The current issue and full text archive of this journal is available on Emerald Insight at:
https://www.emerald.com/insight/2586-940X.htm

Received 11 August 2020
Revised 21 October 2020
17 November 2020
Accepted 4 December 2020

A tale of dual-approach construct validation and reliability testing for a *Zika* infection awareness knowledge questionnaire

Chin Mun Wong, Mohd Rohaizat Hassan, Rozita Hod and Sharifa Ezat Wan Puteh
Department of Community Health, Faculty of Medicine, Universiti Kebangsaan Malaysia, Kuala Lumpur, Malaysia, and
Sazaly Abu Bakar
Tropical Infectious Diseases Research and Education Centre(TIDREC), Universiti Malaya, Kuala Lumpur, Malaysia

Abstract

Purpose – This study aims to validate the English version of a WHO-adapted questionnaire: Zika infection awareness/knowledge questionnaire using a unique dual-approach validation model.

Design/methodology/approach – A cross-sectional pilot study of 30 adult respondents in Malaysia completed the self-administered questionnaire on knowledge and perception to Zika infection. Construct validity was assessed by exploratory factor analysis (EFA) of SPSS and Rasch partial credit. Reliability is tested using pKR20 and Cronbach's alpha.

Findings – Knowledge construct was unidimensional, good model fit, easy to endorse and well discriminative. Five-rating Likert scale for perception domain was appropriate. Knowledge domain should be separated into 6 level of difficulties. Perception domain should remain as one construct. Knowledge domain was highly reliability (pKR20 = 0.96), perception domain was fairly reliable (Cronbach's alpha = 0.641). Respondent's ability to answer knowledge domain and perception domain were separated into 3 and 4 levels.

Research limitations/implications – Small sample size may affect factor analysis.

Practical implications – The questionnaire has good psychometric properties to measure the knowledge and perception of Zika infection among Malaysian community.

Social implications – The questionnaire helped to gauge knowledge and perception of the general community in Malaysia to aid preparation of health education tool for Zika infection.

© Chin Mun Wong, Mohd Rohaizat Hassan, Rozita Hod, Sharifa Ezat Wan Puteh and Sazaly Abu Bakar. Published in *Journal of Health Research*. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at http://creativecommons.org/licences/by/4.0/legalcode

The authors would like to express their gratitude to Universiti Kebangsaan *Malaysia* and the Ministry of Higher Education Malaysia, the Director General of Ministry of Health Malaysia, Perak State Health Department and the Dean of Faculty of Medicine, Universiti Kebangsaan Malaysia for their support and funding to conduct this project. This project is a part of the doctorate research of Dr Wong Chin Mun which is supported and funded by grant LRGS MRUN/F1/01/2018 and UKM grant (code FF-2019-167). Special thanks to those involved directly and indirectly in the research, including the District Health Officer of Batang Padang District of Perak State, Dr. Raja Mohd Azim bin Raja Haron and his team.

Journal of Health Research
Vol. 36 No. 6, 2022
pp. 1078-1090
Emerald Publishing Limited
e-ISSN: 2586-940X
p-ISSN: 0857-4421
DOI 10.1108/JHR-08-2020-0332

Downloaded from http://ftp.nowpublishers.com/jhr/article-pdf/36/6/1078/1488659/jhr-08-2020-0332.pdf by guest on 16 May 2026

Originality/value – This paper validated questionnaire with two biostatistical software programs in bidirectional approach – items difficulty and respondents' ability – is the first field test of WHO questionnaire among general population in Southeast Asia.

Keywords Validation of a questionnaire for Zika, Construct validation, Reliability

Paper type Research paper

Introduction

Zika infection has not attracted much attention in Malaysia despite being a country with an active and current circulation of the virus [1] and as the main contributor of Asian lineage of ZIKV [2]. Malaysia played an important role in the transmission of the *Zika* virus to Latin American regions in 2015 [3]. Underestimation is highly anticipated due to a suboptimal surveillance system [4].

Knowledge and perception can be used to predict the respondents' belief about their susceptibility toward the disease as well as the treatment of the disease according to the health belief model [5]. This maximizes the effort of prevention and control of Zika disease, especially in the design and strengthening of a capacity-building program such as the Community Behavior Intervention Program (COMBI) [6]. However, the majority of questionnaires were targeted at the clinical community and a bias for external validation of the community situation [7]. Sampling through social media or phone surveys may give rise to sampling bias, random bias, external validity and reliability issues [8]. The World Health Organization developed a questionnaire for communities with *Zika* virus transmission or those at risk such as Malaysia [9]. The resource pack provided general and thematic question banks in the domains of knowledge, attitudes and practices [9]. Users of this questionnaire can modify questions in the domains based on operational priorities [9]. The question style, language, answer mode and length can be modified for suitability of usage [9]. The questionnaire is available in various languages including English, Chinese, Portuguese, Arabic and Malay languages but no field validation was performed for knowledge/awareness toward *Zika* infection in South East Asia to the author's best knowledge [9].

This paper attempted to validate the construct validity and reliability of the English version of the WHO resource pack on knowledge, attitudes and practice surveys *Zika* virus disease and potential complications 2016 using a unique dual-approach validation model – the IBM SPSS and Rasch analysis. These two statistical approaches analyzed the data uniquely answering objectives, but the IBM SPSS was unable to measure underlying latent traits such as the respondents' ability, skills and language proficiency (item response theory) [10], which could be complemented by the Rasch, in the expression of item difficulty and individual ability and handling Likert scales category responses [11]. This formed the second unique feature of this paper – a bidirectional validation of a person's ability measured by scores achieved in the questionnaire and reversely validated the item ability of the questionnaire [12]. Third, this paper assessed both the construct validity and reliability of the questionnaire. With these three unique features, a solid validation was achieved to provide a valid and reliable questionnaire.

Methodology

A cross-sectional pilot study of 30 adult respondents in Malaysia completed the questionnaire in a single attempt, assisted by trained enumerators from June to August 2018 via Google Forms. Categorization of sociodemographic characteristics of respondents was based on WHO classification of "elderly" for 60-year-olds and above [13]. Age was divided into "adult" and "elderly" groups as the author hypothesized that the knowledge and perception levels were different between both groups based on their place of residence and activities [14]. Education level was based on the local school system, which was later collapsed into high and low levels of education for distinctive comparison of knowledge.

Downloaded from http://ftp.nowpublishers.com/jhr/article-pdf/36/6/1078/1488659/jhr-08-2020-0332.pdf by guest on 16 May 2026

Instrument

The questionnaire consisted of two domains – "knowledge" and "perception". The Knowledge domain had 14 questions adapted from section A of the General Question Bank of the WHO resource pack for *Zika* [9]. Items and answer option selections went through content validation with content experts in workshop discussion based on literature review and were made up of 63 items [9]. The letter "K" denoted "knowledge items", item KC-1 was on Information/communication, KS2-KS7 on cause/symptoms, KP8-KP9 on prevention, KT10-KT14 on treatment/care-seeking. The "perception" domain consisted of 6 items in a single construct denoted by the letter "P" and was developed by content experts. Perception of *vulnerability, transmission, complication, prevention, treatment, prevention* was in descending order of items (P1-6). This domain would be subjected to criterion validation through Likert category response in the Rasch [11].

Answer options for the "knowledge" domain were "Yes", "No" or "Don't know", in which only correct answers were given 1 mark, option "don't know" always carried no mark. Final scores were subjected to SPSS analysis to generate mean ± 2SD or median ± IQR. Scores were transformed into binary "1, 0" to generate logit data in model fitness statistics including item measure, person measure, mean ± 2SD score and reliability index.

The items in the "perception" domain were presented in a 5-point Likert scale where participants could answer "strongly agree", "agree", "neutral", "disagree" or "strongly disagree". Item P1, P2, P5 and P6 were tested on perception toward positive health-seeking behaviors, "strongly agree" was given a 5 score, "strongly disagree" was given 1 score. Items P3, P4 tested on perception toward a negative health-seeking behavior', scores were given in reverse order. Scores were inverted for Rasch analysis [11].

Statistical analysis

The author attempted a dual-approach of construct validation by marrying IBM SPSS and Rasch model analysis (Winsteps), which is rare but not the first of its kind [12]. Construct validity referred to the appropriateness of inferences made based on measurements to determine if a test measures the intended construct [15]. SPSS conventionally provides convergence and divergence of question items leading to construct formation, but was unable to measure the underlying ability and proficiency of respondents in the endorsement of responses to the question items tested [10]. The score respondents achieved may not accurately reflect their level of ability, as affected by how well-targeted the question items were including the language barrier and culture comprehensibility. The Rasch provided more comprehensive evidence on latent traits measured [12]. Besides, the Rasch was able to confirm the number of scales in measuring the difficulty in endorsing a response with the inference to a psychological attribute [16]. With this, the psychometric explanation was made clear and comparable with the range of item difficulty levels to the respondents' ability in answering the items [11].

Difficulty index and discriminatory index

Difficulty index was the proportion of participants who answered correctly, used to gauge the difficulty of questionnaire toward the respondents, calculated using the formula "D = (students with correct answer × 100)/total students" in excel [15] and later compliment by the Rasch model (In summary statistic) [10].

Discriminatory index explained how well items discriminate different levels of traits observed, calculated by:

$$R = (H - L)/27\% \text{ of Total Respondents } [10, 15]$$

Downloaded from http://ftp.nowpublishers.com/jhr/article-pdf/36/6/1078/1488659/jhr-08-2020-0332.pdf by guest on 16 May 2026

H = number of correct answers from the top 27% of respondents

L = number of correct answers from the bottom 27% of respondents

Discriminatory index of ≥0.4 was considered very good items, 0.09–0.19 were indiscriminative items [10, 15].

Item-correlation measures relationship between items or discrete variables with equally-spaced values (e.g. Likert scale) [12, 15]. Good item-correlation (r = 0.4–0.85) indicated model fitness, characterized by (1) Outfit MNSQ (outlier-sensitive fit) 0.5–1.5 logits, (2) Z-standard −2 to +2.15 Outfit MNSQ referred to mismatched ability to item difficulty or vice-versa [12]. Comparison of item correlation with "Corrected-Item-Correlation" using reliable analysis in SPSS complemented the test. Poor item-correlation (r < 0.32 in Rasch and ITC < 0.3) were items poorly correlating difficulty levels in measuring respondents' ability to answer questions [10, 11].

Reliability

Reliability analysis is the reproducible ability of the questionnaire items to obtain the same result in different populations and/or times and was used as a form of criterion validation in this study. The internal consistency of measurements with dichotomous choices was calculated using pKR20, while those with Likert scales were calculated using Cronbach's alpha, where a value of >0.6 is desirable for both pKR20 and Cronbach's alpha [11]. The Rasch provides reliability analysis to both items and persons and is complemented with the SPSS. Confusing and/or problematic questions were then modified or excluded prior to the initiation of the actual study.

The number of scales or the quality of scale category rating was examined by a few criteria, that is (1) a total of at least ≥10 respondents/category, (2) the rating frequencies of each category should be in regular distributions fashion (i.e. uniform, normal, bimodal or slightly skewed), (3) the average measures of rating should increase in uni-direction across the rating scale, (4) the total logit of adjacent threshold distance should be between 1.4 – 5. A logit >5 should suggest the merge of 2 adjacent scales, while logit <1.4 should suggest a split of scale into 2 or more adjacent categories, (d) a distinct probability curve displayed on each response category and (5) outfit value MNSQ <2 [11]. Distinctive peaks produced in probability curves (Rasch) indicated the number of rating suitable [11].

Construct validation – exploratory factor analysis (SPSS) and dimensionality (Rasch)

Principle component analysis aimed to reduce components to a smaller number by identifying principal components which explain the majority of variance [15]. Covariance with Varimax rotation was chosen to explore the 63 components (factors) of the knowledge domain and 6 factors in the perception domain. Criteria for principal components were 1) Eigenvalues above 1; 2) plots before the Scree plot levels off (Cattell rule); 3) when >80% variance explained [15]. Rasch item-dimensionality further analyzed the spread of variance and dimensionality of components by providing total variance % explained by measured trait and person' responses, besides items itself as a component [12]. Since the knowledge domain was built on 14 questions from four constructs in general question bank (A. Knowledge) WHO resource pack [9], expanded to 63 items by different answer options, the high possibility that a number of components are irreducible, as each option (e.g. sexual route transmission) was contributing significantly. In such a case, exploratory factor analysis (EFA) was used to confirm the validity of the 4 constructs [15].

Kaiser-Meyer-Olkin test of sampling adequacy and Bartlett's test of sphericity were used to determine the suitability of the data for factor analysis procedure [15]. However, the small sample size (<50) may be problematic for EFA, as generally, a 1:3 to 1:6 ratio per item should be available to extract principal components [17]. In the case of PCA not extractable, high

Downloaded from http://ftp.nowpublishers.com/jhr/article-pdf/36/6/1078/1488659/jhr-08-2020-0332.pdf by guest on 16 May 2026

communalities (>0.7–0.9) of components (square root of variance) may explain the significance of the factor loading of the particular component, hence considered valid to assess the underlying trait [15]. A variance of >80% would also be adequate to indicate the adequacy of such a number of components in assessing the trait studied [15]. Factor loading less than 0.35 was suppressed for analysis. High factor loading, high item-correlation, high communalities indicate strong data despite near singular component due to small sample size [17].

Model fitness statistic using Rasch analysis

Model fitness was assessed by SPSS's factors' dimensionality, Item-correlation, discriminatory index and complemented by a summary statistic of Rasch [11]. Rasch partial credit rating affirmed the fitness of the Likert rating scale [11]. The fitness of items to trait measured would be expressed by good variance %, communality, strong factor loadings, good item-correlation and good discriminatory index in SPSS [10, 15]. While Rasch complemented the fitness of person toward items administered, represented by Infit MNSQ = 1.0 (<-2 to <2.0) and Infit ZSTD near 0 (<-2 to <2.0) [11]. Model fitness was also complemented by person reliability measure, items targetability (Wright map), adequacy of variance spread (dimensionality and Wright map), identify individual erratic performance (Infit MNSQ >2SD) and individual item-correlation [16]. All items within mean±1SD of person side (Wright map) were considered as questions items well-targeted to the participants, items >+1SD were too hard, <−1SD was too easy for a person. Erratic items should be revised or removed [11, 12].

Ethical issue

This research obtained ethical approval from the University Kebangsaan Malaysia with UKM/PPI/111/8/JEP-2019-104, National Medical Research Registration (NMRR) with the approval code NMRR-18-3853-45130-IIR and the Department of Welfare of Aborigine with code JAKOA/PP.30.032Jld 45(75).

Results

Sociodemographic characteristics of respondents are summarized as in Table 1. A total of 30 respondents from all over Malaysia completed the questionnaire, with 0% drop out, 70% were female respondents, 93.3% were Chinese, 66.7% were educated to university level, all respondents at least attained secondary education. 2/3 respondents were residing at the forest fringe, 40% were working in the forest fringe. The household size was small, with no pregnant respondents.

Difficulty index versus person's ability – questions too hard or not understood?

The mean score for knowledge was 61.9 ± 24.76%; Rasch model analysis showed max mean measure for knowledge domain as log 18.6 (log 3.60 to -6.06), in descending order of ability. Questions that were most frequently missed were K4-6 "What causes *Zika*?" Breast milk (Answer: Yes); K5-3 "How does a person get *Zika*?" Through sexual intercourse (Answer: Yes); K13-3 "If a pregnant woman has *Zika*, what are the risks she faces?" She may have difficulty giving birth (Answer: Yes). All 3 questions were related to the sexual transmission of *Zika* infection, indicating a low level of knowledge toward this route of transmission for *Zika* virus, which is found uncommon among the *flavivirus* family.

The difficulty index of the knowledge domain of the questionnaire was 61.9 (61–80 = easy), complemented by Rasch to prove it as a set of well-targeted questions, 79.36% (50/63). The questions were not only easy for the Malaysian general public, but the items were easy to endorse with the local culture and language command. Beyond that, the Rasch Wright map showed

Downloaded from http://ftp.nowpublishers.com/jhr/article-pdf/36/6/1078/1488659/jhr-08-2020-0332.pdf by guest on 16 May 2026

Variables		N	Percentage %
Sociodemographic factors			
Age group (years)			
	Adult		
	18–59	27	90
	Elderly		
	60 and above	3	10
Mean age (years) 33.17 ± 12.99 (kurtosis, −0.080, skewness 0.890)			
Sex			
	Male	9	30
	Female	21	70
Ethnicity			
	Malay	0	0
	Chinese	28	93.3
	Indian	2	6.7
Educational level			
	Low		*0*
	No formal school		0
	Primary school		0
	High		*100*
	Secondary school	10	33.3
	Tertiary education	20	66.7
Residence area			
	Forest fringe	21	70
	City	9	30
Work area			
	Forest fringe	12	40
	City	18	60
Mean household number: 4.6 ± 1.673 (kurtosis 0.029, skewness 0.31)			
Mean household women at reproductive age (15-49 year old): 1.3 ± 1.393 (kurtosis 0.203, skewness 0.977)			
Median number of pregnant women in a household 0 ± 0.00 (kurtosis 23.77, skewness 4.782)			
Mean household monthly income (RM): 3000 + (25th centile 2000, 75th 6000) (kurtosis 2.62, skewness 8.958)			
Mean knowledge score: 39.0 ± 15.6/63 * 100% = 61.9 ± 24,76% (kurtosis, 1.318, skewness −1.418)			

Table 1. Sociodemographic characteristics of respondents

20.63% of the items <−1S.D. appeared too easy for the respondents. Participants could be separated into 3 groups: good/acceptable/poor knowledge based on the person separation index.

Difficulty index was not measured by SPSS for perception domain, but Rasch model affirmed 66.7% (4/6) of well-targeted question items, 33.3% of the items tested the ablest respondents. All respondents had good perceptions, in the singular person separation index.

Discriminatory index – how well were question items targeted?

The discriminatory index of the question items in the knowledge domain by SPSS was 0.34, indicating the items as “good items” to assess knowledge of the studied population on *Zika* infection. Rasch completed by showing items KS4-4, KS3-3, K3-5, KS2-1, KS5-7 and KT12 were items with poor correlation ($r < 0.32$) that caused indiscrimination of knowledge, similar to ITC (SPSS), KS4-2, KS4-6, KS5-4, KS6-6, KT10, KT12-3, KT13-3 ($r < 0.3$) (Table 2). A bubble chart identified two erratic items: KS4-2 and KT13-3. Item K4-2, “What causes *Zika*? Polluted water/ Dirty environment”, the answer should be “No”. Possible confusion may arise as pollution water means stagnant water bodies, which may encourage mosquito larvae breeding, leading to some attempts to answer “yes”. K13-3, “If a pregnant woman has *Zika*, what are the risks she faces?” with the option “she may have difficulty giving birth”. The answer to this question should be “Yes”. The general population may not understand about anencephaly/microcephaly may be

Downloaded from http://ftp.nowpublishers.com/jhr/article-pdf/36/6/1078/1488659/jhr-08-2020-0332.pdf by guest on 16 May 2026

Table 2.
Difficulty index and discrimination index of knowledge and perception toward *Zika*

Knowledge domain			
SPSS		Rasch	
Difficulty index (DI)	61.9 (easy)	Wright map	79.36% (50/63) items, well targeted 0% (0/63) items >+1SD, (too hard) 20.63% (13/63) <−1SD, (too easy)
Discriminatory index (Bi)	0.34 (good items)	Fit statistic	Good model fitness item measure 0.96 (1.00) Infit ZSTD −0.0 (closed to 0.00)
Perception domain			
SPSS		Rasch	
Difficulty index (DI)	Not tested (Likert scale)	Wright map	66.7% (4/6) items well targeted 33.3% (2/6) items >+1SD (too hard) 0% (0/6) <−1SD (too easy)
Discriminatory index (Bi)	1.025 (very good items) (Likert scale)	Fit statistic	Good model fitness item measure 0.98 (near 1.00) Infit ZSTD = −0.2 (near 0.00)

associated with polyhydramnios, unstable lie and hence unable to proceed with spontaneous vaginal delivery. Erratic questions should be revised, refined, or removed.

The perception domain had a DI of 1.025 and showed excellent discrimination. Person's ability ranged log 1.28 to–1.8. Two respondents (1.5%) show erratic performance (infit ZSTD 2.6). Only one perception item, P6 had a poor correlation with a person's ability in Rasch, $r = 0.24$ ($r < 0.32$), but three items were poorly correlated by SPSS (P1, P2, P6). Item P6, "I am confident that I can protect myself and my family from getting *Zika* infection", which "Strongly Agree" is given a 5 score. However, most respondents appeared not confident to protect themselves or family from getting *Zika* infection possibly due to fear of the disease, lack of knowledge of the disease and its prevention, over perception on the lethal nature of the disease likely due to sampling bias as researcher emphasis of the importance of the study to the studied population.

Rasch item separation index for knowledge domain was 6, items could be categorized into very easy, easy, moderate, above average, hard and very hard; and for perception, the domain was 4, which were easy, moderate, hard and very hard. Table 2 summarized the difficulty index and discriminatory index for both domains.

Perception domain response was assessed by the Likert scale. Rasch analysis enabled the Likert scale rating evaluation by checking on the probability curve for distinctive peaks (bond). Figure 1 shows the separation index yielded logit 2.72 ($1.4 < s < 5.0$), indicating a clear separation of 5 scales by distinctive intersections.

Model fit by Rasch model analysis

Overall, both domains showed good model fitness to assess knowledge and risk perception of the studied population toward Zika disease. All question items for *Zika* infection knowledge and perception assessment had good person measure (Bi) and good item measure (Di). A bubble chart identified 2 erratic items for the knowledge domain, K4-2 & K13-3, perfect fitness for perception domain. This analysis completed discriminatory analysis bi-directionally, both on the items and the persons' abilities.

Reliability

This questionnaire had very good reliability for the knowledge domain (pKR20 = 0.962) and fair reliability for the perception domain, Cronbach's alpha 0.641 with 4 items retained.

Downloaded from http://ftp.nowpublishers.com/jhr/article-pdf/36/6/1078/1488659/jhr-08-2020-0332.pdf by guest on 16 May 2026

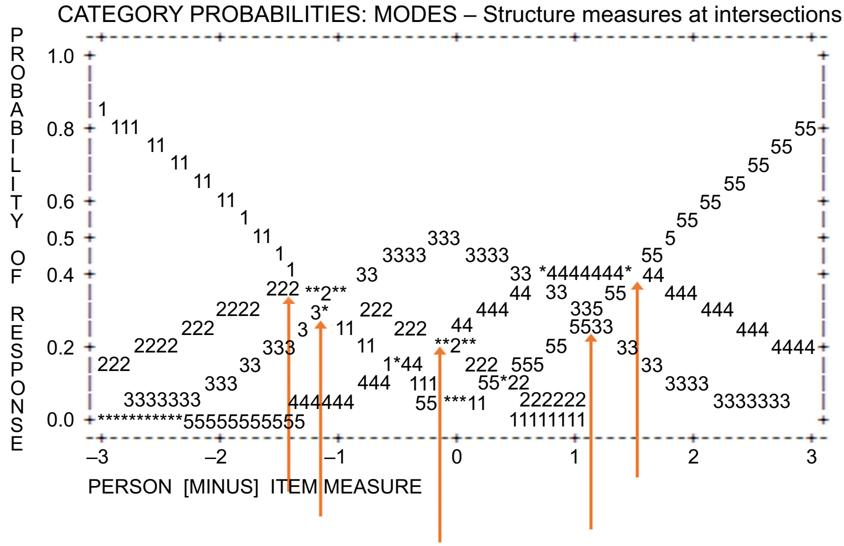

Figure 1.
Probability curves for Likert scale in perception domain

Cronbach's alpha was 0.418 for the initial 6 items in the perception domain. The value increased to 0.527 post removal of P2 and 0.641 with P2 & P6 removed. Question item P2 was "I will not get *Zika* infection if I practice safe sex, keep my environment clean and wear protective clothing" aimed to test on the perception of *Zika* disease transmission; while question item P6, "I am confident that I can protect myself and my family from getting *Zika* infection" aimed to test the perception of *Zika* disease prevention. In which an answer of "strongly agree" was given 5 points for both questions, and "strongly disagree" was given one point. The number of factor recovery was close to the expected constructs if all two problematic items were removed. Discussion with content experts on the revision of sentences and the number of items in the perception domain is needed to maintain the validity and reliability of the questionnaire. Table 3 summarizes the EFA and item analysis for both domains.

Construct validation

The Kaiser-Meyer-Olkin measure of sampling adequacy and Bartlett test of sphericity could not be generated and indicated inadequacy of sampling, most probably due to the small sample size or high multicollinearity issues. Component transformation matrix extracted four components which explained 54.66% of the variance, while 11 components were needed to achieve 81% variance (the above 80% rule) [15]. Scree plot levels off at 4 components, represented principal components, despite there being 15 components above Eigenvalue of 1 as described in Figure 2. A small sample size (<50) often gives rise to problematic near singular covariance matrix [17]. However, high commonality is seen in all items in knowledge (all above 0.7), indicating all components had significant loadings [15]. The standard error of measurement was 3.03 (<5), indicating that the finding is good (significant), Table 4 [15]. This offset the low variance with good psychometric properties [17]. This is most likely because the question items were all measuring the same trait – knowledge, while each item represented an option of the answer to the root question [18]. Therefore, all knowledge items should remain in one construct, four principal components explained nearly half of the variance.

Rasch item dimensionality affirmed that 38.7% of raw variance were explained by measure, 18.2% by items. (total up to 56.9%), while a person's ability to answer the question

Downloaded from http://ftp.nowpublishers.com/jhr/article-pdf/36/6/1078/1488659/jhr-08-2020-0332.pdf by guest on 16 May 2026

Items	Mean (S.D)	Factor 1 Information/ communication	Factor 2 Cause/ symptom	Factor 3 prevention	Factor 4 Treatment/ care-seeking	ITC
KC1	38.37	0.959				0.609
KS2	38.33		0.960			0.518
KS3-1	38.37		0.959			0.623
KS3-2	38.20		0.959			0.578
KS3-3	38.13		0.959			0.736
KS3-4	38.30		0.959			0.708
KS3-5	38.23		0.959			0.698
KS3-6	38.07		0.959			0.730
KS3-7	38.17		0.959			0.738
KS3-8	38.30		0.959			0.665
KS4-1	38.13		0.960			0.555
KS4-2	38.47		0.961			0.009
KS4-3	38.63		0.960			0.372
KS4-4	38.53		0.960			0.546
KS4-5	38.17		0.959			0.575
KS4-6	38.50		0.961			0.269
KS4-7	38.10		0.960			0.443
KS5-1	38.53		0.959			0.705
KS5-2	38.50		0.960			0.487
KS5-3	38.57		0.960			0.434
KS5-4	38.23		0.961			0.247
KS5-5	38.57		.960			0.428
KS5-6	38.57		0.960			0.302
KS5-7	38.27		0.960			0.434
KS5-8	38.57		0.960			0.509
KS5-9	38.37		0.960			0.516
KS5-10	38.13		0.960			0.494
KS6-1	38.23		0.959			0.736
KS6-2	38.27		0.959			0.601
KS6-3	38.30		0.959			0.666
KS6-4	38.40		0.959			0.564
KS6-5	38.70		0.960			0.469
KS6-6	38.47		0.961			0.240
KS6-7	38.70		0.960			0.545
KS7	38.13		0.960			0.310
KP8	38.10			0.959		0.742
KP9-1	38.10			0.959		0.705
KP9-2	38.23			0.959		0.705
KP9-3	38.47			0.960		0.489
KP9-4	38.20			0.960		0.411
KP9-5	38.17			0.960		0.546
KP9-6	38.33			0.959		0.569
KP9-7	38.73			0.960		0.387
KT10	38.37				0.961	0.066
KT11-1	38.33				0.959	0.725
KT11-2	38.37				0.959	0.621
KT11-3	38.23				0.959	0.665
KT11-4	38.13				0.959	0.713
KT12-1	38.17				0.959	0.711
KT12-2	38.73				0.959	0.614
KT12-3	38.67				0.961	0.222
KT12-4	38.33				0.960	0.389

(*continued*)

Table 3.
Exploratory factor analysis (knowledge and perception of *Zika*)

Downloaded from http://ftp.nowpublishers.com/jhr/article-pdf/36/6/1078/1488659/jhr-08-2020-0332.pdf by guest on 16 May 2026

Items	Mean (S.D)	Factor 1 Information/ communication	Factor 2 Cause/ symptom	Factor 3 prevention	Factor 4 Treatment/ care-seeking		ITC	
KT12-5	38.47				0.959		0.569	
KT12-6	38.30				0.960		0.545	
KT13-1	38.23				0.960		0.435	
KT13-2	38.50				0.960		0.463	
KT13-3	38.23				0.961		0.110	
KT14-1	38.27				0.959		0.601	
KT14-2	38.33				0.959		0.745	
KT14-3	38.33				0.959		0.564	
KT14-4	38.27				0.960		0.550	
KT14-5	38.27				0.959		0.686	
KT14-6	38.10				0.959		0.627	
Cronbach's alpha	*0.960*	–	*0.926*	*0.838*	*0.904*			
P1	13.17	6.63			0.341	0.252	0.254	0.685
P2	12.47	–			0.527	–	−0.041	–
P3	13.77	7.23			0.252	0.521	0.421	0.514
P4	13.73	7.20			0.263	0.535	0.348	0.482
P5	13.37	6.83			0.272	0.415	0.371	0.578
P6	12.67	–			0.500	–	−0.053	–
Cronbach's alpha					*0.418*	*0.641*		

Table 3.

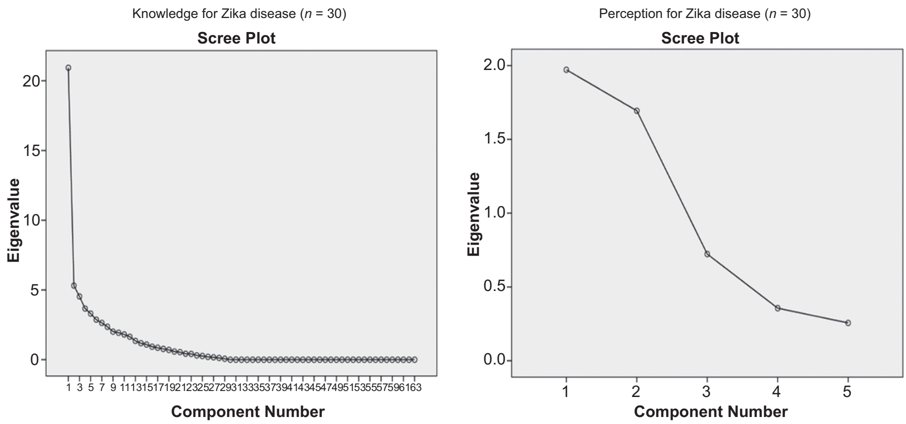

Figure 2. Scree plots for knowledge & perception domains

(latent trait) explained 20.5% of the variance. Rasch analysis calculated construct separation by item separation index (G) using the formula:

$$G_{\text{item}} = \text{true SD}/\text{Real RMSE} = 1.46/0.37 = 3.95 = 4$$

Indicating four constructs were found to discriminate items at different levels of the trait being measured, in this case, "knowledge". The result compliments EFA in SPSS for construct validation.

For the perception domain, the Kaiser-Meyer-Olkin measure of sampling adequacy was 0.526, Bartlett's test of sphericity was significant at $p < 0.0001$, and the data were suitable

Downloaded from http://ftp.nowpublishers.com/jhr/article-pdf/36/6/1078/1488659/jhr-08-2020-0332.pdf by guest on 16 May 2026

Summary statistic	Knowledge	Perception
Cronbach's α (Rasch)	0.960	0.530
Cronbach's α (SPSS)	0.960	0.642
Item reliability (Rasch)	0.80	0.89
Item separation (Rasch)	4	3
Person reliability (Rasch)	0.960	0.50
Person separation (Rasch)	2	1
Standard error (Rasch)	0.36	0.27
Dimensionality		
Measured variance, % (Rasch)	38.7%	39.5%
Measured variance, % (SPSS)	54.5%	64.3%
Eigenvalue (Rasch)	36.9	4
Eigenvalue (SPSS)	22.0	2
Meaningful pattern of residuals (Rasch)	No	No
Rating scale (Rasch)		
≥10 responses per category	No	Yes
Regular distributions	N/A	Yes
Distance adjacent threshold	N/A	Yes
Outfit MNSQ <2	Yes	Yes
Probability curves	2 distinctive peaks	5 distinctive peaks
Number of construct (SPSS/Rasch)	4/4	2/3
Number of level trait measured (person's ability) (SPSS/Rasch)	3/3	–/1
Number of constructs for item difficulty (Rasch)	6	4
Number of intersection	–	5 ratings Likert scale

Table 4.
Construct validity for knowledge and perception domains using SPSS and Rasch analysis

for factor analysis procedures despite the barely adequate sample size [17]. Communality was good, at 0.6–0.86, two components have Eigenvalues greater than 1, initial six-factor solution accounted for 64.3% variance. Total variance improved to 73.3% after problematic item P2 was removed. Pattern matrix showed strong factor loadings for all items into two principal components.

Rasch item dimensionality affirmed that 39.5% of raw variance were explained by measure, 29.6% by items; (total up to 56.9%), while a person's ability to answer the question (latent trait) explained 9.9% of the variance. Rasch analysis calculated construct separation by item separation index (G) using the formula:

$$G_{\text{item}} = \text{True SD}/\text{Real RMSE} = 0.577/0.22 = 2.59 = 3$$

Separating perception items to three constructs. The result compliments EFA in SPSS for construct validation (before rounded to nearest number, observe also scree plot pattern).

Discussion

This study is a rare duet that assessed construct validity and reliability in a complementary and completing way bidirectionally [10, 12]. Validation with dual approaches provides extra value to validation of a questionnaire by the assessment of latent traits such as language barriers and subjects' cultural understanding toward a questionnaire design [12]. The questionnaire is suitable to be used in a large-scale study to assess the knowledge and perception of general communities in other developing countries in Southeast Asia with similar cultural understanding and language commands, to help governments tailor the capacity-building program in Zika disease prevention. The small sample size may serve as a limitation to factor analysis but does not leave researchers with no solution [17]. Beyond the rules-of-thumb of EFA ratio 5:1, a strong and uniform set of commonalities, the number of

Downloaded from http://ftp.nowpublishers.com/jhr/article-pdf/36/6/1078/1488659/jhr-08-2020-0332.pdf by guest on 16 May 2026

items/factors and the item factor loading strengths all contribute to strong data [18] which produces reliable, reproducible, stable factor solution [19].

The questionnaire has good reliability and validity, good difficulty index & discriminatory index. The fair reliability of the perception domain is most probably related to inadequate health information received for *Zika* among the Malaysian community [15]. Erratic items KT13-3 & KS4-2 was possibly due to lack of clarity of phrases, wrong inference on the subject matter, or not testing the level of the respondents [15]. Respondents have good knowledge of *Zika*, similar to previous studies [7]. Nonetheless, the breakdown of construct analysis showed ineffective risk communication of *Zika* disease prevention among the population. There is a possibility of knowledge (on cause, symptom, prevention) inference toward *dengue* fever which presented similarly to *Zika* (KS2-KS7, KP8-KP9). Respondents failed to identify sexual transmission as a possible route (KS4-6, KS5-3, KS5-6, KS5-7, KS5-9). The vector prevention and control approach should be strengthened. Lack of knowledge on sexual and blood transmissions prevention potentially leads to devastating ZIKV endemicity [20]. However, positive health-seeking behaviors and good doctor-client relationship were found to be established (KT10-KT14), clients trusted the modality of treatment given by the ministry of health. This is important as the ZIKV disease monitoring period is much longer than that of *dengue* fever [21].

More than half of the respondents think that Malaysia is at risk of ZIKV, in contrast to research published [7] shortly after the Public Health Emergency of International Concern for *Zika* was declared in Malaysia. Most respondents had a good perception of *Zika* infection prevention and positive health-seeking behavior (P1-P4). Respondents' perceived risk of infection at forest fringe, thereby expressed keenness for self-protection. However, respondents expressed poor risk perception toward the threat & severity imposed by *Zika* infection (P5-P6), reflected respondents had not received adequate information on disease prevention from the local health authority, possibly overshadowed by the more common dengue fever and falsely assured by the undetected cases [4]. WHO risk communication guidelines should be implemented effectively to increase public resilience and participation in the collaborative effort to control the transmission of *Zika* [21].

Conclusion

The field-validated English version *Zika* infection awareness/knowledge questionnaire had good psychometric properties to measure the knowledge and perception of *Zika* infection among the Malaysian community. Respondents had good knowledge and risk perception toward *Zika*. The questionnaire could aid the prevention and control of Zika disease in developing countries in Southeast Asia by designing a behavioral modification program and ensure effective risk communication. Further development should focus on the items revised for the Malaysian forest fringe population who are at higher risk of facing sylvatic to urban cycle transmission of *Zika* infection and should be tested on a larger scale.

Conflict of Interest: None

References

1. Woon YL, Lim MF, Tg Abd Rashid TR, Thayan R, Chidambaram SK, Syed Abdul Rahim SS, *et al.* Zika virus infection in Malaysia: an epidemiological, clinical and virological analysis. BMC Infect Dis. 2019; 19(1): 152. doi: 10.1186/s12879-019-3786-9.

2. Marchette NJ, Garcia R, Rudnick A. Isolation of Zika virus from Aedes aegypti mosquitoes in Malaysia. Am J Trop Med Hyg. 1969; 18(3): 411-5. doi: 10.4269/ajtmh.1969.18.411.

3. Duong V, Dussart P, Buchy P. Zika virus in Asia. Int J Infect Dis. 2017; 54: 121-8. doi: 10.1016/j.ijid.2016.11.420.

4. Ministry of Health Malaysia. Kenyataan Akhbar KPK Minggu ke-28 2020 – Situasi Semasa Zika di Malaysia. [updated 2020 Jul 25; cited 2020 Sep 30]. Available from: http://www.myhealth.gov.my/en/situasi-terkini-demam-denggi-dan-zika-di-malaysia/.

Downloaded from http://ftp.nowpublishers.com/jhr/article-pdf/36/6/1078/1488659/jhr-08-2020-0332.pdf by guest on 16 May 2026

5. Vadhariya A, Sansgiry SS. Use of health belief model to understand knowledge, attitudes and behaviors of people towards the ebola outbreak. Value Health. 2015; 18(3): A241. doi: 10.1016/j.jval.2015.03.1405.
6. Ministry of Health Malaysia. COMBI. MyInfoSihat. [cited 2019 Sep 19]. Available from: https://www.infosihat.gov.my/index.php/projek-khas/combi.
7. Arief M, Hassali MAA, Saleem F, Khan MU, Ahmad A, Bhagavathulha AS, *et al.* A cross-sectional survey on the knowledge and attitudes towards Zika virus and its prevention among residents of Selangor, Malaysia. Journal of Pharmacy Practice and Community Medicine. 2017; 3(2): 81-9. doi: 10.5530/jppcm.2017.2.20.
8. Samuel G, DiBartolo-Cordovano R, Taj I, Merriam A, Lopez JM, Torres C, *et al.* A survey of the knowledge, attitudes and practices on Zika virus in New York City. BMC Publ. Health. 2018; 18(1): 98. doi: 10.1186/s12889-017-4991-3.
9. World Health Organization [WHO]. Knowledge, attitudes and practice surveys: Zika virus disease and potential complications: resource pack. Geneva: WHO; 2016.
10. IBM. Item response theory/Rasch Models in SPSS statistics. [cited 2019 Sep 10]. Available from: http://www-01.ibm.com/support/docview.wss?uid=swg21488442.
11. Bond TG, Fox CM. Applying the Rasch model: fundamental measurement in the human sciences. 2nd ed. New York, NY: Routledge; 2012.
12. Rosnah I, Mohd Zali MN, Noor Hassim I, Azmi MT. A tale of two construct validation analysis: Rasch model and exploratory factor analysis approach for Three-Factor Eating Questionnaire (TFEQ-R21) among Malaysian male workers. Med J Malaysia. 2015; 70(3): 169-76.
13. Department Government Service Malaysia. Health - what is the definition of senior citizen in Malaysia. [cited 2020 Oct 9]. Available from: http://www.jpapencen.gov.my/english/senior_citizen.html.
14. Department of Statistics Malaysia. Five year internal migration in Malaysia. [cited 2020 Aug 30]. Available from: https://www.dosm.gov.my/v1/index.php?r=column/cdatavisualization&menu_id=a2ZOQ084R0ZEK2VrWjArRkxndFE3Zz09&bul_id=RGZWOFl2eWREa2lWa2FhMkdvdWxqUT09.
15. Chan YH. Biostatistics 302. Principal component and factor analysis. Singapore Med J. 2004; 45(12): 558-65, quiz 66.
16. Christensen KB, Engelhard G Jr, Salzberger T. Ask the experts: Rasch vs. factor analysis. Rasch Measurement Transactions. 2012; 26(3): 1373-8.
17. Jung S, Lee S. Exploratory factor analysis for small samples. Behav Res Methods. 2011; 43(3): 701-9. doi: 10.3758/s13428-011-0077-9.
18. Costello AB, Osborne J. Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. Practical Assess. Res. Eval.. 2005; 10: 1-9. doi: 10.7275/jyj1-4868.
19. Wang LL, Watts AS, Anderson RA, Little TD. Common fallacies in quantitative research methodology. In: Little TD, (Ed.) The Oxford handbook of quantitative methods in psychology: Vol. 2 Statistical analysis. New York, NY: Oxford University Press; 2013. 718-58.
20. Goodnough LT, Marques MB. Zika virus and patient blood management. Anesth Analg. 2017; 124(1): 282-9. doi: 10.1213/ANE.0000000000001770.
21. Pan American Health Organization [PAHO]; World Health Organization [WHO]. Zika virus infection: step by step guide on risk communication and community engagement. Washington, DC: PAHO/WHO; 2016.

Corresponding author
Mohd Rohaizat Hassan can be contacted at: rohaizat@ppukm.ukm.edu.my

For instructions on how to order reprints of this article, please visit our website:
www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: **permissions@emeraldinsight.com**

Downloaded from http://ftp.nowpublishers.com/jhr/article-pdf/36/6/1078/1488659/jhr-08-2020-0332.pdf by guest on 16 May 2026