Psychometric item analysis and validation of the Indonesian version of the Readiness for Interprofessional Learning Scale (RIPLS)

Dwi Tyastuti¹, Hirotaka Onishi¹, Fika Ekayanti² and Kiyoshi Kitamura¹

¹Faculty of Medicine, International Research Center for Medical Education (IRCME), The University of Tokyo, Bunkyo-ku, Tokyo, Japan and ²Faculty of Medicine and Health Sciences, State Islamic University Syarif Hidayatullah Jakarta, Indonesia, Jakarta, Indonesia

Abstract

Complex health care needs in developing countries are stimulating development and implementation of interprofessional education (IPE). To better understand IPE, it is necessary to develop and evaluate an educational program that focuses on interprofessional learning (IPL) in Indonesia. However, no instrument in the Indonesian language has been developed to measure attitudes toward IPL. The aim of this study is to describe the process of a cross-cultural adaptation of the Readiness for Interprofessional Learning Scale (RIPLS) in an Indonesian version including determining its reliability and validity. The study was conducted among students enrolled in medical, nursing, pharmacy and public health courses at the State Islamic University, Jakarta, Indonesia, in 2012. The completed responses to RIPLS were collected from 755 students. The psychometric properties were analyzed by both exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). EFA on 18-items revealed three factors accounting for 59.9% of the total variance. CFA resulted in a three-factor model over 16 items with satisfactory reliability (alpha coefficients >0.7), construct validity and acceptable indices of goodness of fit. We conclude that this Indonesian version of RIPLS with a three-factor model over 16 items is a valid tool to measure students’ attitudes toward IPL.

Introduction

Currently in developing countries, the health sector faces several issues such as the double-burden diseases, a shortage of health professionals and low quality of health services (WHO, 2010). Moreover, globalization has caused increasingly complex healthcare needs (Loxley, 1997; Mickan, Hoffman, & Nasmith, 2010). Those conditions have stimulated the development and implementation of interprofessional education (IPE) and collaborative practices in many areas worldwide (WHO, 2010).

It is important to develop and apply IPE in Indonesia in accordance with the movement launched by WHO, a strategy of Framework for Action on Interprofessional Education and Collaborative Practice (WHO, 2010). It is part of WHO’s longstanding commitment to improve health professional education through IPE (Barr, 2010).

To understand interprofessional learning (IPL) and the best model for collaborative practice among health providers in Indonesia, it is necessary to develop an initial program that focuses on IPL in community-based settings. The reasons are as follows: primary healthcare is at the forefront of health services in Indonesia; home care services have become necessary in Indonesia due to increased life expectancy and the prevalence of chronic diseases; students can learn about patient care management in integrated and holistic approaches; and students can work with a large number of other health professions in the community such as paraprofessionals (educators and lawyers), community volunteers and indigenous workers (Islamic scholars, traditional birth attendants, non-registered nurses, etc.) who have unique roles in supporting the health programs in Indonesia.

One of the indicators of success of an IPL program is high scores on measures of students’ knowledge, skills and attitudes. We presume that the assessment of students’ attitude is the most important element in promoting IPL in Indonesia since attitudes are the biggest barriers to overcome (Gilbert, 2005; Parsell & Bligh, 1999). Many scales have been constructed to measure various aspects of attitude in IPL since IPE was initially introduced 30 years ago. A scale used to evaluate students’ attitude toward IPE with a good level of validity (Freeth, Reeves, Hammick, Koppel, & Barr, 2005) is the Readiness for Interprofessional Learning Scale (RIPLS) published by Parsell and Bligh in 1999. It is an important instrument for evaluating interprofessional activities (Reid, Bruce, Allstaff, & McLernon, 2006).

The RIPLS is preferred to other existing measures of IPE since the scale has been examined and reported in the peer-reviewed literatures (Thannhauser, Russell-Mayhew, & Scott, 2010), thus the scale is used widely in the IPE literatures (Williams, Brown, & Boyle, 2012) and has been adapted for use in different languages such as in Swedish and Japanese (Lauffs et al., 2008; Tamura, Seki, Usami, Taku, & Bontje, 2012). The Indonesian version of RIPLS is needed as an instrument to evaluate students’
attitudes toward IPL in various educational programs in Indonesia. To our knowledge, there is no information on the RIPLS scale in Indonesia.

Parsell and Bligh performed validation tests on the RIPLS scale in 1998 (n = 914 respondents) and in 1999 (n = 120 respondents), and both studies yielded different factors (Parsell & Bligh, 1999; Parsell, Stewart, & Bligh, 1998). The first study (1998) generated a RIPLS scale that consisted of 19 items and was grouped into two main constructs: teamwork and collaboration (TWC) and professional identity (PI) (Parsell et al., 1998). The second study (1999) yielded three subscales including a subscale of TWC (item 1–9), PI (item 10–16) as well as roles and responsibilities (RR; items 17–19) (Parsell & Bligh, 1999). The study by Lauffs et al. (2008) generated constructs similar to Parsell & Bligh’s study (1999) and other studies yielded different constructs (McFadyen, Webster, & Maclaren, 2006; McFadyen et al., 2005; Tamura et al., 2012). El-Zubeir, Rizky, and Al-Khalil (2006) and Reid et al. (2006) modified RIPLS to adjust for existing conditions, and this modified RIPLS yielded a satisfactory reliability. The validation of the RIPLS Japanese version generated a three-factor model, but the items that constructed these factors were different from other models (Tamura et al., 2012).

The aim of the study presented in this article is to describe the psychometric properties, reliability and validity of the Indonesian version of the RIPLS scale, adapted from the original scale through a cross-cultural process, in order to be able to assess the readiness of undergraduate students to engage in interactive and collaborative learning during community-based health care.

Methods

Cross-cultural adaptation process

Translation into the Indonesian language and cross-cultural adaptation of the original version of the RIPLS was conducted according to guidelines (Guillemin, Bombardier, & Beaton, 1993). The process of adaptation started with forward and backward translations (van de Vijver & Hambleton, 1996). The first phase was the translation of the questionnaire (forward translation) by two language experts. The second phase aimed to assess the consistency between the original version and the translated version. The scale in the Indonesian version was retranslated into English (backward translation) by different language experts to check for inconsistencies.

Participants

The main study was carried out at the Faculty of Medicine and Health Sciences (FMHS), State Islamic University, Syarif Hidayatullah Jakarta, Indonesia, in February 2012. First to third-year students from four courses (medicine, nursing, pharmacy, and public health) participated in the study. We distributed 800 questionnaires. Of the 776 questionnaires returned by participants, 21 (2.8%) questionnaires were invalid and excluded. The sample size of 755 met the minimum criteria (Hair, 2005; Kline, 2005).

Table 1. Summarize of the internal consistency of the RIPL scale from previous studies.

| Study                      | Number of samples | Number of items | Total | TWC | PI  | RR  |
|----------------------------|-------------------|-----------------|-------|-----|-----|-----|
| Parsell, Stewart, & Bligh (1998) | 914               | 19              | 0.85  | 0.46|     |     |
| Parsell & Bligh (1999)      | 120               | 19              | 0.90  | 0.88| 0.63| 0.32|
| McFadyen (2005), data 2003a| 308               | 19              | 0.84  | 0.79|     |     |
| McFadyen (2005), data 2004a| 247               | 19              | 0.89  | 0.88|     |     |
| McFadyen (2006)a            | 65                | 19              | 0.89  | 0.71|     |     |
| Lauffs et al. (2008)        | 214               | 19              | 0.89  | 0.89| 0.63| 0.48|
| Reid et al. (2006)b         | 66                | 23              | 0.76  |     |     |     |
| El-Zubeir et al. (2006)b    | 178               | 20              | 0.61  |     |     |     |
| Tamura et al. (2012)b       | 132               | 19              | 0.74  |     |     |     |

*McFadyen’s studies produced a four-factor model.

*bAll of these studies yielded three factors, but the items that made up these factors were different from other studies.

NPI = negative professional identity; PPI = positive professional identity.

Instrument

We only translated and used 18 items from the original scale in this study. We omitted item 17 (the function of nurses and therapists is mainly to provide support for doctors). The reasons for omitting this question are as follows: (1) there are no therapist students who participated in this study and (2) the term ‘‘therapist’’ is used ambiguously in Indonesian society, since the term is used in fields other than the medical one.

Item answers were given on five-point Likert scales (1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree and 5 = strongly agree) with higher scores indicating more positive attitudes. Table 1 described the reliability coefficient of the RIPLS from previous studies in both the original version and modified versions.

Data collection

Data were collected by distributing the questionnaire in classrooms over a two-week period. A faculty member made a schedule to set the day, time and length of time required to fill out the questionnaire. Faculty members who were not involved in teaching activities assisted in the distribution and collection of questionnaires.

Data analysis

To ensure validity and to determine the best factor structure of the RIPLS scale in the Indonesian version, both exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) were conducted. EFA was run to identify latent variables or factors of a set of variables (Harrington, 2009), whereas CFA was performed to assess the fit of a model obtained from the EFA result or a previously existing theoretical model and to confirm convergent and discriminant validity of a fit model (Hair, 2005; Harrington, 2009).

The total samples (n = 755) were randomly divided into two split-half samples by SPSS, version 16.0 (SPSS Inc., Chicago, IL), one for EFA (n = 377) and one for CFA (n = 378). Data from the sub-sample (n = 377) were analyzed by EFA on 18 items of the Indonesian version with maximum likelihood methods and oblique rotation. Acceptance of EFA was checked with three
well-recognized criteria for EFA analysis, the Kaiser–Meyer–Olkin (KMO), Bartlett’s test of sphericity and the diagonals of the anti-image correlation matrix. Two criteria were set up to retain factors produced by EFA, i.e. eigenvalues >1 and scree plot evidence (DeVellis, 2012).

The second sub-sample \((n = 378)\) was used for the analysis of factor structure using CFA. In this study, the CFA procedure with maximum likelihood estimates (Hair, 2005), and the goodness-of-fit (GOF) of the model was evaluated by multiple criteria (Hair, 2005; Harrington, 2009). To indicate a fit model, the criteria for GOF include the Normed Fit Index (NFI), the Goodness of Fit Index (GFI), the Tucker–Lewis fit index (TLI) and the Comparative Fit Index (CFI) all of which are acceptable if \(>0.90\). The Standardized Root Mean Square and the Root Mean Square Error of Approximation (RMSEA) are also acceptable if \(<0.08\) (Browne & Cudeck, 1993; Byrne, 2001; Schumacker & Lomax, 2004). Reliability of the Indonesian version was assessed by internal consistency, item reliability and composite reliability (CR) (Fornell & Larcker, 1981; Hair, 2005). The validity of the scale was tested by calculating construct and discriminant validity. All analyses were conducted using SPSS version 16.0 and AMOS version 18.

Ethical considerations
All students gave written and oral informed consent after receiving both written and oral information about this study. The Ethics Committees of both the University of Tokyo and the State Islamic University, Syarif Hidayatullah Jakarta, Indonesia, approved the research. The University of Tokyo collaborates with the State Islamic University to improve the quality of teaching staff at FMHS.

Results
Descriptive and statistical assumptions
Among 800 questionnaires distributed, 776 questionnaires were collected. A total of 21 questionnaires were excluded because (1) eight questionnaires had missing items, and the result of the missing value analysis showed that the values of missing completely at random were \(>0.05\); and (2) 13 questionnaires showed outliers (Z-score values of less than \(-3\) or more than \(+3\)) (Kline, 2005). Valid questionnaires were 755 and the response rate was 94.4%. The proportion of female student \((n = 557, 73.8\%)\) was higher than male students. Students who participated in this study are from public health (33.6%), medicine (30.5%), pharmacy (29.8%) and nursing (61.1%) courses. Participants consisted of first-year students (45.8%), second-year students (28.7%) and third-year students (25.5%).

Factor analysis
EFA on 18 items was accepted because three measures met the criteria. The KMO (0.905) and Bartlett’s test of sphericity \((p < 0.001)\) indicated a factorable correlation matrix. Third, the diagonals of the anti-image correlation matrix for all items were between 0.799 and 0.944, above the standard of 0.5 (Hair, 2005).

Initially, EFA was performed by using maximum likelihood and promax rotation, and we found the cross loading of some items on some factors. Subsequently, we performed direct oblimin rotation, which produced factors that were free of cross loading (Table 2). The EFA showed 59.9% of the total variance including 37.3% for the first factor (F1), 10.8% from the second factor (F2), 6.1% from the third factor (F3) and 5.7% from the fourth factor (F4). The composition of items for each sub-scale in the Indonesian version was unlike the previous studies. The results showed that F1 covered items 1–6, F2 did items 10–12 and 18, F3 did items 13–16 and 19 and F4 covered items 7–9. The factor loadings for item 19 was 0.255, whereas the other 16 items were more than 0.4 (between 0.401 and 0.905).

Measurement models
The CFA was conducted for the Indonesian version to examine reliability and validity of the scale and generate a fit model. The maximum likelihood method of estimation was chosen since the normality assumption was not violated (Kline, 2005).

Initial model
The initial model was a measurement model derived from the results of the EFA. The four-factor model with 18 items (F1: items 1–6; F2: item 7–9; F3: items 10–12 and 18; F4: items 13–16 and 19) showed that the t-value of all items were significant by the criterion of \(p < 0.05\). However, a Heywood case was found with item 19 [communality > 1] (Hair, 2005; Kline, 2005). The loading factors of items 1–16 were more than 0.5, but item 18 and 19 had loadings of 0.433 and −0.133, respectively, suggesting that items 18 and 19 were not represented in the measured constructs (Figure 1).

The initial model was not a fit model because items 18 and 19 had loading factors of less than 0.5, and there was a Heywood case on item 19 (Hair, 2005; Kline, 2005). Hence, we eliminated items 18 and 19 from the model and reconstructed a 16 item model for the next analysis by using previously existing theoretical models.

Three-factor version
We conducted CFA on several models to find the fit model and the three-factor model with 16 items was a satisfactory model for this study. The three-factor model with 16 items showed that factor loadings for all items ranged from 0.55 to 0.86. Although Hair (2005) suggested that the best factor loading was more than 0.7 (Hair, 2005), factor loading more than 0.5 was also satisfactory, especially in explorative research (Johnson & Stevens, 2001) (Figure 2).

Therefore, we retained all items (16 items) for model modification and further analysis. Furthermore, a three-factor model with 16 items yielded an acceptable Normed chi-square of 2.485, and all other GOF indices met criteria such as GFI (0.92), NFI (0.91), TLI (0.93), CFI (0.94), Standardized Root Mean Square Residual (SRMR) (0.043) and RMSEA (0.063) (Table 3).

Assessment of internal structure of a model
Reliability analysis showed that the overall alpha coefficient of the initial model (0.69) was lower than modified models (0.87). Table 3 shows that three alpha coefficients of the modified model were more than 0.70.

Furthermore, average variance extracted (AVE) and CR were calculated to check construct validity of the measurement model. CR values for each construct of the modified model were satisfied \((>0.7)\) (Fornell & Larcker, 1981). AVE of each factor is expected to be \(>0.5\). AVE for two factors, negative professional identity (NPI) and positive professional identity (PPI), met the criterion and AVE for TWC factor was close to the criterion (see Table 3). Thus, the modified version had good construct validity.

Discriminant validity upon a scale was tested to check the confidence in subsequent research findings (Fornell & Larcker, 1981) to confirm unidimensionality of each construct. Discriminant validity is demonstrated if the AVEs of both constructs are greater than the squared correlation (Chau, 1997; Fornell & Larcker, 1981). This study used this method to determine whether the constructs in the three-factor model were
different from the others. Focusing on Table 4, the results shows questionable discriminant validity for constructs of TWC since the AVE of TWC (0.46) was lower than the square correlation between TWC and PPI (0.59). Only the constructs of NPI and PPI indicated acceptable discriminant validity because the AVE was greater than the squared correlation. The results demonstrated that NPI and PPI were unidimensional.

**Discussion**

In this study, the original RIPLS was translated and adapted to an Indonesian setting, and the Indonesian version proved to be a model with three constructs on 16 items by EFA and CFA. It was also demonstrated to be a reliable and valid scale. This final model is as McFadyen et al. (2005) without the RR sub-scale.

EFA on the Indonesian version was run for 18 items by using direct oblimin, and this produced four-factor model. The total variance of this study was higher than the previous study (McFadyen et al., 2005; Parsell & Bligh, 1999; Parsell et al., 1998). However, this study only used 18 of 19 items from the original study, and the sample study was first to third year students. It was different from previous studies.

Through CFA processes, the four-factor model of EFA result was tested and did not offer the most desirable fit to our data (see Table 3). Reconstructing the measurement model by omitting items 18 and 19, due to low and negative loading factors, generated a three-factor model, which can be labeled as TWC, PPI and NPI, because these factors have the same formation as McFadyen’s study (2005, 2006). Referring to criteria of the fitted CFA model (Brown, 2006), we can conclude that the final CFA with this three-factor model showed satisfactory data. Absolute fit indices in this study consisted of Normed chi-square, GFI, NFI, TLI, CFI, RMSEA and SRMR, and all generated acceptable results.

Each absolute fit index was unique and cannot be used individually because the indices are sensitive to several factors such as the number of samples, whether the data is based on population or not, parsimony or complexities of a model and number of parameters of a model (Baggozi & Yi, 1988; Tanaka, 1993). As for the comparative fit indices (Brown, 2006), we also referred to as incremental fit indices (Hu & Bentler, 1998), we analyzed three indices, i.e. indices of CFI, TLI and NFI, which generate satisfactory data (CFI = 0.94, NFI = 0.91 and TLI = 0.93). Although NFI is very sensitive to sample size less than 200, CFI is an index that is not affected by sample size and also can be used as an indicator to ensure no misspecification models. TLI is an important index to compensate for the effect of the model complexity (Brown, 2006).

Internal consistency, as determined by the alpha coefficient for each factor, was adequate in this study (alpha coefficient = 0.71 for the instrument; 0.75–0.88 for each factor). The results presented in Table 3 attested to the high internal
consistency of the instrument in which all values were above the suggested 0.70 level for scale robustness (Nunnally & Bernstein, 1994).

Construct validity of a three-factor model showed that the AVE and CR of all constructs yielded acceptable values, but the test of discriminant validity resulted in unidimensional constructs for NPI and PPI, whereas the construct of TWC was two-dimensional. A high alpha value does not necessarily indicate that a factor is unidimensional. The alpha coefficient is not sufficient for measuring the dimensionality of a construct or factor (Tavako & Dennick, 2011). This is evidenced by TWC’s factor in this study in which the alpha coefficients of this factor was high but TWC’s factor was multidimensional. Several factors contributed to the emergence of discriminant validity problems and low factor loading. In this study, the existence of cultural diversity among respondents such as gender, ethnicity, belief and local cultures (Pashaei, Razaghi, Foroushani, & Tabatabaei, 2013) as well as students’ experiences in multi disciplinary work (McFadyen et al., 2005) may have influenced the responses to the items on the Indonesian version. In addition, the health and education systems prevailing in Indonesia and the effect of translation into a construct was multidimensional. Several factors contributed to the emergence of discriminant validity problems and low factor loading. In this study, the existence of cultural diversity among respondents such as gender, ethnicity, belief and local cultures (Pashaei, Razaghi, Foroushani, & Tabatabaei, 2013) as well as students’ experiences in multi disciplinary work (McFadyen et al., 2005) may have influenced the responses to the items on the Indonesian version. In addition, the health and education systems prevailing in Indonesia and the effect of translation into a particular language played a significant role in this study. The important issue is that when translating an instrument to another language and testing its validity, the role of cultural differences must be considered (Pashaei et al., 2013).

Several studies on the validation of the RIPL scale showed that the factor of RR is unstable. Parsell and Bligh’s study (1998) only generated two factors (TWC and PI), and the study on 1999 produced three factors (TWC, PI and RR) but the internal consistency of RR factor was unacceptable (<0.3). Similar results were also found in subsequent studies (Lauffs et al., 2008; McFadyen et al., 2005) and generated an inadequate internal consistency of RR factor, and another study failed to establish RR factor (Tamura et al., 2012; Williams et al., 2012). Only the study by McFadyen et al. (2006) yielded an adequate internal coefficient RR factor (RR = 0.62). This condition may be related to the ability of the respondents to comprehend the meaning of role and responsibilities in clinical setting (McFadyen et al., 2005). In this study, we failed to produce an RR factor, and this may have been due to omitting item 17. It may have affected the items 18 and 19 as they load on the same factor (RR factor).

In the main study, items 18 and 19 had low loadings and indicated that both items were not appropriate in Indonesian context. Item 18, “I am not sure what my professional role will be/is”, had loadings of 0.436 by CFA. It is possible that the item wording introduced error variance. The participants in this study were first- to third-year students. The first-year students could answer this question by seeing this statement as contrary to their understanding. Many studies of students’ perception about career and profession show that first-year students give positive responses about their future career as a physician, nurses or pharmacist (Kritikos, Watt, Krass, Sainsburry, & Bosnic-Anticevich, 2003; Law & Arthur, 2003; Rudland & Mires, 2005; Watmough, Waddelove, & Jaeger, 2009), whereas the advanced students’ views toward their profession or other professions were acquired from teachers, media or prior work experiences. If students encountered unclear tasks and the overlapping of RR, then students would think that their professional role is not clear (Hall, 2005). Furthermore, the understanding of the term “role and professional responsibility” in the clinical field may not be so obvious to young professionals as compared to the advanced students (McFadyen et al., 2005).

Item 19, “I have to acquire much more knowledge and skill than other students/professionals in my own faculty/organisation”, had loadings of 0.255 on factor 2 by EFA and had loading of −0.133 by CFA. The wording of this question might have been a problem, because having confident in one’s performance on “much more knowledge and skills” measure not only academic self-efficacy but also generalized self-esteem. In social learning theory, self-efficacy reflects an individual’s understanding of what knowledge and skills he/she can offer to the members in a group setting (Ormrod, 1999). However, item 19 implied not only high self-efficacy but also implied the meaning “to be more superior than the others” in Indonesian context. The meaning of this statement is contrary to the concept of equanimity in Indonesian society. The concept of “equanimity” in Indonesian society is influenced by beliefs and religion and implies doing everything for the good of humanity and nature and not for selfish purposes. More than two-thirds of the participants in this study were from the countryside, and they studied in Islamic boarding schools so this concept is deep-rooted.

Reconstructing the measurement model in an Indonesian version was successful because the fit model, 16 items three-factor model, yielded the highest reliability and construct validity and the acceptable GFI when compared to the initial model (18 items four-factor model). Although the 16 items three-factor model of Indonesian version is favorable, there are several limitations. First, our sample was confined to a school of health professionals in Indonesia that might not represent all Indonesian students. The findings may be difficult to generalize because the sample was only derived from one institution. Second, the sampling method was non-probability which may produce sampling bias. Nonetheless, the sample represented students from various multicultural and social backgrounds, which is reassuring. Finally, this study only validated 18 items of 19 items...
Conclusion comments

This study contributed to the literature in a number of ways. First, to our knowledge, it is the first study to investigate RIPLS in the Indonesian context. Second, the translation and adaptation of the RIPLS into the Indonesian language by a cross-cultural adaptation process was successful, and the Indonesian version produced a

Table 3. Confirmatory factor analysis: reliability, validity and fit indices of initial and modified models (N = 378).

| Indicator | Reliability–validity | Index fit |
|-----------|----------------------|-----------|
| FA1 (items 1–6) | 0.84 | 0.85 | 0.48 |
| FA2 (items 7–9) | 0.77 | 0.77 | 0.43 |
| FA3 (items 10–12, 18) | 0.72 | 0.74 | 0.42 |
| FA4 (items 13–16, 19) | 0.59 | 0.68 | 0.45 |
| B. Modified model (16 items) (0.87) | Three factors | | |
| TWC (items 1–9) | 0.88 | 0.89 | 0.46 |
| NPI (items 10–12) | 0.73 | 0.75 | 0.51 |
| PPI (items 13–16) | 0.86 | 0.86 | 0.62 |

Table 4. Analyzing of discriminant validity in the three-factor model (N = 378).

| Three factor model | TWC | NPI | PPI |
|--------------------|-----|-----|-----|
| TWC | 0.46* | | |
| NPI | 0.14 | 0.51* | |
| PPI | 0.59 | 0.18 | 0.62* |

*Diagonal elements report of the AVE and other matrix entries report the squared correlation estimation between two factors.
valid and reliable scale as verified by structural equation modeling. Third, two RIPLS items in the Indonesian version exhibited misfit measurement within the model and were excluded. Consequently, further revision and assessment of the RIPLS’ psychometric properties in the Indonesian version is recommended.

Declaration of interest

The authors report no conflicts of interest. The authors alone are responsible for the writing and content of this paper.

The research was funded from the University of Tokyo.

References

Bagus, R.P., & Yi, Y. (1988). On the evaluation of structural equation models. Journal of the Academy of Marketing Science, 16, 74–94.

Barr, H. (2010). The WHO framework for action. Journal of Interprofessional Care, 24, 475–478.

Brown, T.A. (2006). Confirmatory factor analysis for applied research. New York: The Guilford Press.

Browne, M.W., & Cudeck, R. (1993). Alternative ways of assessing model fit. Testing Structural Equation Models (pp. 136–162). Beverly Hills, CA: Sage.

Byrne, B.M. (2001). Structural equation modeling with AMOS, EQS, and LISREL: Comparative approaches to testing for the factorial validity of a measuring instrument. International Journal of Testing, 1, 55–86.

Chau, P.Y.K. (1997). Reexamining a model for evaluating information center success using a structural equation modeling approach. Decision Sciences, 28, 309–334.

DeVellis, R.F. (2012). Scale development: Theory and applications. Chapel Hill: Sage Publishers.

El-Zubeir, M., Rizky, D., & Al-Khalil, R. (2006). Are senior UAE medical and nursing students ready for interprofessional learning? Validating the RIPL scale in a Middle Eastern Context. Journal of Interprofessional Care, 20, 619–632.

Fornell, C., & Larcker, D.F. (1981). Evaluation of structural equation model with unobservables variable and measurement error. Journal of Marketing Research, 18, 39–50.

Freed, D., Reeves, S, Hammick, M., Koppel, I., & Barr, H. (2005). Evaluating interprofessional education: A self help guide. Higher Education Academy, Health Sciences and Practice Network. Retrieved from http://www.usask.ca/ipe/Documents/Evaluating%20IPE%20-%20A%20Self%20Help%20Guide.pdf.

Gilbert, J.H.V. (2005). Interprofessional education for collaborative, patient-centred practice. Nursing Leadership, 18, 32–38.

Guillemin, F., Bombardier, C., & Beaton, D. (1993). Cross-cultural adaptation of health-related quality of life measures: Literature review and proposed guidelines. Journal of Clinical Epidemiology, 46, 1417–1432.

Hair, J.F. (2005). Multivariate data analysis. Upper Saddle River, NJ: Prentice Hall.

Hall, P. (2005). Interprofessional teamwork: Professional cultures as barriers. Journal of Interprofessional Care, 19, 188–196.

Harrington, D. (2009). Confirmatory factor analysis. New York, NY: Oxford University Press.

Hui, L.-T., & Bentler, P.M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. Psychological Methods, 3, 424–453.

Johnson, B., & Stevens, J.J. (2001). Exploratory and confirmatory factor analysis of the school level environment questionnaire (SLEQ). Learning Environment Research, 4, 325–344.

Kline, R.B. (2005). Principles and practice of structural equation modeling. New York: Guilford.

Kritikos, V., Watt, H.M.G., Krass, L., Sainsbury, E.J., & Bosnic-Anticevich, S.Z. (2003). Pharmacy student’s perceptions of their profession relative to other health care professions. International Journal of Pharmacy Practice, 11, 121–129.

Laufi, M., Ponzer, S., Saboonchi, F., Lonka, K., Hylin, U., & Mattiasson, A.-C. (2008). Cross-cultural adaptation of the Swedish version of Readiness for Interprofessional Learning Scale (RIPLS). Medical Education, 42, 405–411.

Law, W., & Arthur, D. (2003). What factors influence Hong Kong school students in their choice of a career in nursing? International Journal of Nursing Studies, 40, 23–32.

Loxley, A. (1997). Collaboration in health and welfare: Working with difference. London: Jessica Kingsley Publishers.

McFadyen, A.K., Webster, V., Strachan, K., Figgins, E., Brown, H., & Mckechnie, J. (2005). The Readiness for Interprofessional Learning Scale: A possible more stable sub-scale model for the original version of RIPLS. Journal of Interprofessional Care, 19, 595–603.

McFadyen, A.K., Webster, V.S., & Maclaren, W.M. (2006). The test-retest reliability of a revised version of the Readiness for Interprofessional Learning Scale (RIPLS). Journal Of Interprofessional Care, 20, 633–639.

Mickan, S., Hoffman, S.J., & Nasmith, L. (2010). Collaborative practice in a global health context: Common themes from developed and developing countries. Journal of Interprofessional Care, 24, 492–502.

Nunnally, J.C., & Bernstein, I.H. (1994). Psychometric theory. New York: McGraw-Hill.

Ormrod, J.E. (1999). Human learning. Upper Saddle River, NJ: Prentice Hall.

Parsell, G., & Bligh, J. (1999). The development of a questionnaire to assess the readiness of health care students for interprofessional learning (RIPLS). Medical Education, 33, 95–100.

Parsell, G., Stewart, A., & Bligh, J. (1998). Testing the validity of the ‘‘Readiness for Interprofessional Learning Scale’’ (RIPLS). Paper presented at the 8th Ottawa International Conference, Philadelphia, PA, 12–15 July 1998.

Pushaee, T., Razaghi, O.M., Foroushani, A.R., & Tabatabaei, M.G. (2013). Assessing the validity and reliability of the Farsi version of inventory drug-taking situations. Iranian Journal of Psychiatry, 8, 80–85.

Reid, R., Bruce, D., Allstaff, K., & McMerton, D. (2006). Validating the readiness for interprofessional learning scale (RIPLS) in the post-graduate context: are health care professionals ready for IPL? Medical Education, 40, 415–422.

Rudland, J.R., & Mires, G.J. (2005). Characteristics of doctors and nurses as perceived by students entering medical school: Implications for shared teaching. Medical Education, 39, 448–455.

Schumacker, R.E., & Lomax, R.G. (2004). A beginner’s guide to structural equation modeling. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Tamura, Y., Seki, K., Usami, M., Taku, S., & Bontje, P. (2012). Cultural adaptation and validating a Japanese version of the readiness for interprofessional learning scale (RIPLS). Journal of Interprofessional Care, 26, 56–63.

Tanaka, J.S. (1993). Multifaceted conceptions of fit in structural equation models (pp. 10–39). Newbrury Park, CA: Sage.

Tavak, M., & Dennick, R. (2011). Making sense of Cronbach’s alpha. International Journal of Medical Education, 2, 53–33.

Thannhauser, J., Russell-Mayhew, S., & Scott, C. (2010). Measures of interprofessional education and collaboration. Journal of Interprofessional Care, 24, 336–349.

van de Vijver, F.J.R., & Hambleton, R.K. (1996). Translating tests: Some practical guidelines. European Psychologist, 1, 89–99.

Watmough, S., Waddelove, C., & Jaeger, L. (2009). First year medical students in their choice of a career in nursing? In Constructing the future: Career guidance for changing contexts. Stourbridge: Institute of Career Guidance.

WHO. (2010). Framework for action on interprofessional education and collaborative practice. Geneva: World Health Organization.

Williams, B., Brown, T., & Boyle, M. (2012). Construct validation of the readiness for interprofessional learning scale: A Rasch and factor analysis. Journal of Interprofessional Care, 26, 326–332.