focusing on item-level modeling, and there has been a claim that IRT is theoretically more feasible and generally more robust than CCT.³ Although Wright made a detailed theoretical comparison between factor analysis and Rasch measurement,⁴ comparative studies on the effectiveness of the two approaches are scarce.⁵⁻⁷ The results of some of the existing empirical comparative studies have shown that the two strategies had similar efficacy, and so far no conclusive remarks on the preference can be made. Table 1 contains examples of the results obtained from previous studies. Although the table shows a comparison between the two methods based on the number of items reduced, the correlation between the original and shortened scales, and the internal consistency of the shortened scale, none of them have demonstrated the relative efficacy of the two methods, and it is left almost entirely up to the researcher as to which item-reduction approach to take.

It has been noted that the scales tested in these studies were relatively short (ranging from 16 to 38 items). It is therefore in our academic interest to find out if the method of confirmatory factor analysis (CFA)¹ and RA for item reduction works equally well in the case of a longer scale such as the 111-item Substance Abuse and Mental Health Services Administration – Recovery Inventory for Chinese (SAMHSA-RIC).^{8,9} Readers may refer to a publication by Chiu et al.⁸ for the background of the recovery model, and some elaboration on the conceptual basis of the construct of mental health recovery.

Material and methods

Item reduction

The first strategy, CFA, is a popular method for confirming the number of factors in a scale. In contrast to exploratory factor analysis, where the number of factors is not known in advance, a path diagram defining the casual relationship between the factor(s) and items of a questionnaire has to be defined explicitly. In this study, the technique of struc-

tural equation modeling (SEM) was used for estimating the path coefficient of the path diagrams. As path coefficients represent the strength of relationship between items and factor(s), items with small path coefficients will be discarded from the scale. Hence, the number of items in the scale can be reduced while keeping the loss of information to a minimum.¹

The second strategy used in the study was RA.² One of the characteristics of RA is locating the response of each respondent on a single dimensional item-person map according to a probabilistic relation between persons' "ability" and items' "difficulty". This allows further analysis of the items' discriminative power on the respondents' ability through the use of item-person map. Items with low discriminative power will be deleted; the definition of low discriminative power depends on each scale. Item reduction using RA is not uncommon in different research areas. Similar analysis with the use of Rasch model can be found in different fields such as psychology,^{10,11} optometry,^{12,13} rehabilitation^{14,15} and education.¹⁶ The increasing popularity demonstrates that RA could be a new and effective tool for item reduction.

In this research, we aimed to compare the performance of CFA and RA. We hypothesized that the novel RA would be superior to the traditional CFA. Throughout the research, the CFA was done using Amos 16, the RA was completed with RUMM2020 (www.rummlab.com.au), and SPSS 16 was used for the rest of the analysis. The RA item reduction process was carried out on each subscale separately. Five statistical indexes were used in sequence for reducing the items: the individual item's χ^2 fit statistics, differential item functioning, item residual score, item-person maps, and item-trait interaction χ^2 fit statistics. The detailed steps of the RA item reduction process have been documented by Chiu et al.⁹ For CFA, items with standardized path coefficient loadings smaller than 0.30 are removed from the subscales and the subsequent structural model. The detailed steps of the CFA item reduction process have been documented by Ho.¹⁷ To compare the performance of CFA and RA, the number of reduced items and the reliability of each subscale were compared first. Afterwards, SEM was used to compare the structural models' direct, indirect, and total effects, and their different fit statistics (i.e. χ^2 , Tucker-Lewis index, root mean square error of approximation, comparative fit index, and percentage of variance explained for the response variable) before and after item reduction.

Instruments

The SAMHSA-RIC includes eleven (sub)scales in total. A six-item Adult State Hope Scale (ASHS), which assesses a respondent's optimism about achieving their goals, was used for assessing patients' hope. It shows good internal consistency, with an α value ranging from 0.79 to 0.95.¹⁸ The Recovery Attitude Questionnaire (RAQ-7) was used to assess the non-linear understanding of the recovery process. It was developed in the USA by consumer groups, and aims at measuring adherence to recovery values. Its test-retest reliability and internal consistency ($\alpha = 0.84$) meet conventionally accepted standards in American subjects.¹⁹ The scale consisted of seven items in the original version; how-

Table 1 Previous research on comparing item response theory and classical test theory on item reduction.

	Number of items reduced		Correlation between the original and shortened scales		Cronbach's alpha of the shortened scale	
	CTT ^a	Rasch	CTT ^a	Rasch	CTT ^a	Rasch
Prieto et al. ⁵	38→20	38→22	0.97	0.97	0.82-0.88	0.87-0.88
Nijsten et al. ⁶	16→10	16→11	0.94	0.96	0.52-0.85	0.83
Erhart et al. ⁷	19→13	19→11	0.96	0.93	0.86	0.85

^a CTT: Classical test theory.

ever, the Cronbach's alpha of the scale dropped to 0.56 when it was translated into Chinese and tested among the Hongkongers, who held widely different recovery beliefs. In view of this development, two items were removed from the scale because of low item-total correlation, and the subsequent internal consistency was raised to 0.63. The Health Care Climate Questionnaire (HCCQ) – a scale measuring the degree of autonomy support that clients perceive their psychiatrists to provide was used to assess person-centered treatment. The six-item scale has demonstrated good predictive validity and internal consistency ($\alpha=0.96$) in previous studies.²⁰ The self-responsibility and initiative subscale of the Exercise of Self Care Agency Scale (ESCA) was adopted to assess self-responsibility.²¹ The twelve-item subscale assessed the respondent's initiative in regard to maintaining their health, and it was validated in different studies.²² The 17-item personal competence subscale of the Resilience Scale (RS) was adopted to assess the strength of patients.²³ The range of internal consistency α value of RS was found between 0.76 and 0.91 in previous studies. The seven-item Mastery Scale (MS) was used to assess the sense of self direction.²⁴ It measured the subjective rating of a respondent's ability to exercise control in the daily course of life. It showed good reliability and validity among people with severe mental disorders ($\alpha=0.73$).^{25,26} The nine-item self-efficacy and self-esteem subscale of the Making Decision Empowerment scale (MDES) was adopted to assess empowerment.²⁷ It was constructed to measure respondents' level of empowerment. The scale has been validated in the Sweden and USA, and good reliability and validity was found with $\alpha=0.90$.²⁸ The perceived discrimination, alienation, and the social withdrawal subscale of the Internalized Stigma of Mental Illness Scale (ISMI) were adopted to measure perceived respect. Its reliability and validity were well-established and so reflected the subjective experience of stigma among populations with mental disorders.²⁹

Holistic recovery was measured through three aspects: social support (community), spirituality, and psychosocial symptoms (emotion and mind). The psychosocial subscale of the Schizophrenia Quality of Life Scale (SQLS) (15-item) was used for measuring the frequency of psychosocial symptoms.³⁰ It has shown excellent internal validity and reliability, with internal consistency α value 0.93. The multi-dimensional scale of Perceived Social Support-Chinese version (MSPSS-C) was adopted to assess social support. The scale has twelve items, indicating perceived social support from family members and friends, and with a Cronbach's

alpha 0.89.³¹ The WHO Spirituality Religion and Personal Belief Scale – Hong Kong version (WHO-SRPB-HK) was used to assess spirituality. Within the eight facets of the scale, three (connectedness to a spiritual being or force, faith, and spiritual strength subscales) were chosen for this study because they were the only three facets that provided a pure assessment of respondents' spirituality; therefore, they were not complicated by a respondent's mental health condition.³² These four-item subscales have shown excellent validity and reliability, with an internal consistency α value ranging from 0.77 to 0.95.³³ The Hong Kong Chinese WHO Quality of Life Measure (abbreviated version) (WHOQOL-BREF(HK)) was used to assess the respondents' quality of life. It has well-established validity and reliability among Chinese subjects with schizophrenia. The WHOQOL perception among schizophrenia subjects has been found having negative correlation with psychiatric ratings.^{34,35} The WHOQOL-BREF(HK) (28-item) has four domains and two specific questions that indicate overall health and overall quality of life. The internal consistency α of the four domains and the test-retest reliability of items ranged from 0.67 to 0.79 and 0.64 to 0.90 respectively. Subjective rather than objective quality of life indicators was adopted because they tapped into multiple aspects of actual experiences and provided a more accurate reflection of personal wellbeing.

Structural model of the SAMHSA-RIC and WHOQOL-BREF

Ho et al.³⁶ proposed a structural model for the casual relationship of the eleven recovery components and quality of life. It was able to explain 81 % of the WHOQOL-BREF outcome. The model mainly stated that psychosocial symptoms (SQLS) contribute to internal stigma (ISMI), but also make an indirect and direct contribution to WHOQOL-BREF through perceived support. At the same time, perceived support is manifested in three observable domains: autonomy support from clinicians (HCCQ); spirituality (WHOQOL-SRPB-HK); and social support from family and friends (MSPSS-C). Moreover, there are two mediators between perceived support and WHOQOL-BREF: optimism and personal agency. Optimism is manifested in two observable domains, motivational hope (ASHS) and general hope (RAQ-7). On the other hand, personal agency is manifested in four observable domains including empowerment (MDES), sense of mastery in daily setting (MS), resilience (RS), and responsibility in self-care (ESCA). The original scale has 111 items and is considered the first attempt to operationalize the recovery concept in

the Asian context,⁸ yet the lengthy list of items mean there is a need for item reduction, be it CFA or RA.

Ethical statement

Ethics approval of the original study had been obtained from the Hong Kong Hospital Authority Cluster Research Ethics Committees before collecting the data.

Results

Sample

This RA is essentially a secondary analysis of the anonymous measurement data of a larger study that has obtained ethics approval from the university and health authority. The original study involved more than 200 eligible subjects from two psychiatric outpatient clinics in Hong Kong. The inclusion criteria of this study were: a) aged between 18 and 60; and b) with a primary diagnosis of "schizo-affective", "schizophreniform", or "schizophrenia" disorder. The exclusion criteria of this study were: a) lack of ability to communicate in Cantonese; b) global score of less than four in the CapQOL screening assessment; and c) having been discharged from a psychiatric ward during the 30 days preceding the interview. Written consent was obtained from the participants after they were given a complete description of the study.

Sample characteristics

Among the 204 subjects, around half were male with a mean age equaled to 42 (SD = 9.15). Around one-third (29 %) were married while nearly 80 % were living with family members. More than 80 % had education above primary school level, and around half were currently employed. 16 %, 6%, and 12 % of the patients had engaged in vocational, residential, and community support services in the past 12 months respectively. Detailed demographic information can be found in Chiu et al.⁸

Item reduction

The SAMHSA-RIC was successfully shortened by CFA and RA from the original 111-item scale to 72 and 41 items respectively. At subscale level, the number of items reduced by RA was consistently more than by CFA. In particular, for the subscales "Holistic Wellbeing-Social Support: MSPSS" and "Holistic Wellbeing-Psychosocial symptoms: SQLS", CFA could not reduce any item lists, but the RA was able to reduce most of them (Table 2).

Scale reliability

Table 3 shows that after item reduction by either CFA or RA, most of the SAMHSA-RIC subscales' Cronbach alphas had a certain degree of decrease. The Cronbach alpha of the eleven original subscales ranged from 0.65 to 0.95. After reduction, the range dropped to 0.62-0.95 and 0.62-0.89 for CFA and RA respectively. All the alpha values of the scales

shortened by either method were within acceptable range ($\alpha > 0.6$). In addition, some of the subscales, such as holistic wellbeing-spirituality and respect-stigma, maintained a high level of internal consistency alpha value ($\alpha > 0.8$).

Effects of independent variables on WHOQOL-BREF

The three versions of the SAMHSA-RIC (i.e. the 111, 72, and 41-item versions) were analyzed by confirmatory factor analysis based on the proposed model so that the percentage explained for WHOQOL-BREF and the direct, indirect, and total effects of the response variables on WHOQOL-BREF in the model can be compared between the three scales.

Table 4 is the result of the direct, indirect, and total effects of the independent variables on WHOQOL-BREF and its sub-domains. We can see that for the independent variables of support, optimism, and personal agency, the parameters representing their effects on WHOQOL-BREF and the sub-domains were strengthened. Although the parameters obtained by using RA were not as large as the parameters obtained by CFA, the differences were actually very small. This is because the average and maximum differences of pre-post item reduction were 0.04 (8.3 %) and 0.06 (13.7 %) respectively for CFA, whereas the corresponding values for RA were 0.02 (4.7 %) and 0.03 (7.7 %) respectively.

In contrast, for psychosocial symptoms, the effects on WHOQOL-BREF and the sub-domains were weakened after item reduction. Both CFA and RA gave similar average and maximum differences. For CFA, the average and maximum differences of pre-post item reduction were 0.02 (4.3 %) and 0.03 (5.6 %) respectively, whereas the corresponding values for RA were 0.02 (4.6 %) and 0.03 (5.7 %) respectively.

Finally, the situation for internal stigma was relatively complicated because CFA yielded weakened effects whereas RA yielded strengthened effects. For CFA, the average and maximum decreases of pre-post item reduction were 0.02 (6.8 %) and 0.03 (8.0 %) respectively, whereas the corresponding increases for RA were 0.02 (5.8 %) and 0.03 (7.7 %) respectively.

Model fit statistics

Table 5 provides a comprehensive assessment of the three scales. It is clear that all three scales showed acceptable fit for the comparative fit index (CFI), Tucker-Lewis index (TLI), and χ^2/df , whereas the root mean square error of approximation (RMSEA) had adequate fit. Although the above four indexes showed no difference, the performance of the three models can be distinguished by the percentage variance explained for WHOQOL-BREF. The results showed that the scale shortened by RA had a percentage variance explained for WHOQOL-BREF that was higher than the original 111-item scale (81.3 % vs 80.7 %). On the other hand, for the scale shortened by CFA, the percentage variance explained for WHOQOL-BREF was lower than the original 111-item scale (78.4 % vs 80.7 %).

Table 2 Number of items in the SAMHSA-RIC after and before item reduction.

Scale	Number of items		
	Pre item reduction	Post item reduction by CFA	Post item reduction by Rasch
Hope: ASHS	6	6	4
Non-linear Recovery: RAQ-7	4	2	2
Person Centered: HCCQ	6	5	3
Self-Responsibility: ESCA-ISR	10	5	4
Strength Based: RS-PC	6	5	5
Self-Direction: MS	5	4	4
Empowerment: MDES-SESE	18	6	4
Respect-Stigma: ISMI	17	7	6
Holistic Wellbeing-Social Support: MSPSS	12	12	2
Holistic Wellbeing-Spirituality: WHOQOL-SPRB	12	12	4
Holistic Wellbeing-Psychosocial symptoms: SQLS	15	8	3
Total	111	72	41

Table 3 Cronbach alpha of the SAMHSA-RIC before and after item reduction.

Scale	Cronbach alpha		
	Pre item reduction	Post item reduction by CFA	Post item reduction by Rasch
Hope: ASHS	0.77	0.78	0.68
Non-linear Recovery: RAQ-7	0.62	0.62	0.62
Person Centered: HCCQ	0.84	0.84	0.73
Self-Responsibility: ESCA-ISR	0.83	0.79	0.73
Strength Based: RS-PC	0.76	0.76	0.68
Self-Direction: MS	0.70	0.68	0.66
Empowerment: MDES-SESE	0.92	0.84	0.73
Respect-Stigma: ISMI	0.91	0.87	0.82
Holistic Wellbeing-Social Support: MSPSS	0.92	0.92	0.69
Holistic Wellbeing-Spirituality: WHOQOL-SPRB	0.95	0.95	0.89
Holistic Wellbeing-Psychosocial symptoms: SQLS	0.93	0.90	0.74

Discussion

CFA has probably been the most popular method for psychometric evaluation of inventories.¹ It examines the proposed structure that represents how manifest and latent constructs are fitted in. The theoretical model is often drawn from the existing literatures of both theoretical and empirical understanding on the subject matter. The measurement model has thereafter been built to test whether there is a good fit. The criterion is whether the application of data collected meets the conventional threshold of CFA statistics. Although CFA is not designed for item reduction of scales, by identifying items with weaker loading or setting an arbitrary but reasonable cut-off loading, items can be removed in order to test whether the remaining ones can still maintain the same factor structure. In this way, both items with insignificant factor loadings and items under the cut-off threshold could be reduced.

On the other hand, RA has grown out of the response theory that good item responses are those that can differentiate between the respondents in the intended outcome.²

Items that everybody either scored or missed are poor items – unable to differentiate. Although one may argue that the use of RA would have presumed the relevance of the item in measuring the construct that the scale is supposed to measure, RA has left it to the researcher to decide which items are there and why. RA makes no assumption about the underlying structure or model. Therefore, RA would have a wider application, and its technicality ranges from simple hand-scoring analysis to complicated software applications.^{37,38}

This paper has set out to compare the differences in shortening the SAMHSA-RIC using CFA and RA. The intention is similar to what Wright did,⁴ but this study is focused on comparing the performance of the two models empirically^{5–7} rather than comparing their model structure theoretically.⁴ We found that RA could reduce around two-thirds of the number of items, but only one-third could be reduced with the use of CFA. In some particular subscales (holistic wellbeing-social support and holistic wellbeing-spirituality), no item lists could be reduced by CFA, but RA could reduce most of them in the process. It is clear that CFA keeps the items because they are internally coherent in relation to

Table 4 Total, direct and indirect effects of independent variables on WHOQOL-BREF.

Parameters for the models			Pre item reduction			Post item reduction by CFA			Post item reduction by Rasch		
			Direct effect	Indirect effect	Total effect	Direct effect	Indirect effect	Total effect	Direct effect	Indirect effect	Total effect
Support	→	WHOQOL-BREF	–	0.465	0.465	–	0.520	0.520	–	0.492	0.492
Support	→	WHOQOL-Physical	–	0.362	0.362	–	0.408	0.408	–	0.390	0.390
Support	→	WHOQOL-Psychological	–	0.428	0.428	–	0.474	0.474	–	0.449	0.449
Support	→	WHOQOL-Social	–	0.293	0.293	–	0.327	0.327	–	0.308	0.308
Support	→	WHOQOL-Environmental	–	0.358	0.358	–	0.407	0.407	–	0.381	0.381
Optimism	→	WHOQOL-BREF	–	0.543	0.543	–	0.577	0.577	–	0.560	0.560
Optimism	→	WHOQOL-Physical	–	0.423	0.423	–	0.453	0.453	–	0.444	0.444
Optimism	→	WHOQOL-Psychological	–	0.500	0.500	–	0.526	0.526	–	0.511	0.511
Optimism	→	WHOQOL-Social	–	0.342	0.342	–	0.363	0.363	–	0.351	0.351
Optimism	→	WHOQOL-Environmental	–	0.418	0.418	–	0.451	0.451	–	0.434	0.434
Personal Agency	→	WHOQOL-BREF	0.584	–	0.584	0.620	–	0.620	0.610	–	0.610
Personal Agency	→	WHOQOL-Physical	–	0.455	0.455	–	0.487	0.487	–	0.484	0.484
Personal Agency	→	WHOQOL-Psychological	–	0.538	0.538	–	0.565	0.565	–	0.558	0.558
Personal Agency	→	WHOQOL-Social	–	0.368	0.368	–	0.390	0.390	–	0.383	0.383
Personal Agency	→	WHOQOL-Environmental	–	0.450	0.450	–	0.484	0.484	–	0.473	0.473
Psychosocial Symptom	→	WHOQOL-BREF	–0.288	–0.352	–0.640	–0.265	–0.347	–0.611	–0.289	–0.320	–0.609
Psychosocial Symptom	→	WHOQOL-Physical	–	–0.498	–0.498	–	–0.480	–0.48	–	–0.483	–0.483
Psychosocial Symptom	→	WHOQOL-Psychological	–	–0.590	–0.590	–	–0.557	–0.557	–	–0.556	–0.556
Psychosocial Symptom	→	WHOQOL-Social	–	–0.403	–0.403	–	–0.384	–0.384	–	–0.382	–0.382
Psychosocial Symptom	→	WHOQOL-Environmental	–	–0.493	–0.493	–	–0.478	–0.478	–	–0.472	–0.472
Internal Stigma	→	WHOQOL-BREF	–0.244	–0.176	–0.421	–0.198	–0.193	–0.391	–0.221	–0.223	–0.444
Internal Stigma	→	WHOQOL-Physical	–	–0.327	–0.327	–	–0.307	–0.307	–	–0.352	–0.352
Internal Stigma	→	WHOQOL-Psychological	–	–0.388	–0.388	–	–0.357	–0.357	–	–0.406	–0.406
Internal Stigma	→	WHOQOL-Social	–	–0.265	–0.265	–	–0.246	–0.246	–	–0.278	–0.278
Internal Stigma	→	WHOQOL-Environmental	–	–0.324	–0.324	–	–0.306	–0.306	–	–0.344	–0.344

WHOQOL-BREF, World Health Organization Quality of Life Measure Abbreviated version.

Table 5 Summary of model fit with percentage of variance explained for recovery measured by the three versions of the SAMHSA-RIC.

Scales	χ^2	df	χ^2 / df	TLI	CFI	RMSEA	% variance explained for WHOQOL-BREF
Pre item reduction (111 Items)	202.72	84	2.41	0.90	0.92	0.08	80.7 %
Post item reduction by CFA (72 Items)	185.60	84	2.21	0.90	0.93	0.07	78.4 %
Post item reduction by Rasch (41 Items)	178.02	84	2.12	0.91	0.93	0.08	81.3 %

CFI, Comparative fit index (close to 1, excellent fit; ≥ 0.90 , acceptable fit; < 0.90 , poor fit); RMSEA, root mean square error of approximation (< 0.05 , good fit; 0.05 – 0.08 , adequate fit; 0.08 – 0.10 , mediocre fit; > 0.10 , poor fit); TLI, Tucker-Lewis index (≥ 0.95 , good fit; 0.90 – 0.95 , acceptable fit; < 0.90 , poor fit); χ^2 / df = normed χ^2 (< 1 , poor model fit; 1 – 2 , excellent fit; 2 – 5 , acceptable fit).

the construct. However, close resemblance in items, though internally coherent, does not help with classifying respondents by response. RA does not have this issue. As long as the responses to these items are not identical, there will be differences for RA in determining which items among the group go first, and dropping items that have poor classification power.

It is important to note that the RA not only managed to remove more items than CFA, but also brought a higher percentage of explained variance in quality of life measures compared to the original SAMHSA-RIC and the CFA-shortened SAMHSA-RIC. The result demonstrates that it is not necessary for more information to be lost if more items are deleted. It is highly plausible that when some items are deleted, distractive information associated with the item is also removed, resulting in a better model.

Although the path coefficients of the SAMHSA-RIC shortened by CFA were generally larger, the differences between the two item reduction strategies were practically negligible. Moreover, the final scale will not have the problem of biased weighting on certain components. This is because the scale was originally developed to view the recovery status of each component separately. Although considering each component separately will make uni-dimensionality unfeasible, this could be achieved if all items are aligned to the same scoring options.⁹ In this study, we considered RA as the better method for shortening the SAMHSA-RIC. The result we obtained not only means a shorter and better version of the SAMHSA-RIC, it also demonstrates that RA possesses great potential for item reduction while the validity and reliability of the scales can still be maintained. Use of RA should be encouraged in research studies of psychiatry³⁹ and rehabilitation.^{40,41}

Conclusions

This study has shown that both CFA and RA can be effective item reduction strategies for inventory development, in this case SAMHSA-RIC. CFA successfully reduced the quantity of the original scale by 35 % and RA 63 %. Apparently, RA enjoys better efficacy than CFA in terms of number of items deleted. The Structural Equation Model of the RA-shortened scale also had a higher percentage of explained variance than that of the CFA-shortened scale. The results show the promising potential of RA, at least for those scales with multiple dimensions and long lists of items. However, this is not to suggest that RA be used generally for scale reduction.

More systematic empirical studies are warranted to confirm this advantage in long inventories. It is safe at least to say that RA can be used as an additional method for reference with CFA, if it is not intended to replace CFA.

Ethical statement

Ethics approval of the original study had been obtained from the Hong Kong Hospital Authority Cluster Research Ethics Committee before collecting the data.

Funding

There was no funding for this work.

Conflict of interest

The authors have no conflict of interest to declare.

Acknowledgment

None.

References

1. DeVellis RF. Scale development: theory and applications. 4th edn. Los Angeles: Sage Publications; 2017.
2. Bond TG, Fox CM. Applying the Rasch model: fundamental measurement in the human sciences. 3rd edn. Routledge; 2015.
3. Humphrey N, Kalambouka A, Wigelsworth M, Lendrum A, Deighton J, Wolpert M. Measures of social and emotional skills for children and young people: a systematic review. *Educ Psychol Meas*. 2011;71:617–37.
4. Wright BD. Comparing factor analysis and Rasch measurement. *Rasch Meas Trans*. 1994;8:350.
5. Prieto L, Alonso J, Lámarcha R. Classical test theory versus Rasch Analysis for quality of life questionnaire reduction. *Health Qual Life Outcomes*. 2003;1:27.
6. Nijsten T, Unaëze J, Stern RS. Refinement and reduction of the Impact of Psoriasis Questionnaire: Classical Test Theory vs. Rasch Analysis. *Br J Dermatol*. 2006;154:692–700.
7. Erhart M, Hagquist C, Auquier P, Rajmil L, Power M, Ravens-Sieberer U. European KIDSCREEN Group: a comparison of Rasch item-fit and Cronbach's alpha item reduction analysis for the development of a Quality of Life scale for children and adolescents. *Child Care Health Dev*. 2010;36:473–84.

8. Chiu MY, Ho WW, Lo WT, Yiu MG. Operationalization of the SAMHSA model of recovery: a quality of life perspective. *Qual Life Res.* 2010;19:1–3.
9. Chiu MYL, Wong FHT, Ho WWN. Rasch Analysis of the SAMHSA Recovery Inventory for Chinese (SAMHSA-RIC). *Int J Soc Psychiatry.* 2014;60:254–62.
10. Meads DM, Bentall RP. Rasch Analysis and item reduction of the hypomanic personality scale. *Pers Individ Dif.* 2008;44:1772–83.
11. Dreer LE, Berry J, Rivera P, Snow M, Elliott TR, Miller D, Little TD. Efficient assessment of social problem-solving abilities in medical and rehabilitation settings: a Rasch Analysis of the social problem-solving inventory-revised. *J Clin Psychol.* 2009;65:653–69.
12. Pesudovs K, Garamendi E, Keeves JP, Elliott DB. The activities of daily vision scale for cataract surgery outcomes: re-evaluating validity with Rasch Analysis. *Invest Ophthalmol Vis Sci.* 2003;44:2892–9.
13. Ryan B, Court H, Margrain TH. Measuring low vision service outcomes: Rasch Analysis of the Seven-Item National Eye Institute Visual Function Questionnaire. *Optom Vis Sci.* 2008;85:12–21.
14. Vidotto G, Carone M, Jones PW, Salini S, Bertolotti G. Quess Group: Maugeri Respiratory Failure questionnaire reduced form: a method for improving the questionnaire using the Rasch model. *Disabil Rehabil.* 2007;29:991–8.
15. Siegert RJ, Tennant A, Turner-Stokes L. Rasch Analysis of the Beck Depression Inventory-II in a neurological rehabilitation sample. *Disabil Rehabil.* 2010;32:8–17.
16. Waugh R, editor. *Specialized Rasch measures applied at the forefront of education.* New York: Nova Science Publishers; 2010.
17. Ho WW. *Recovery Model among Chinese People with Schizophrenia.* PhD Thesis. Hong Kong: Hong Kong Baptist University; 2008.
18. Snyder CR, Simpson SC, Ybasco FC, Borders TF, Babyak MA, Higgins RL. Development and validation of the state hope scale. *J Per Soc Psychol.* 1996;70:321–35.
19. Ralph RO, Kidder K, Philips D. *Can we measure recovery? A compendium of Recovery Measures.* Cambridge, MA: Human Services Research Institute; 2000.
20. Williams GC, Rodin GC, Ryan RM, Grolnick WS, Deci EL. Autonomous regulation and long-term medication adherence in adult outpatients. *Health Psychol.* 1998;17:269–76.
21. Kearney BY, Fleischer BJ. Development of an instrument to measure exercise of self-care agency. *Res Nurs Health.* 1979;2:25–34.
22. Riesch SK, Hauck MR. The Exercise of Self-Care Agency: an analysis of construct and discriminant Validity. *Res Nurs Health.* 1988;11:245–55.
23. Wagnild GM, Young HM. Development and psychometric evaluation of the Resilience Scale. *J Nurs Meas.* 1993;1:165–78.
24. Pearlin L, Schooler C. The structure of coping. *J Health Soc Behav.* 1978;19:2–21.
25. Bengtsson-Tops A. Mastery in patients with schizophrenia living in the community: relationship to sociodemographic and clinical characteristics, needs for care and support, and social network. *J Psychiatr Ment Health Nurs.* 2004;11:298–304.
26. Rosenfield S. Factors contributing to the subjective quality of life of the chronic mentally ill. *J Health Soc Behav.* 1992;33:299.
27. Rogers ES, Chamberlin J, Ellison ML, Crean T. A consumer-constructed scale to measure empowerment among users of mental health services. *Psychiatr Serv.* 1997;48:1042–7.
28. Hansson L, Björkman T. Empowerment in people with a mental illness: reliability and validity of the Swedish version of an empowerment scale. *Scand J Caring Sci.* 2005;19:32–8.
29. Boyd RJ, Otilingam PG, Grajales M. Internalized stigma of mental illness: psychometric properties of a new measure. *Psychiatry Res.* 2003;121:31–49.
30. Wilkinson G, Hesdon B, Wild D, Cookson R, Farina C, Sharma V, Fitzpatrick R, Jenkinson C. Self-report quality of life measure for people with schizophrenia: the SQLS. *Br J Psychiatry.* 2000;177:42–6.
31. Chou KL. Assessing Chinese adolescents' social support: the multidimensional scale of perceived social support. *Pers Individ Dif.* 2000;28:299–307.
32. Moreira-Almeida A, Koenig HG. Retaining the meaning of the words religiousness and spirituality: a commentary on the WHO-QOL SRPB group's a cross-cultural study of spirituality, religion, and personal beliefs as components of quality of life(62: 6, 2005, 1486-1497). *Soc Sci Med.* 2006;63:843–5.
33. Group WS. A cross-cultural study of spirituality, religion, and personal beliefs as components of quality of life. *Soc Sci Med.* 2006;62:1486–97.
34. Chan GW, Ungvari GS, Shek DT, Leung Dagger JJ. Hospital and community-based care for patients with chronic schizophrenia in Hong Kong: quality of life and its correlates. *Soc Psychiatry Psychiatr Epidemiol.* 2003;38:196.
35. Chan S, Yu IW. Quality of life of clients with schizophrenia. *J Adv Nurs.* 2004;45:72–83.
36. Ho WW, Chiu MY, Lo WT, Yiu MG. Recovery components as determinants of the health-related quality of life among patients with schizophrenia: structural equation modelling analysis. *Aust N Z J Psychiatry.* 2010;44:71–84.
37. RUMM Laboratory Pty Ltd. RUMM for analysing assessment and attitude questionnaire data. Available at: <http://www.rummlab.com.au/>. [Accessed on 29 January 2020].
38. Winsteps. WINSTEPS: Multiple-Choice, Rating Scale and Partial Credit Rasch Analysis. Available at: <https://winsteps.com/winsteps.htm>. [Accessed on 29 January 2020].
39. da Rocha NS, Chachamovich E, de Almeida Fleck MP, Tennant A. An introduction to Rasch Analysis for psychiatric practice and research. *J Psychiatr Res.* 2013;47:141–8.
40. Tesio L. Measuring behaviours and perceptions: Rasch Analysis as a tool for rehabilitation research. *J Rehabil Med.* 2003;35:105–15.
41. Tesio L, Simone A, Bernardinello M. Rehabilitation and outcome measurement: where is Rasch Analysis-going? *Eura Medicophys.* 2007;43:417.