

# Rasch analysis of Stamps's Index of Work Satisfaction in nursing population

Nora Ahmad<sup>1</sup> | Nelson Ositadimma Oranye<sup>2</sup> | Alyona Danilov<sup>3</sup>

<sup>1</sup>Department of Nursing, Brandon University, Brandon, Manitoba, Canada

<sup>2</sup>Department of Occupational Therapy, University of Manitoba, Winnipeg, Manitoba, Canada

<sup>3</sup>Neil John Maclean Health Sciences Library, University of Manitoba, Winnipeg, Manitoba, Canada

## Correspondence

Nelson Ositadimma Oranye, Department of Occupational Therapy, University of Manitoba, Winnipeg, Manitoba, Canada.  
Email: nelson.oranye@umanitoba.ca

## Abstract

**Aim:** One of the most commonly used tools for measuring job satisfaction in nursing is the Stamps Index of Work Satisfaction. Several studies have reported on the reliability of the Stamps' tool based on traditional statistical model. The aim of this study was to apply the Rasch model to examine the adequacy of Stamps's Index of Work Satisfaction for measuring nurses' job satisfaction cross-culturally and to determine the validity and reliability of the instrument using the Rasch criteria.

**Design:** A secondary data analysis was conducted on a sample of 556 registered nurses from two countries.

**Methods:** The RUMM 2030 software was used to analyse the psychometric properties of the Index of Work Satisfaction.

**Results:** The persons mean location of -0.018 approximated the items mean of 0.00, suggesting a good alignment of the measure and the traits being measured. However, at the items level, some items were misfitting to the Rasch model.

## KEYWORDS

nurses, nursing, Rasch analysis, job satisfaction, Stamp's Index

## 1 | BACKGROUND

Job satisfaction remains an important topic in organizational studies and has been extensively studied in many fields, including nursing. Studies on job satisfaction dates back to as early as 1920 (Snarr & Krochak, 1996) and have been studied with numerous tools and in different populations. The existing evidence shows that job satisfaction is influenced by multiple factors operating at the level of the job, individual, professional, organizational and the general work environment (Pittman, 2007; Ravari, Bazargan, Vanaki, & Mirzaei, 2012). Some of the specific factors that have been found to affect nurses' job satisfaction are job stress (Flanagan & Flanagan, 2002), management style of nursing leadership (Pietersen, 2005; Yamashita, Takase, Wakabayashi, Kuroda, & Owatari, 2009), empowerment (Cicolini, Comparcini, & Simonetti, 2014; Manojlovich & Laschinger, 2002), nursing autonomy

(Castaneda & Scanlan, 2014; Hayes, Bonner, & Pryor, 2010), co-worker interactions, group cohesion and salary (Curtis & Glacken, 2014; Wielenga, Smit, & Unk, 2008). The multiplicity of factors that impinge on nurses' job satisfaction have made the development of measurement tools that are valid and reliable across different work and cultural environments very challenging. However, several measurement tools have emerged over time, most of which have demonstrated high reliability and validity.

There are several reasons why job satisfaction among nurses has remained a persistent and hot topic in the nursing literature. Many researchers recognize the need to monitor job satisfaction of nurses because nurses' dissatisfaction could be disruptive to patient care delivery and reduce healthcare organizational effectiveness (Cheung & Ching, 2014; Curtis, 2007; Djukic, Kovner, Budin, & Norman, 2010; Taunton et al., 2004). Also, job satisfaction has been linked to different outcomes

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

for the nurses, which includes nurses' perceived ability to express caring behaviours with patients (Amendolair, 2012), new immigrant nurses' acculturation (Ea, Griffin, L'Eplattenier, & Fitzpatrick, 2008) and 'lower levels of job-stress, burnout and career abandonment among nurses' (Foley, Lee, Wilson, Cureton, & Canham, 2004, p. 94). Nurses job satisfaction has also been associated with positive patient outcomes, such as reduced patient falls (Alvarez & Fitzpatrick, 2007). However, it is important that measurement tools used for job satisfaction are constantly reviewed to ensure that they are measuring what they are intended to measure and that users are made aware of any pitfalls, should they choose to use such tools.

### 1.1 | Measurement tools for job satisfaction

A large body of research on job satisfaction has been accumulated, either using or attempting to validate well-known measurement tools or new tools that assess nurses' job satisfaction. Our search of the literature on nurses' job satisfaction from 1986 to May, 2015 identified 100 studies that reported measurement of job satisfaction in nursing. Among these studies, there were 20 different instruments used to measure nurses' job satisfaction. Some of the tools that showed good reliability and validity and which were most commonly used include: Minnesota Satisfaction Questionnaire developed by Weiss and colleagues in 1967 (Kaplan, Boshoff, & Kellerman, 1991; Lamarche & Tullai-McGuinness, 2009; Stamps, 1997; Weiss, Dawis, & England, 1967); Index of Work Satisfaction (IWS) developed by Stamps and Piedmont in 1970s (Slavitt, Stamps, Piedmont, & Haase, 1978; Stamps & Piedmonte, 1986); Quinn and Staines's Facet-free Job Satisfaction Scale developed by Quinn and Staines in 1979 (Djukic et al., 2010; Kovner, Brewer, Wu, Cheng, & Suzuki, 2006); Mueller and McCloskey's Satisfaction Scale (MMSS) developed by Mueller and McCloskey in 1990 (Misener, Haddock, Gleaton, & Abuajamieh, 1996; Mueller & McCloskey, 1990; Price, 2002; Tourangeau, Hall, Doran, & Petch, 2006).

The Minnesota Satisfaction Questionnaire has Cronbach  $\alpha$  range of 0.83–0.84 and validity between 0.32–0.75 (Lamarche & Tullai-McGuinness, 2009). Zurmehly (2008) noted that Hoyt reliability coefficient between 0.59–0.97 has been reported for the Minnesota Satisfaction Questionnaire, while Kaplan et al. (1991) reported a Cronbach  $\alpha$  ranging between 0.82–0.90 for the different components, which demonstrate adequate reliability. The internal consistency of the MMSS was reported as 0.89 in Mueller and McCloskey (1990) and 0.90 in Misener et al. (1996). The test–retest reliability for the subscales ranged from 0.08–0.64 (Misener et al., 1996; Mueller & McCloskey, 1990). With regard to Quinn and Staines's Facet-free Job Satisfaction Scale, Kovner et al. (2006) reported reliability coefficients for the scales ranging from 0.70–0.95. The psychometric properties of IWS have been reported in multiple studies (Ahmad & Oranye, 2010; Huber et al., 2000; Stamps, 1997; Wade et al., 2008), which reported on the internal consistency reliability and the validity of the IWS scales. Zangaro and Soeken (2005) explored the reliability and validity of the IWS through a meta-analysis of 14 studies that used the IWS to measure nursing job satisfaction. The meta-analysis by Zangaro and Soeken (2005) included only articles that reported the reliability of part B of the IWS and concluded that the part B of the IWS was reliable and valid in different

settings, including university, community and acute care hospitals and for multisite studies. The internal consistency reliability and validity of the IWS scale and its subscales ranged from 0.50–0.92 Cronbach's  $\alpha$  (Bjork, Samdal, Hansen, Torstad, & Hamilton, 2007; Hayes, Douglas, & Bonner, 2015; Itzhaki, Ea, Ehrenfeld, & Fitzpatrick, 2013; Manojlovich & Laschinger, 2007; Penz, Stewart, D'Arcy, & Morgan, 2008; Zangaro & Soeken, 2005). The highest subscale coefficient of 0.92 was reported by Manojlovich and Laschinger (2007), while the lowest Cronbach's alpha was reported by Medley and Larochelle (1995). The Cronbach's  $\alpha$  originally reported by Stamps (1997) ranged from 0.82–0.91. Content validity (Kovner, Hendrickson, Knickman, & Finkler, 1994) and construct validity through factor analysis (Stamps, 1997) have been established.

Among these job satisfaction measurement tools, the IWS has been one of the most widely used. The IWS measures 'the extent to which people like their jobs' (Stamps, 1997, p. 13) and provides a quantitative estimation of nurses' job satisfaction. The tool was amplified in 1986 by Stamps and Piedmont based on a critical review of occupational theories in the social sciences (Amendolair, 2012; Kovner et al., 1994; Slavitt et al., 1978; Stamps & Piedmonte, 1986). The strong theoretical foundation of Stamps's IWS was intended to address the seemingly atheoretical plunge of many of the extant job satisfaction measurement tools. Stamps and Piedmonte (1986, p. 19), noted that they '...proceeded to develop a valid and reliable scale for measuring nurses' work satisfaction, one general enough to be used in many settings...' The IWS scale assesses the level of nurses' professional satisfaction in six work dimensions: payment, professional status, task requirements, interactions, organizational policies and autonomy (Stamps & Piedmonte, 1986) and is rated on a seven-point Likert scale. The level of professional satisfaction for each of the six dimensions (subscales' scores) and the overall professional satisfaction level (entire IWS score) have been reported in previous studies.

Hitherto, the statistical methods typically used for psychometric measurement in nursing research were based on the traditional statistical model. The Rasch analysis model provides an alternative to the traditional psychometric measurement that is sophisticated, comprehensive and is based on the Item Response Theory (Belvedere & de Morton, 2010; Hagquist, Bruce, & Gustavsson, 2009). The Rasch model was originally developed for measuring the psychometric properties of educational testing tools (Andrich, 2005), but nowadays, has been increasingly used in health sciences and many other disciplines. However, not many studies have been undertaken using the Rasch model in health sciences (Hagquist et al., 2009). A successful implementation of the Rasch measurement requires that the assumptions of local independence and unidimensionality are satisfied (Brentari & Golia, 2008). In addition to the criteria of unidimensionality and local independence, Rasch uses the criteria of differential item functioning (DIF), person separation index (PSI) and fit statistics to determine the reliability and validity of a measurement tool.

Few studies have been undertaken using the Rasch model in the health sciences (Hagquist et al., 2009) and very few studies have used Rasch to measure nurses' job satisfaction. There were three articles that applied the Rasch model in nursing (Clinton, Dumit, & El-Jardali, 2015; Flannery, Resnick, Galik, Lipscomb, & McPhaul, 2012; Hagquist et al., 2009); however, despite the wide use of Stamps' IWS in nursing research

and in diverse environments, no study has applied the Rasch model to evaluate its reliability and validity. The purpose of this study was to apply the Rasch model to examine the adequacy of Stamps's Index of Work Satisfaction for measuring nurses' job satisfaction cross-culturally and to determine the validity and reliability of IWS using the Rasch criteria.

## 2 | METHODOLOGY

This is a secondary data analysis that uses data from Ahmad and Oranye (2010) survey of registered nurses in two teaching hospitals in Malaysia and England. A total of 556 registered nurses participated in that study and are included in this analysis. The survey used four previously developed scales of Structural Empowerment scale, The Psychological Empowerment scale, Meyer and Allen Organizational Commitment Scale and the Index of Work Satisfaction scale. Details of the study design and description of the tools are reported in Ahmad and Oranye (2010).

### 2.1 | Ethics

The study complies with the international human research ethics guideline and the Declaration of Helsinki code of ethics. The Ethical approval for the study was obtained from the University of Sheffield Ethics Committee, the NHS and Hospital directors in the two hospitals in England and Malaysia (Ahmad & Oranye, 2010).

### 2.2 | Procedure

The current descriptive study uses the data related to part B of Stamps (1997) IWS tool to determine the adequacy of the IWS tool in measuring job satisfaction cross-culturally, by applying the Rasch model. The IWS contains 44 items with six components of pay, autonomy, task requirements, professional status, interaction and organizational policies. There are six items in the pay subscale, eight in autonomy, six in task requirements, seven in professional status, 10 in interaction and seven in the organizational policies subscale (Ahmad & Oranye, 2010; Stamps & Piedmonte, 1986). The reliability index of the IWS has been reported in previous studies (Ahmad & Oranye, 2010; Medley & Laroche, 1995; Wade et al., 2008) and in the Manual (Stamps & Piedmonte, 1986). This study, conducted a systematic search of the literature related to

nurses' job satisfaction and research in four major databases of PubMed, CINAHL, PsycINFO and SCOPUS from 1986 - May 2015, to find studies that have relevance to this study and to ascertain if Rasch model has been applied to the IWS. The following inclusion and exclusion criteria were applied: (1). Papers published in English language; (2). Publications with a study sample that included nurses; (3). Job satisfaction was measured using the IWS; (4). Reliability and validity of the IWS were reported for the study sample; (5). Papers that applied Rasch analysis model to measures of job satisfaction. The search resulted in 100 papers, which were further screened for relevance. Finally, 53 of the papers and four other papers on Rasch model were included in this study.

### 2.3 | Analysis

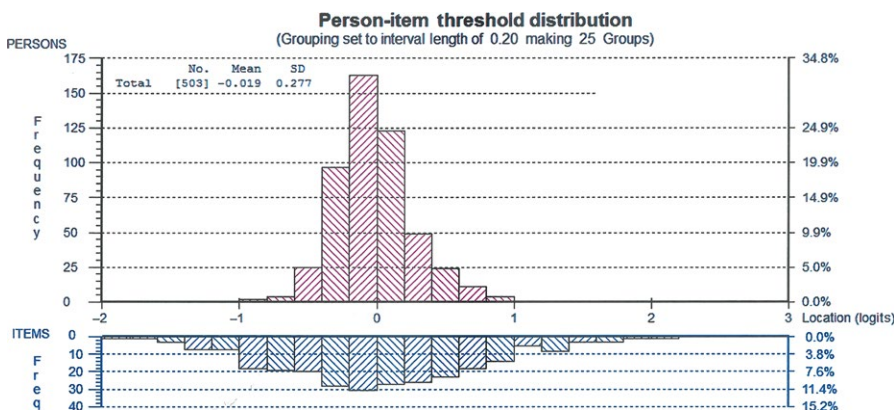
The Rasch analysis model was applied using the Rumm 2030 software. Data from the two countries were stacked for comparative analysis purpose. The Rumm 2030 software performs an item by item analysis, providing the capability to examine each item at different levels, including individual, country and other group levels, such as age, work status etcetera. The data stacking enables a simultaneous analysis of variables across the multiple levels. An analysis of the fit statistics was used to determine if IWS scale fits the Rasch model expectations. The statistics were examined to determine whether the criteria of unidimensionality, differential item functioning (DIF) and person separation index (PSI) were satisfied by the IWS scale (Brentari & Golia, 2008).

## 3 | RESULTS

### 3.1 | Descriptive analysis

Of the 554 subjects in the data, 70% were from Malaysia and 30% were British. The majority of the subjects were female (96.4%), which is a reflection of the gender composition of the nursing profession in many environments. The majority were married (62.5%), while 34.5% were single and the others were divorced, widowed or unknown. A smaller proportion had a university degree (11%), but the most common level of education was Diploma (65.9%) and a certificate in nursing (23.1%). Most of the nurses worked as full time staff (90.4%).

In Rasch analysis, one way to measure the adequacy of a tool is the targeting of the traits of interest in the population. In Fig. 1, the spread



**FIGURE 1** Person-item threshold distribution

**TABLE 1** Test of reliability

Scales	With extremes		Without extremes		N
	PSI	Conbach $\alpha$	PSI	Conbach $\alpha$	
All 44 items	0.8578	0.851	0.8578	0.851	503
Professional status	0.6441	0.5567	0.6101	0.5418	540
Task requirement	0.5824	0.5611	0.5747	0.5611	550
Pay	0.4814	0.4819	0.4275	0.4612	540
Interaction	0.7811	0.7429	0.7811	0.7429	544
Organizational policies	0.591	0.5575	0.5697	0.5502	545
Autonomy	0.7258	0.6853	0.7115	0.6791	546

PSI, Person Separation Index.

of the items in the scale vis-à-vis the persons location shows that the tool has a good targeting of the person characteristics in the sample. The persons mean location of  $-0.018$  is approximately equal to the items mean of  $0.00$ . However, the negative persons mean value suggests the possibility of very few participants whose scores were lower than the theoretical expected average level of job satisfaction. This could also suggest a slightly lower level of job satisfaction among the population.

### 3.2 | Reliability indices

The Rasch model provides two estimates that confirm the reliability of a tool and the precision of the estimate of each person trait in the sample. The person separation index (PSI)= $0.8578$  was approximately equal to the Cronbach  $\alpha$  coefficient= $0.851$ , both of which indicate a very good reliability and internal consistency of the IWS (Table 1).

### 3.3 | Fit analysis

Table 2 shows fit statistics for the item-person interaction. The Rasch analysis indicates an excellent power of analysis of fit for the data, which means that the analysis was strong enough to detect any differences where there was one. The standard deviation of the fit

residuals for the items at the subscale levels and the total scale were high, suggesting poor fit to the Rasch model. Generally, these suggest the presence of some mis-fitting items and individuals in the data set whose response patterns deviated substantially from the expectation of the Rasch model (Tennant & Conaghan, 2007). In all the six subscales, the residual standard deviation for the items were higher than the residual standard deviations for persons. The persons residual standard deviations for the Professional status, Task requirement and Pay subscales are below  $1.4$ , suggesting that it is very unlikely there were persons whose responses deviated significantly from the Rasch model expectation in those subscales. The significant Chi Square,  $p < .0001$ , equally indicates a misfit to the Rasch model. Cummings, Hayduk, and Estabrooks (2006) have argued that the lack of model fit in a measurement tool is an indication of a validity problem. The Rasch model identified 51 extreme cases, which were subsequently dropped from the analysis. The removal of these extreme cases did not alter the power of analysis fit and the Chi Square fit statistics remained significant. So, the misfitting was not caused by the extreme values. The RMSEA was calculated to further evaluate the model fit. The result shows that the Pay subscale has a poor fit to the Rasch model, while the two subscales of Interaction and Organizational Policies closely approximate the Rasch Model. For the other three subscales, there is a reasonable error of approximation to the model fit (Browne & Cudeck, 1993).

#### 3.3.1 | Unidimensionality test

Another important statistics considered in this study is the unidimensionality test, using the paired  $t$  test statistics. The paired  $t$  test =  $-2.8$ , shows that 187 of the sample estimates were significantly different at  $p < .05$  and 110 were significantly different, at  $p < .01$ . The  $t$  test statistic gave a significant value much higher than the 5% required for Rasch unidimensionality. This analysis supports the multidimensionality of the IWS scale, which was originally designed to measure six dimensions of pay, autonomy, task requirement, organizational requirement, job status and interaction.

The individual item fit residuals were examined to identify those items that may be causing the model miss fit. The result from the

**TABLE 2** Summary test-of-fit on item-person interaction

Scales	Items		Persons		RMSEA	$\chi^2$	p value
	Location Mean (SD)	Residual Mean (SD)	Location Mean (SD)	Residual Mean (SD)			
Total scale	0.00 (0.39)	0.63 (1.65)	-0.02 (0.26)	-0.42 (2.34)	0.069	1189.42	<.0001
Professional status	0.00 (0.31)	0.76 (1.62)	0.35 (0.5)	-0.29 (1.13)	0.055	146.33	<.0001
Task requirement	0.00 (0.48)	-0.03 (1.91)	-0.31 (0.51)	-0.4 (1.08)	0.053	121.02	<.0001
Pay	0.00 (0.39)	0.59 (3.36)	-0.42 (0.43)	-0.36 (1.15)	0.115	390.5	<.0001
Interaction	0.00 (0.37)	0.78 (1.61)	0.23 (0.53)	-0.37 (1.42)	0.036	135.06	<.0001
Organizational policies	0.00 (0.19)	0.83 (1.57)	-0.28 (0.43)	-0.46 (1.48)	0.045	102.21	<.0001
Autonomy	0.00 (0.23)	0.31 (2.25)	0.11 (0.53)	-0.56 (1.58)	0.058	180.5	<.0001

SD, standard deviation; RMSEA, root mean square error of approximation.

**TABLE 3** Analysis of miss fitting items based on subscales

Subscales	Item	Location	SE	FitR	df	$\chi^2$	df	p values
Task requirement	22	-0.67	0.04	2.43	454.5	46.87	8	<.0001*
	36	0.50	0.04	-2.57*	454.5	18.6	8	.0172
Pay	01	-0.2	0.03	-0.21	443.7	27.68	8	.0005*
	14	0.08	0.03	-2.57*	443.7	69.24	8	.0000*
	21	0.05	0.03	0.01	443.7	28.6	8	.0004*
	32	-0.64	0.03	6.75*	443.7	215.09	8	<.0001*
	44	0.53	0.04	-1.997	443.7	37.3	8	<.0001*
Interaction	03	-0.05	0.03	3.83	485.7	11.85	8	.1582
	10	0.15	0.03	3.26*	485.7	14.22	8	.0762
Autonomy	07	-0.19	0.03	2.26*	473	26.61	8	.0008*
	31	0.00	0.03	-2.5	473	25.71	8	.0012*
	43	0.02	0.03	3.43*	473	23.33	8	.003
Organizational policy	18	-0.1	0.03	2.62*	462.4	11.00	7	.1384
	25	0.00	0.03	-1.19	462.4	25.87	7	.0005*
	33	-0.01	0.03	2.63	462.4	20.13	7	.0051
Professional status	02	0.47	0.03	3.66	456.4	34.38	8	<.0001*
	09	0.04	0.03	0.60	456.4	29.29	8	.0003*
	15	-0.47	0.04	-0.9	456.4	30.42	8	.0002*

Note \* denotes significant p values. SE, standard error; FitR, Fit Residual.

subscales shows that items 7, 10, 14, 18, 32, 36 and 43 had extreme fit residual values. The fit residuals for items 7 and 32 were consistently high at subscale and combined scales levels. Also, the Table 3 shows that a total of 12 items had significant Bonferroni Adjusted  $\chi^2$  probability <0.00125, indicating that these items were miss fitting of the Rasch model.

### 3.3.2 | Analysis of Differential Item Functioning (DIF)

The DIF is a test of item bias or how each item in the scale functions for each individual, irrespective of 'ability level'. In this study, the DIF was examined with respect to age, gender, years of experience, work status (full time or part time) and country. A primary factor of interest is the country, whether participants were British or Malaysian nurses, which by extension implies socioeconomic and cultural differences.

Table 4 shows items with significant DIF for the six person factors. The difference between participants was considered significant, if the *F*-statistics has an adjusted Bonferroni probability <.001667. A total of 18 items had significant DIF at country level, 10 of which have significant or high fit residuals. It is known that the presence of DIF can cause a misfit to the model. One item, 42 had significant DIF for sex,  $p = .0014$ . A few items showed evidence of significant DIF for age, work status, years of experience, education and marital status. The large number of items with significant DIF for country raises questions about the cross-cultural validity of the IWS tool.

One-way ANOVA was performed to determine the significance of the variation in the IWS items with regard to the person factors. The result in Table 5 shows significant variation at country and education levels, adjusted Bonferroni probability <.000379. The high significant variation for country,  $p < .0001$  supports the result from Table 4 that the IWS functions very differently for nurses in Malaysia than those in the UK.

The person-item distribution (Fig. 2) shows that Malaysian nurses had lower levels of job satisfaction, with a group mean = -0.091, compared with the group mean = 0.158 for UK nurses. The wide spread of the items could mean that some of the items may not be relevant for understanding the constructs, given that there were many measures on the two extremes.

## 4 | DISCUSSION

The PSI and Cronbach's  $\alpha$  for the Index of Work Satisfaction in this study are consistent with previous studies, which reported good reliability, ranging from 0.54–0.92 (Bjork et al., 2007; Curtis & Glacken, 2014; Manojlovich & Laschinger, 2007; Oermann, 1995; Stamps & Piedmonte, 1986). There is a strong evidence, both from this and previous studies that support the reliability of IWS for assessing job satisfaction among nurses. However, it is possible that the variation in the reliability reported across studies is an indication that the meanings or values of some of the items may not always be consistent across populations. For instance, Karanikola and Papatthanasoglou (2015) found that two items in the IWS scale were not consistent with other items and as such affected the internal consistency of the tool. Essentially, the reliability of a measurement tool focuses on the consistency of the measurement in measuring what it is intended to measure. However, what is measured, especially in the social world, is often inequivalent across social environments, because the meanings and values vary from one place to another. Therefore, it is important that attention is paid not just to the consistency of a scale, but the meanings and values of the concept or construct being measured, across cultures.

The high standard deviations of the fit residual for the items (range: 1.57–3.36) points to the possibility of mis-fitting items, while the fit residual for the persons (range: 1.08–1.58) shows the less likelihood



**TABLE 4** Analysis of DIF by subscale

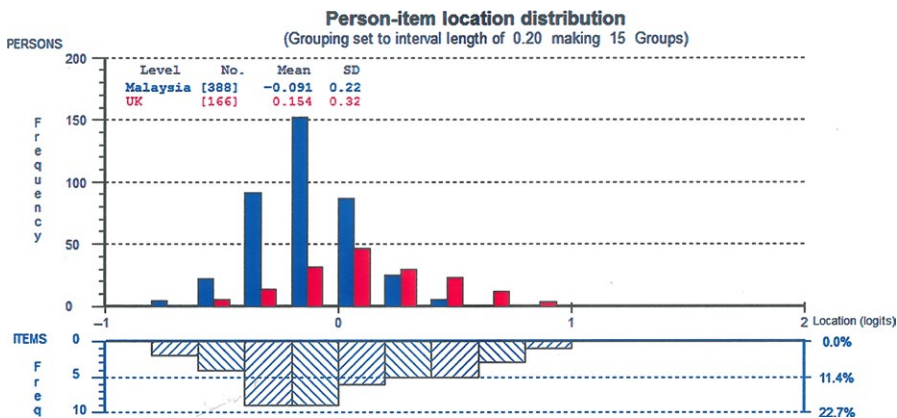
Subscales	Items	Country		Age <i>p</i> value	Work status <i>p</i> value	Experience <i>p</i> value	Education <i>p</i> value	Gender <i>p</i> value	Marital status <i>p</i> value
		<i>F</i>	<i>p</i> value						
Professional status	2	11.0799	.0009						
	15			<.0001		<.0001			
	27	84.3628	<.0001				.0011		
	38	20.0570	<.0001						
	41	27.485	<.0001						
Task requirement	4			.0006		.0009			
	22	31.1457	<.0001	<.0001		.0001	.0002		
Pay	1	19.4680	<.0001						
	32	76.4063	<.0001						.0012
	44	52.5904	<.0001						.0015
Interaction	6	13.2544	.0003						
	19	21.8351	<.0001						
	35	21.5627	<.0001						
Organizational policies	18	27.6627	<.0001						
	33	23.9983	<.0001						
	40	26.5460	<.0001						
	42	61.7807	<.0001					.0014	
Autonomy	7	50.2147	<.0001		.0002				
	13			<.0001		<.0001	.0006		
	17					.0016			
	26						.0003		
	31	13.9779	.0002						
	43	82.4976	<.0001						

Note. The *p* values reported in the table were below adjusted bonferroni probability 0.001667. \*Lowest adjusted bonferroni probability <.001667 for all the subscales.

of individuals whose response patterns deviated substantially from the expectation of Rasch model, compared with the items (Tennant & Conaghan, 2007). The Chi Square statistics and RMSEA were used to determine if the person-item interaction in the IWS provides a good fit to the Rasch model. While the significant Chi Square,  $p < .0001$ , indicates a deviation from the Rasch model, the RMSEA suggests a mixed bag, with some subscales presenting a better fit than others. Apart from the Pay subscale which has a poor fit to the Rasch model, the two subscales of Interaction and Organizational Policies closely approximate the Rasch Model. The other three subscales of Autonomy, Professional Status and Task Requirement demonstrated reasonable errors of approximation to the model fit (Browne & Cudeck, 1993), which may not necessarily imply a complete misfit to the Rasch model. Cummings et al. (2006) have pointed out that the lack of model fit in a measurement tool is an indication of validity problem. There are several reasons why a measurement tool or data may have a model misfit. The lack of Fit to the Rasch Model in this study could be due to cultural differences between the two countries, the size of the sample or because the IWS is a multidimensional scale. It is also known that the presence of DIF can cause a misfit to the Rasch model. All of these factors are true in this study analysis. Essentially, our primary interest in this analysis was to determine whether the IWS tool functions differently for different groups (DIF), whether the groups are at country level, gender, work status etc. Some of the items identified in this study would require further analysis, to determine why they have high or significant fit residual.

The IWS items align very well with the persons measure, but overall, the spread of the items were wider on both tails of the graph (Fig. 1) than the person traits being measured. The spread seems to suggest that some of the measures were either above or below the respondents' 'ability' level. In the context of the measurement of job satisfaction, the items at the extreme were possibly measuring traits that may not be directly relevant to understanding participants' job satisfaction. Again, it is important to note that what makes for job satisfaction would very likely vary in time, place and people. A detailed individual item-response analysis will be required to identify those items in the tool that are probably irrelevant or contributing very little to the measurement of the construct of job satisfaction or the underlying concepts of pay, professional status, interaction, task requirements and organizational policies.

The Rasch model expects a good measurement tool to be invariant across the sample and traits being measured. In other words, each item on a measurement scale is expected to measure the attribute of interest between different participants without any bias. Linacre and Wright (1987) have emphasized the importance of identifying and quantifying differential item functioning for contrasting groups and to clearly understand the differences between groups. For a measurement tool, such as the IWS that is designed to be used in different environments, it is important to understand how the different items in the tool function for different participants and groups. The presence of a substantial number of items with



**FIGURE 2** Person-item distribution by country

**TABLE 5** ANOVA for DIF in participants

Factor	F	df <sub>B</sub>	df <sub>W</sub>	p value
Country	98.93	1	501	.0000
Age	4.84	3	499	.0025
Education	8.97	2	500	.0002
Experience	4.82	3	499	.0026
Sex	4.76	1	501	.0295
Marital status	3.22	2	500	.0408
Work status	1.55	1	501	.2132

Note. p values for the subscales are significant if below 0.000379.  
\*Adjusted Bonferroni probability <.000379 for all the 44 items.

significant DIF,  $p < .0017$  for participants from Malaysia and UK in this study is an important measurement issue that researchers who use the IWS scale need to pay attention to. A significant DIF could be an evidence of measurement bias, or may result from the nature of the constructs being measured. Such a bias will have important implications for the cross-cultural validity of how job satisfaction is being measured by IWS. Also, it should be noted that the meanings of the concepts of payment, professional status, task requirements, professional interactions, organizational policies and autonomy could differ for the nurses in these two countries. The meanings of the item questions in each of the six components and the values of what is being measured may not be equivalent across cultures and countries. This significant DIF for the countries raises another important question on the use of IWS for measuring and comparing nurses' job satisfaction across cultural groups and whether the findings from different countries are actually comparable or generalizable across countries, or even among cultural groups in the same country or workplace. Belvedere and de Morton (2010) have pointed out that the presence of DIF calls to question the validity and generalizability of a measurement result.

The negative mean of  $-0.091$  indicates that on the average, job dissatisfaction was lower among nurses in Malaysia than those in the UK. Ahmad and Oranye (2010) have reported a significant difference in job satisfaction between the English and Malaysian nurses and pointed out that the factors that determine job satisfaction was different for both groups. Several studies (Adwan, 2014; Alvarez & Fitzpatrick,

2007; Andrews, Stewart, Morgan, & D'Arcy, 2012; Ea et al., 2008) have reported total job satisfaction based on the IWS tool. Adwan (2014) reported high scores in most of the IWS subscales among paediatric patient care nurses, while Alvarez and Fitzpatrick (2007) reported (67%) moderate job satisfaction and (33%) low job satisfaction at the unit levels. The calculation of total scores on job satisfaction was made on the assumption that the scores on the subscales are additive. However, given the multidimensionality of the IWS scale and the differences in the meanings of what is being measured, it is questionable that total scores are realistically comparable. The study by Ahmad and Oranye (2010) points to the fact that what determines job satisfaction can vary between groups and countries. For instance, while the pay was the significant determinant of job satisfaction among the English nurses, 'interaction—the opportunities presented for both formal and informal contacts during working hours' (Ahmad & Oranye, 2010, p. 589), was the primary determinant of job satisfaction among Malaysian nurses. These differences in perceived job satisfaction may be a function of DIF as evident in this study, than other workplace or condition of work factors.

#### 4.1 | Limitations

There were 51 extreme cases in this study sample; however, their removal did not significantly change the result. The findings from this single study may not be sufficient to draw definitive conclusions on the miss fit of IWS to the Rasch model. Further studies across countries and work environments that apply Rasch model and a review of local dependency and item difficulty levels is needed.

## 5 | CONCLUSION

The IWS is a very reliable tool, especially at the composite level, as indicated by this study and several others. The IWS has been used in several nursing studies, but its cross-cultural validity has not been well evaluated based on item-response theory and using Rasch model statistics of DIF. Caution should be exercised in comparing results of IWS across cultural groups, in view of the evidence on possible DIF for culturally diverse societies. It is important that further studies are

conducted to test for DIF across cultural groups. Given that the IWS is a multidimensional tool, it may not be realistic to sum up the scores from the different dimensions as an index for comparison between significantly different groups, since the issues they measure may vary over time and place. Equally, it may not be meaningful to compare the total score on job satisfaction between two different groups, since the meaning of job satisfaction or any of the components may differ significantly between groups.

## FUNDING

There is no funding for this study.

## CONFLICT OF INTEREST

The authors declare that there is no conflict of interest in this study.

## AUTHOR CONTRIBUTIONS

All authors have agreed on the final version and meet at least one of the following criteria [recommended by the ICMJE (<http://www.icmje.org/recommendations/>)]:

- substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data;
- drafting the article or revising it critically for important intellectual content.

## REFERENCES

- Adwan, J. Z. (2014). Pediatric nurses' grief experience, burnout and job satisfaction. *Journal of Pediatric Nursing*, 29(4), 329–336.
- Ahmad, N., & Oranye, N. O. (2010). Empowerment, job satisfaction and organizational commitment: A comparative analysis of nurses working in Malaysia and England. *Journal of Nursing Management*, 18, 582–591.
- Alvarez, C. D., & Fitzpatrick, J. J. (2007). Nurses' job satisfaction and patient falls. *Asian Nursing Research*, 1(2), 83–94.
- Amendolair, D. (2012). Caring behaviors and job satisfaction. *Journal of Nursing Administration*, 42(1), 34–39.
- Andrews, M., Stewart, N., Morgan, D., & D'Arcy, C. (2012). More alike than different: A comparison of male and female RNs in rural and remote Canada. *Journal of Nursing Management*, 20(4), 561–570.
- Andrich, D. (2005). Rasch, Georg. In K. Kempf-Leonard (Ed.), *Encyclopedia of social measurement*, Vol. 3 (pp. 299–306). Amsterdam: Academic Press.
- Belvedere, S. L., & de Morton, N. A. (2010). Application of Rasch analysis in health care is increasing and is applied for variable reasons in mobility instruments. *Journal of Clinical Epidemiology*, 63(12), 1287–1297.
- Bjork, I. T., Samdal, G. B., Hansen, B. S., Torstad, S., & Hamilton, G. A. (2007). Job satisfaction in a Norwegian population of nurses: A questionnaire survey. *International Journal of Nursing Studies*, 44(5), 747–757.
- Brentari, E., & Golia, S. (2008). Measuring job satisfaction in the social services sector with the Rasch Model. *Journal of Applied Measurement*, 9(1), 45–56.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. *Sage Focus Editions*, 154, 136–136.
- Castaneda, G., & Scanlan, J. (2014). Job satisfaction in nursing: A concept analysis. *Nursing Forum*, 49(2), 130–138.
- Cheung, K., & Ching, S. S. Y. (2014). Job satisfaction among nursing personnel in Hong Kong: A questionnaire survey. *Journal of Nursing Management*, 22(5), 664–675.
- Cicolini, G., Comparcini, D., & Simonetti, V. (2014). Workplace empowerment and nurses' job satisfaction: A systematic literature review. *Journal of Nursing Management*, 22(7), 855–871.
- Clinton, M., Dumit, N. Y., & El-Jardali, F. (2015). Rasch measurement analysis of a 25-item version of the Mueller/McCloskey Nurse Job Satisfaction Scale in a Sample of Nurses in Lebanon and Qatar. *SAGE Open*, 5(2), 1–10.
- Cummings, G. G., Hayduk, L., & Estabrooks, C.A. (2006). Is the nursing work index measuring up? Moving beyond estimating reliability to testing validity. *Nursing Research*, 55(2), 82–93.
- Curtis, E. A. (2007). Job satisfaction: A survey of nurses in the republic of Ireland: Original article. *International Nursing Review*, 54(1), 92–99.
- Curtis, E. A., & Glacken, M. (2014). Job satisfaction among public health nurses: A national survey. *Journal of Nursing Management*, 22(5), 653–663.
- Djukic, M., Kovner, C., Budin, W., & Norman, R. (2010). Physical work environment testing an expanded model of job satisfaction in a sample of registered nurses. *Nursing Research*, 59(6), 441–451.
- Ea, E. E., Griffin, M. Q., L'Eplattenier, N., & Fitzpatrick, J. J. (2008). Job satisfaction and acculturation among Filipino registered nurses. *Journal of Nursing Scholarship*, 40(1), 46–51.
- Flanagan, N. A., & Flanagan, T. J. (2002). An analysis of the relationship between job satisfaction and job stress in correctional nurses. *Research in Nursing and Health*, 25(4), 282–294.
- Flannery, K., Resnick, B., Galik, E., Lipscomb, J., & McPhaul, K. (2012). Reliability and validity assessment of the job attitude scale. *Geriatric Nursing*, 33(6), 465–472.
- Foley, M., Lee, J., Wilson, L., Cureton, V. Y., & Canham, D. (2004). A multi-factor analysis of job satisfaction among school nurses. *Journal of School Nursing (Allen Press Publishing Services Inc.)*, 20(2), 94–100.
- Hagquist, C., Bruce, M., & Gustavsson, J. P. (2009). Using the Rasch model in nursing research: An introduction and illustrative example. *International Journal of Nursing Studies*, 46(3), 380–393.
- Hayes, B., Bonner, A., & Pryor, J. (2010). Factors contributing to nurse job satisfaction in the acute hospital setting: A review of recent literature. *Journal of Nursing Management*, 18(7), 804–814.
- Hayes, B., Douglas, C., & Bonner, A. (2015). Work environment, job satisfaction, stress and burnout among haemodialysis nurses. *Journal of Nursing Management*, 23(5), 588–598.
- Huber, D. L., Maas, M., McCloskey, J., Scherb, C. A., Goode, C. J., & Watson, C. (2000). Evaluating nursing administration instruments. *Journal of Nursing Administration*, 30(5), 251–272.
- Itzhaki, M., Ea, E., Ehrenfeld, M., & Fitzpatrick, J. J. (2013). Job satisfaction among immigrant nurses in Israel and the United States of America. *International Nursing Review*, 60(1), 122–128.
- Kaplan, R. A., Boshoff, A. B., & Kellerman, A. M. (1991). Job involvement and job satisfaction of South African nurses compared with other professions. *Curationis*, 14(1), 3–7.
- Karanikola, M. N. K., & Papatthanassoglou, E. D. E. (2015). Measuring professional satisfaction in Greek nurses: Combination of qualitative and quantitative investigation to evaluate the validity and reliability of the index of work satisfaction. *Applied Nursing Research*, 28(1), 48–54.
- Kovner, C., Brewer, C., Wu, Y. W., Cheng, Y., & Suzuki, M. (2006). Factors associated with work satisfaction of registered nurses. *Journal of Nursing Scholarship*, 38(1), 71–79.
- Kovner, C. T., Hendrickson, G., Knickman, J. R., & Finkler, S. A. (1994). Nursing care delivery models and nurse satisfaction. *Nursing Administration Quarterly*, 19, 74–85.
- Lamarche, K., & Tullai-McGuinness, S. (2009). Canadian nurse practitioner job satisfaction. *Nursing Leadership*, 22(2), 41–57.
- Linacre, J.M., & Wright, B.D. (1987). *Item-bias: Mantel-Haenszel and the Rasch Model*. Memorandum No. 39. Chicago: MESA Psychometric Laboratory, Department of Education, University of Chicago.



- Manojlovich, M., & Laschinger, H. (2002). The relationship of empowerment and selected personality characteristics to nursing job satisfaction. *Journal of Nursing Administration*, 32(11), 586–595.
- Manojlovich, M., & Laschinger, H. (2007). The nursing worklife model: Extending and refining a new theory. *Journal of Nursing Management*, 15(3), 256–263.
- Medley, F., & Larochelle, D.R. (1995). Transformational leadership and job satisfaction. *Nursing Management*, 26, 64JJ–64NN.
- Misener, T., Haddock, K., Gleaton, J., & Abuajamieh, A. (1996). Toward an international measure of job satisfaction. *Nursing Research*, 45(2), 87–91.
- Mueller, C., & McCloskey, J. (1990). Nurses job satisfaction: A proposed measure. *Nursing Research*, 39(2), 113–117.
- Oermann, M. H. (1995). Critical care nursing education at the baccalaureate level: Study of employment and job satisfaction. *Heart and Lung—The Journal of Acute and Critical Care*, 24(5), 394–398.
- Penz, K., Stewart, N. J., D'Arcy, C., & Morgan, D. (2008). Predictors of job satisfaction for rural acute care registered nurses in Canada. *Western Journal of Nursing Research*, 30(7), 785–800.
- Pietersen, C. (2005). Job satisfaction of hospital nursing staff. *South African Journal of Human Resource Management*, 3(2), 19–25.
- Pittman, J. (2007). Registered nurse job satisfaction and collective bargaining unit membership status. *Journal of Nursing Administration*, 37(10), 471–476.
- Price, M. (2002). Job satisfaction of registered nurses working in an acute hospital. *British Journal of Nursing*, 11(4), 275–280.
- Ravari, A., Bazargan, M., Vanaki, Z., & Mirzaei, T. (2012). Job satisfaction among Iranian hospital-based practicing nurses: Examining the influence of self-expectation, social interaction and organizational situations. *Journal of Nursing Management*, 20(4), 522–533.
- Slavitt, D. B., Stamps, P. L., Piedmont, E. B., & Haase, A. M. (1978). Nurses' satisfaction with their work situation. *Nursing Research*, 27(2), 114–120.
- Snarr, C. E., & Krochak, P. C. (1996). Job satisfaction and organizational characteristics: Results of a nationwide survey of baccalaureate nursing faculty in the United States. *Journal of Advanced Nursing*, 24(2), 405–412.
- Stamps, P. L. (1997). *Nurses and work satisfaction: An index for measurement*. Chicago, IL: Health Administration Press.
- Stamps, P. L., & Piedmonte, E. B. (1986). *Nurses and work satisfaction: an index for measurement*. Ann Arbor, MI: Health Administration Press.
- Taunton, R. L., Bott, M. J., Koehn, M. L., Miller, P., Rindner, E., Pace, K., ... Dunton, N. (2004). The NDNQI-adapted index of work satisfaction. *Journal of Nursing Measurement*, 12(2), 101–122.
- Tennant, A., & Conaghan, P. G. (2007). The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied and what should one look for in a Rasch paper? *Arthritis Care and Research*, 57(8), 1358–1362.
- Tourangeau, A., Hall, L., Doran, D., & Petch, T. (2006). Measurement of nurse job satisfaction using the McCloskey/Mueller Satisfaction Scale. *Nursing Research*, 55(2), 128–136.
- Wade, G. H., Osgood, B., Avino, K., Bucher, G., Bucher, L., Foraker, T., ... Sirkowski, C. (2008). Influence of organizational characteristics and caring attributes of managers on nurses' job enjoyment. *Journal of Advanced Nursing*, 64(4), 344–353.
- Weiss, D. J., Dawis, R. V., & England, G. W. (1967). *Manual for the Minnesota satisfaction questionnaire*. Minneapolis: Work Adjustment Project, Industrial Relations Center, University of Minnesota, Minneapolis.
- Wielenga, J. M., Smit, B. J., & Unk, K. A. (2008). A survey on job satisfaction among nursing staff before and after introduction of the NIDCAP model of care in a level III NICU in the Netherlands. *Advances in Neonatal Care*, 8(4), 237–245.
- Yamashita, M., Takase, M., Wakabayashi, C., Kuroda, K., & Owatari, N. (2009). Work satisfaction of Japanese public health nurses: Assessing validity and reliability of a scale. *Nursing and Health Sciences*, 11(4), 417–421.
- Zangaro, G. A., & Soeken, K. L. (2005). Meta-analysis of the reliability and validity of part B of the index of work satisfaction across studies. *Journal of Nursing Measurement*, 13(1), 7–22.
- Zurmehly, J. (2008). The relationship of educational preparation, autonomy and critical thinking to nursing job satisfaction. *Journal of Continuing Education in Nursing*, 39(10), 453–460.