Construct Validity and Measurement Invariance of the Research Skill Inventory

Sanit Srikoon[1]

Tassanee Bunterm[1*]

Teerachai Nethanomsak[1]

Tang Keow Ngang[2]

[1]Faculty of Education, Khon Kaen University, Khon Kaen 40002, Thailand
**Corresponding Author Email: tassaneebun@gmail.com*
[2]International College, Khon Kaen University, Khon Kaen 40002, Thailand

Doi:10.5901/mjss.2016.v7n3p366

Abstract

This paper aimed to validate the developmental process of instrument to create an instructional design model. Therefore this study was designed to test the construct validity of research skill inventory and examine the measurement invariance of research skills across gender and grade. The effect sampling size was 1969 samples. However the actual samples participated was 1712, given a response rate as 86.94 percent. Data was analyzed to test construct validity and multi-group confirmatory factor analysis across gender and grades. Confirmatory factor analysis supports the six models consists of male, female, Grade 7, Grade 8, Grade 9, and total CFA models. MGCFA was comprised of measurement invariance and structural invariance. Findings indicated that measurement invariance had been proved to support the four models as non-invariance namely gender, Grade 7-Grade 8, Grade 8-Grade 9, and Grade 7-Grade 9 models. Structural invariance testing had been proved to support the three invariance models except Grade 8-Grade 9 model.

Keywords: *Research skills, construct validity, measurement invariance*

1. Introduction

A growing body of literature suggests research skills (RSs) are related to student's learning, attitude, research competency, critical thinking, and academic achievement (Wannapiroon, 2014; Srikoon, Bunterm, Samranjai & Wattanathorn, 2014; Zehra, Hassaan & Mushtaq, 2015; Stappenbelt, 2013). Consequently RSs are recognized as important contributors to children behavior and have become the focus of learning process or so called as 'student as researcher' (Elizabeth & Grant, 2013; Leat & Reid, 2012; Wilkin, 2014).

2. Problem Statement

Despite its wide-spread mention, little research has investigated the RSs, and only a few studies have defined the factor structure of RSs. For example, Stokking*, Schaaf, Jaspers, and Erkens (2004) refer RSs as teachers' assessment for students' RSs were complex skills but were undefined them. Although Kiley, Moyes and Clayton (2009), and Czarneski (2013) reported RSs in their research but they were not clearly defined it, too. To date, only a few studies have referred the factor structure of RSs (Meerah et al., 2012; Willison, 2007).

3. Significant of the Study

Construct validity (CV) is designed to measure the theoretical latent construct of variables in order to assess the quality of measures of a behavioral model (Hair, 2006). In addition, confirmatory factor analysis (CFA) is utilized when research has some knowledge of the underlying latent variable structure (Bryne, 2012). The combination of CFA results and CV tests can obtain a better understanding of the quality of these measures (Hair, 2006). Therefore, many researchers used CFA to confirm CV, like Canivez and Sproul (2005), Canivez, Neitzel, and Martin (2005), Martin and Marsh (2008), and Weis and Smenner (2007). In conclusion, CFA is the approximately method to test how well the measured variables represent

the construct.

One of the social science variables for managing learning in classrooms were gender and grade (Lowe, 2014). To date, gender and grade are considered for learning measurement condition (Bas & Yurdabakan, 2012; Harrell-Williams, Sorto, Pierce, Lesser, & Murphy, 2014; Lowe, 2014; Siegling, Furnham, & Petrides, 2015).

4. Conceptual Framework

Research skill inventory is the self-assessment inventory for diagnosing face validity. It consists of 35 items of seven factors. These seven RSs are Research Questioning Skill (ROS), Research Literacy Skill (RLS), Research Design Skill (RDS), Research Collecting Data Skill (RCDS), Research Organizing Data Skill (RODS), Research Conclusion Skill (RCS), and Research Result Presenting Skill (RRPS). Figure 1 shows the conclusion of the seven research skills. Each statement of research instrument is using a 5-point Likert scale namely 1 represents 'Does not perform', 2 represents 'Does not really perform', 3 represents 'Neutral, not sure', 4 represents 'Perform somewhat', and 5 represents 'Absolutely perform'

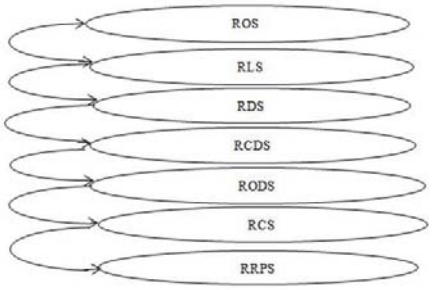

Figure 1: Conceptual framework

5. Aim of the Study

RSs should be precisely measured to examine the nature of construct and to make valid explanations for each individual's social activity. This study aimed to validate the RSs and further address measurement invariance across gender and grade. The specific purposes of this study are as follow:

1. To test the CV of the RSs with the consistency idea in each gender and grade.
2. To examine the measurement invariance of RSs across gender and grade.

6. Research Methodology

6.1 Sample of the Study

A total of 1969 (N) samples were drawn from Grade 7 (n_1 = 648), Grade 8 (n_2 = 660), and Grade 9 (n_3 =661). Owing to incomplete response, a total of 257 cases have been excluded for further analyzing namely 2 from Grade 7, 83 from Grade 8, and 172 from Grade 9, given the actual samples as 1712. As a result, there were 646 (37.73%) Grade 7 cases consisted of 415 males (24.24%) and 231 females (13.49%), 577 (33.71%) Grade 8 cases encompassed 206 males (12.04%) and 371 females (21.67%), and 489 (28.56%) Grade 9 cases included 204 males (11.92%) and 285 females (16.64%).

6.2 Method of the Study

The rapid growing of advanced methodology would be provided the ability of researchers to analyze the CV across group so called 'multiple group invariance confirmatory factor analysis' (MGCFA). MGCFA is used to compare latent variable means, variances, and co-variances across groups while holding measurement parameter invariant (Asparouhov & Muthen, 2014; Wang & Wang, 2012). In other word, MGCFA consists of two different kinds of invariance namely

measurement invariance and structural invariance.

Measurement invariance is tested first and followed by structural invariance. Four different level of measurement invariance are measurement configurable invariance, weak measurement invariance, strong measurement invariance, and strict measurement invariance. Meanwhile structural invariance consists of three levels namely invariance of factor variance, invariance of factor covariance, and factor mean invariance.

MGCFA can give the elaborately results about metric invariance, scalar invariance, and invariance of structural parameter across group (Wang & Wang, 2012; Dimitrov, 2010). There are two reasons of using MGCFA namely to ensure underlying construct has the same theoretical structure and psychological meaning across the groups (Bryne, 2008) as well as to examine the equivalence of factorial validity across gender and grade.

6.3 Data Analysis

Other than descriptive statistics, CFA Mplus program version 6 is utilized to test the hypothesized of seventh-factor model consisted of 35 items in all samples, for each gender and grade. Covariance structures are fitted with the maximum likelihood method. The model fitted indicator is evaluated by means of several fit indices. Universally, the model is considered acceptable when probability value (p-value) >0.05, value of ratio between chi-square statistic and degree of freedom (χ^2/df) in 2:1 (Hair, 2006), the Comparative Fit Index (CFI) ≥ 0.90 and good when ≥0.95 (Bentler, 1992; Hu & Bentler, 1999) and Tucker-Lewis coefficient (TLI) ≥ 0.95 is good fit (Sharma, Mukherjee, Kumar, & Dillon, 2005). Moreover, Standardized Root Means Square Residual (SRMR) should not exceed 0.08 for a good fit (Hu & Bentler, 1999). Furthermore, Root Means Square Error of Approximation (RMSEA) value ≤ 0.06 are considered indicative of a good fit, ≤ 0.08 of fair fit, between 0.08 and 0.01 of mediocre fit and >0.01 of poor fit (Hu & Bentler, 1999; MacCallum, Browne, & Sugawara, 1996).

Analyzing of MGCFA were based on the Satorra-Bentler scaled chi-square statistics and usual the Maximum Likelihood ($ML_{x2statistic}$) because it serves a correction of chi-square when distributional assumption are violated. In testing model can evaluate the goodness-of-fit of model by multiple criteria including CFI, RMSEA, 90 percent confidence interval, and SRMR (Byrne, 2008). CFI values in the range of 0.92 to 0.94 may also be considered as reasonable indicators of good model fit (Marsh, Hau & Wen, 2004). RMSEA values less than 0.05 indicates good fit, and value as high as 0.08 represent reasonable errors of approximation in the popular (M.W. Browne, 1992). For completeness, Bryne (2008) guides to report the 90 percent confidence interval provide for RMSEA (Steiger, 1990). Lastly, SRMR values range from 0.00 to 1.00, with a value less than 0.08 being indicative of a well-fitting model (Hu & Bentler, 1999).

7. Research Findings

Findings are presented in three sections namely descriptive findings, CV, and measurement invariance of gender and grade.

7.1 Descriptive Findings

Descriptive statistics like mean value ($\bar{x}$), standard deviation (*S.D.*), skewness (sk) and kurtosis (ku) are used to analyze the 35 items of the RSs. All items in the research skill inventory were distributed normally according to gender, grade and the total when compared with criteria because skewness values between 2 and -2 ad kurtosis values between 7 and -7 have been estimated as normal as presented in Table 1.

7.2 Construct Validity

CFA was used to evaluate the goodness of fit. Model fit was assessed using χ^2, χ^2/df , CFI, TLI, SRMR and RMSEA. Fit indices of all CFA models, including total, male, female, grade 7, grade 8, and grade 9 models are presented in Table 2. Results showed that all indicators indicated that there was a goodness of fit between the empirical data and the hypothetical measurement model for all models. In other word, all the *p*-value in CFA models are not significant (.053-.067), χ^2/df values are fallen in 2:1 (1.095-1.102), all CFA ≥ 0.95 indicate the good fit (.988-.997). Similarity all the TLIs ≥ 0.95 are good fit (.987-.996). Moreover all the RMSEAs are ≤ 0.06 are considered indicative of a good fit (.008-.014). Finally all the SRMRs≤ 0.08 are accepted for a good fit (.019-.033) too.

Individual parameter estimated for all the CFA models were also examined. Table 3 lists the standardized factor

loadings each latent variables of any CFA models. Meanwhile Table 4 shows the factor correlations between latent variables and factor loading of items in each CFA models were all positive and statistically significant.

Table 1. Mean, standard deviation, skewness, and kurtosis

Items	gender								Grade												Total (n=1712)			
	Male (n=641)				Female (n=1071)				Grade7(n=646)				Grade 8(n=577)				Grade 9(n=489)							
	$\bar{x}$	S.D.	sk	ku	$\bar{x}$	S.D.	sk	ku	$\bar{x}$	S.D.	sk	ku	$\bar{x}$	S.D.	sk	Ku	$\bar{x}$	S.D.	sk	ku	$\bar{x}$	S.D.	sk	ku
Item 1	3.73	.84	.02	-.48	3.77	.82	.00	-.62	3.64	.84	.20	-.55	3.90	.80	-.08	-.80	3.74	.81	-.13	-.20	3.76	.83	.01	-.57
Item 2	3.59	.89	-.02	-.37	3.70	.88	-.17	-.29	3.42	.87	.09	.09	3.79	.88	-.23	-.41	3.82	.83	-.21	-.52	3.66	.88	-.11	-.34
Item 3	3.64	.85	.00	-.40	3.67	.82	-.01	-.42	3.53	.82	.24	-.15	3.79	.81	-.12	-.62	3.68	.85	-.20	-.27	3.66	.83	-.01	-.41
Item 4	3.55	.85	-.05	-.18	3.50	.83	.16	-.18	3.45	.84	.23	-.02	3.56	.82	.00	-.06	3.57	.85	-.04	-.44	3.52	.84	.08	-.19
Item 5	3.23	.88	.15	.00	3.21	.82	.38	.28	3.15	.82	.34	.52	3.30	.82	.33	.02	3.21	.87	.19	-.07	3.22	.84	.28	.16
Item 6	3.17	.84	.06	.29	3.16	.85	.09	.22	3.17	.81	.22	.67	3.20	.87	.01	.05	3.11	.86	.00	.00	3.16	.85	.08	.24
Item 7	3.39	.84	.01	.01	3.37	.81	.12	.08	3.32	.77	.23	.50	3.44	.85	.02	-.29	3.36	.85	-.03	.05	3.37	.82	.08	.05
Item 8	3.41	.84	.03	.19	3.37	.77	.11	.10	3.36	.79	.20	.44	3.47	.77	.10	-.01	3.33	.82	-.03	-.01	3.39	.79	.08	.16
Item 9	3.34	.82	.15	-.08	3.35	.78	.14	.04	3.33	.78	.20	.20	3.38	.80	.07	-.04	3.32	.82	.17	-.18	3.34	.80	.15	-.01
Item 10	3.30	.87	.10	.03	3.35	.83	.20	.08	3.29	.85	.20	.25	3.38	.84	.15	-.13	3.31	.86	.11	.09	3.33	.85	.15	.07
Item 11	3.40	.83	-.03	.21	3.44	.83	.05	-.05	3.36	.81	.25	.14	3.51	.83	-.23	.25	3.41	.84	.03	-.08	3.42	.83	.02	.05
Item 12	3.38	.87	.13	-.30	3.39	.84	.07	-.03	3.35	.83	.32	-.02	3.44	.87	-.07	-.20	3.37	.84	.00	-.12	3.39	.85	.09	-.14
Item 13	3.33	.92	.09	-.28	3.40	.90	.13	-.51	3.39	.87	.25	-.21	3.47	.92	-.02	-.61	3.25	.92	.11	-.39	3.38	.91	.11	-.41
Item 14	3.24	.86	.22	.03	3.22	.85	.11	.07	3.23	.84	.18	.35	3.28	.84	.16	-.15	3.18	.88	.14	-.08	3.23	.85	.15	.05
Item 15	3.35	.86	.06	-.05	3.27	.86	.09	-.06	3.27	.84	.13	.13	3.31	.88	-.03	-.04	3.33	.87	.14	-.30	3.30	.86	.08	-.06
Item 16	3.39	.91	-.05	-.12	3.37	.86	.06	-.15	3.33	.83	.23	.00	3.42	.89	-.10	-.13	3.39	.91	-.09	-.22	3.38	.88	.02	-.14
Item 17	3.37	.89	.06	-.24	3.33	.87	.14	-.10	3.34	.87	.19	.05	3.36	.92	.05	-.39	3.33	.84	.08	-.12	3.34	.88	.11	-.16
Item 18	3.44	.85	.07	-.02	3.50	.82	.14	-.36	3.39	.81	.29	.08	3.60	.84	.10	-.57	3.46	.83	-.14	-.06	3.48	.83	.11	-.21
Item 19	3.46	.85	.11	-.36	3.53	.86	-.04	-.18	3.41	.84	.20	.08	3.61	.84	-.09	-.33	3.50	.88	-.09	-.44	3.50	.85	.02	-.26
Item 20	3.39	.89	.07	-.20	3.43	.87	.06	-.25	3.39	.87	.23	-.12	3.47	.85	-.01	-.30	3.37	.92	-.04	-.29	3.41	.88	.06	-.23
Item 21	3.40	.92	-.04	-.11	3.58	.89	-.05	-.44	3.42	.85	.16	-.02	3.61	.93	-.16	-.43	3.51	.92	-.22	-.28	3.51	.90	-.05	-.29
Item 22	3.42	.89	-.10	-.15	3.56	.86	.09	-.47	3.45	.82	.29	-.11	3.59	.92	-.19	-.48	3.49	.89	-.11	-.15	3.51	.87	.01	-.30
Item 23	3.43	.88	-.09	.07	3.51	.86	.06	-.38	3.41	.83	.18	.21	3.57	.90	-.09	-.45	3.47	.88	-.14	-.20	3.48	.87	.00	-.19
Item 24	3.39	.90	-.03	-.18	3.39	.87	.11	-.37	3.35	.86	.07	.06	3.45	.89	.09	-.65	3.38	.90	-.01	-.29	3.39	.88	.05	-.29
Item 25	3.37	.87	.11	-.04	3.34	.90	.04	-.14	3.29	.87	-.02	.35	3.42	.88	.12	-.32	3.34	.92	.09	-.43	3.35	.89	.06	-.10
Item 26	3.43	.87	-.09	.03	3.47	.86	.19	-.30	3.37	.86	.26	.15	3.60	.85	-.03	-.32	3.41	.86	.00	-.22	3.46	.86	.08	-.16
Item 27	3.43	.86	-.02	-.27	3.46	.86	.01	-.06	3.38	.80	.06	.38	3.55	.91	-.14	-.27	3.42	.87	.02	-.44	3.45	.86	-.01	-.14
Item 28	3.41	.87	.08	-.02	3.44	.86	-.02	-.12	3.42	.83	.11	.25	3.53	.88	-.03	-.40	3.33	.88	-.05	-.10	3.43	.86	.01	-.09
Item 29	3.47	.89	-.01	-.13	3.47	.86	.12	-.30	3.42	.86	.12	.00	3.56	.87	.01	-.40	3.42	.88	.06	-.27	3.47	.87	.07	-.23
Item 30	3.43	.84	.12	-.09	3.46	.86	.19	-.31	3.41	.80	.44	-.11	3.55	.90	-.12	-.22	3.39	.84	.18	-.20	3.45	.85	.17	-.23
Item 31	3.52	.92	-.02	-.40	3.70	.89	-.20	-.39	3.60	.87	-.03	-.09	3.77	.93	-.33	-.51	3.53	.91	-.06	-.49	3.64	.91	-.14	-.41
Item 32	3.39	.94	-.04	-.40	3.50	.90	.00	-.36	3.40	.90	.11	-.13	3.62	.90	-.18	-.41	3.34	.93	.00	-.46	3.46	.92	-.02	-.36
Item 33	3.48	.86	.13	-.33	3.55	.86	.02	-.26	3.47	.83	.13	.09	3.60	.88	-.13	-.38	3.49	.86	.17	-.53	3.52	.86	.06	-.30
Item 34	3.41	.85	.16	-.21	3.46	.83	.13	-.16	3.38	.79	.29	.30	3.54	.84	.05	-.35	3.41	.89	.06	-.43	3.44	.84	.14	-.19
Item 35	3.45	.93	-.03	-.23	3.53	.86	.20	-.50	3.45	.86	.25	-.13	3.59	.91	-.10	-.42	3.47	.88	.11	-.37	3.50	.89	.09	-.34

Note: $SE_{sk(malel)}$= 0.097, $SE_{ku(male)}$=0.193,$SE_{sk(femalel)}$= 0.075, $SE_{ku(female)}$=0.149,$SE_{sk(total)}$= 0.059, $SE_{ku(total)}$=0.118, $SE_{sk(grade\ 7)}$= 0.096, $SE_{ku(grade\ 7)}$=0.192,$SE_{sk(grade\ 8)}$= 0.102, $SE_{ku(grade\ 8)}$=0.203,$SE_{sk(grade\ 9)}$= 0.110, $SE_{ku(grade\ 9)}$=0.220

Table 2. Goodness of fit test of alternative models

Group	x^2	df	p-value	x^2/df	CFI	TLI	RMSEA	SRMR
Total	543.058	495	.067	1.097	.997	.996	.008	.019
Male	563.071	514	.066	1.095	.989	.988	.012	.031
Female	548.330	499	.063	1.099	.996	.995	.010	.021
Grade 7	575.115	522	.053	1.102	.992	.990	.013	.028
Grade 8	559.747	511	.067	1.095	.990	.988	.013	.032
Grade 9	567.267	515	.055	1.101	.988	.987	.014	.033

7.3 Measurement Invariance across Gender and Grade

Mean least square (ML) was used to test the measurement invariance. Factorial invariance across gender or grade was presented in the measurement invariance and structural invariance. Additionally, measurement invariance consists of testing configural invariance, weak measurement invariance, and strict measurement invariance. Weak measurement invariance is defined as invariance of factor loading across group. Factor loading invariance is also called metric invariance. Strong measurement invariance is defined as invariance of both factor loading and item intercept across group. Strict measurement invariance requires metric invariance, scalar invariance, and error variance invariance across group. Structural invariance across gender consists of the invariance of factor invariance, the invariance of factor covariance, and the factor mean invariance. Testing invariance of factor invariance implies that the factors or latent constructs have the same distribution dispersion across group. Testing covariance represents the association between the two factors when factor invariance holds. Testing mean variance is comparing the factor mean across group.

7.3.1 Measurement Invariance across Gender

The model fit similarity for males and females as such χ^2 (df = 535) = 1058.210, p<0.001; CFI=.916; TLI=.906; SRMR=.038, RMSEA=.039; RMSEA 90% C.I. = [.036,.043] for males; and χ^2 (df=536) = 1413.328, p<0.001; CFI= .942; TLI=.935; SRMR= .030, RMSEA= .039; RMSEA 90% C.I. = [.037,.042] for females. The males and females baseline models for analyzing configural model fit the data well: χ^2 (df=1071) = 2471.538, p<0.001; CFI=.934; TLI= .927; SRMR= .033, RMSEA= .039; RMSEA 90% C.I. = [.037,.041] (Table 5).

Weak measurement indicates of testing metric invariance is Δx^2 (df=28) = 26.538, p=0.543 which is not statistically significant. Since one of the factor loading of each factor or 'marker items', including item 1, item 6, item 11, item 16, item 21, item 26, and item 31, are not tested in the likelihood ratio (LR) test so they must be tested invariance of the marker items factor loading. The difference in model x^2 statistics between model with and without equality restriction; Δx^2 (df=7) = 6.128, p=0.523 which is not statistically significant, indicating that the factor loading of maker items are invariance across gender.

Strong measurement with LR test of scalar invariance shows as such Δx^2 (df=63) = 114.571, p=0.000 which is statistically significant but ΔCFI = 0.002 is lesser than 0.01. Thus, both factor loading and item intercept of the RSI are non-invariance across gender. Strict measurement with LR test of two nested models of error invariance are Δx^2 (df=98) = 395.44, p=0.000 of no equality restriction on item invariance which are statistically significant. The corresponding Δx^2 (df=35) = 280.869, p=0.000 of equality restriction on item invariance which are statistically significant. Thus, results indicate that item invariances are non-invariant across groups.

Table 3. Parameter estimates of standardized model results

Items	Male		Female		Grade 7		Grade 8		Grade 9		Total	
	β(SE)	R^2	β(SE)	R^2	β(SE)	R^2	β(SE)	R^2	β(SE)	R^2	β(SE)	R^2
RQS												
1	.583** (.039)	.339	.463** (.030)	.214	.519** (.037)	.270	.503** (.038)	.253	.478** (.050)	.228	.501** (.024)	.251
2	.523** (.036)	.273	.512** (.031)	.262	.516** (.038)	.267	.439** (.040)	.193	.525** (.043)	.276	.522** (.024)	.273
3	.590** (.034)	.348	.615** (.028)	.378	.641** (.036)	.410	.478** (.040)	.228	.563** (.037)	.317	.612** (.022)	.375
4	.576** (.032)	.332	.589** (.024)	.347	.623** (.030)	.388	.578** (.034)	.335	.565** (.037)	.319	.587** (.019)	.345
5	.557** (.035)	.310	.694** (.027)	.482	.616** (.034)	.379	.597** (.036)	.356	.629** (.038)	.395	.641** (.022)	.412
RLS												
6	.413** (.043)	.170	.572** (.025)	.327	.553** (.033)	.306	.488** (.039)	.238	.513** (.041)	.263	.520** (.022)	.270
7	.552** (.035)	.305	.667** (.022)	.445	.616** (.030)	.379	.638** (.031)	.408	.589** (.038)	.347	.629** (.019)	.395
8	.536** (.034)	.287	.667** (.024)	.445	.612** (.035)	.375	.595** (.030)	.354	.604** (.037)	.364	.621** (.020)	.386
9	.569** (.032)	.323	.628** (.023)	.395	.608** (.029)	.369	.589** (.034)	.347	.620** (.034)	.384	.602** (.019)	.363
10	.590** (.032)	.349	.693** (.021)	.481	.642** (.028)	.412	.617** (.033)	.380	.671** (.031)	.450	.664** (.019)	.441
RDS												
11	.564** (.035)	.318	.611** (.023)	.373	.636** (.029)	.405	.521** (.037)	.272	.584** (.034)	.341	.577** (.020)	.333
12	.547** (.031)	.300	.672** (.021)	.451	.646** (.027)	.417	.543** (.032)	.295	.669** (.030)	.447	.615** (.018)	.379
13	.595** (.031)	.354	.661** (.019)	.437	.675** (.024)	.455	.567 (.031)	.321	.603** (.035)	.364	.626** (.017)	.392
14	.606** (.033)	.368	.600** (.024)	.360	.633** (.029)	.401	.565** (.034)	.320	.515** (.042)	.266	.581** (.020)	.337
15	.488** (.040)	.238	.596** (.023)	.356	.617** (.030)	.381	.557** (.036)	.310	.494** (.043)	.244	.554** (.021)	.307

Items	Male		Female		Grade 7		Grade 8		Grade 9		Total	
	β(SE)	R^2	β(SE)	R^2	β(SE)	R^2	β(SE)	R^2	β(SE)	R^2	β(SE)	R^2
RCDS												
16	.628** (.029)	.395	.651** (.021)	.423	.685** (.027)	.470	.602** (.030)	.362	.661** (.030)	.437	.643** (.018)	.413
17	.612** (.032)	.374	.619** (.022)	.383	.681** (.025)	.464	.563** (.030)	.317	.663** (.037)	.439	.603** (.019)	.363
18	.613** (.029)	.375	.652** (.020)	.425	.698** (.023)	.487	.584** (.030)	.341	.650** (.030)	.422	.626** (.018)	.392
19	.550** (.032)	.302	.625** (.023)	.390	.652** (.029)	.425	.536** (.034)	.287	.566** (.036)	.321	.594** (.019)	.353
20	.521** (.034)	.271	.620** (.024)	.385	.644** (.030)	.415	.507** (.037)	.257	.563** (.036)	.317	.587** (.020)	.344
RODS												
21	.531** (.035)	.282	.610** (.024)	.372	.557** (.034)	.311	.626** (.028)	.392	.572** (.040)	.327	.592** (.020)	.351
22	.564** (.033)	.318	.654** (.022)	.427	.693** (.024)	.480	.600** (.032)	.360	.634** (.036)	.402	.640** (.018)	.410
23	.512** (.038)	.262	.628**(.022)	.394	.656** (.027)	.430	.552** (.035)	.305	.605** (.037)	.366	.599** (.020)	.359
24	.494** (.036)	.244	.599** (.023)	.359	.690** (.025)	.476	.501** (.034)	.251	.540** (.039)	.292	.582** (.019)	.338
25	.492** (.038)	.242	.603** (.022)	.364	.665** (.027)	.442	.502** (.037)	.252	.532** (.038)	.283	.578** (.020)	.334
RCS												
26	.586** (.034)	.344	.673** (.019)	.453	.677** (.026)	.458	.567** (.028)	.321	.611** (.033)	.374	.661** (.018)	.438
27	.597** (.029)	.356	.704** (.018)	.496	.680** (.023)	.462	.662** (.032)	.438	.644** (.033)	.414	.664**(.017)	.441
28	.547** (.034)	.299	.685** (.020)	.470	.632** (.029)	.400	.618** (.035)	.382	.663** (.030)	.439	.633** (.018)	.400
29	.560** (.029)	.314	.690** (.019)	.476	.653** (.029)	.426	.602** (.034)	.362	.638** (.030)	.408	.658** (.017)	.432
30	.618** (.033)	.382	.635** (.022)	.403	.666** (.027)	.443	.562** (.036)	.315	.580** (.030)	.337	.626** (.019)	.391
RRPS												
31	.073** (.026)	.508	.672** (.021)	.452	.654** (.028)	.428	.650** (.029)	.423	.647** (.032)	.419	.645** (.018)	.417
32	.624** (.029)	.389	.681** (.019)	.463	.718** (.026)	.515	.640** (.030)	.410	.595** (.035)	.354	.616** (.019)	.380
33	.593** (.029)	.351	.698** (.020)	.487	.672** (.028)	.452	.685** (.026)	.469	.640** (.031)	.410	.645** (.018)	.416
34	.656** (.030)	.431	.696** (.020)	.485	.699** (.026)	.489	.602** (.032)	.363	.679** (.031)	.461	.679** (.017)	.461
35	.617** (.037)	.380	.643** (.022)	.413	.629** (.015)	.341	.567** (.036)	.321	.618** (.037)	.381	.607** (.020)	.369

Table 4. Factor correlation for 7 factors CFA of RSI

Model	Male							Female							Total						
Factor	1.	2.	3.	4.	5.	6.	7.	1.	2.	3.	4.	5.	6.	7.	1.	2.	3.	4.	5.	6.	7.
1.RQS	1							1							1						
2.RLS	.805**	1						.712**	1						.724**	1					
3.RDS	.713**	.813**	1					.735**	.904**	1					.739**	.880**	1				
4.RCDS	.744**	.836**	.833**	1				.713**	.803**	.951**	1				.726**	.807**	.935**	1			
5.RODS	.822**	.829**	.812**	.945**	1			.768**	.771**	.929**	.945**	1			.762**	.762**	.893**	.926**	1		
6.RCS	.696**	.766**	.741**	.787**	.904**	1		.681**	.764**	.868**	.885**	.933**	1		.687**	.754**	.842**	.853**	.885**	1	
7.RRPS	.720**	.725**	.712**	.751**	.826**	.824**	1	.665**	.696**	.798**	.861**	.917**	.910**	1	.705**	.719**	.817**	.856**	.901**	.898**	1
Model	Grade 7							Grade 8							Grade 9						
Factor	1.	2.	3.	4.	5.	6.	7.	1.	2.	3.	4.	5.	6.	7.	1.	2.	3.	4.	5.	6.	7.
1.RQS	1							1							1						
2.RLS	.795**	1						.756**	1						.730**	1					
3.RDS	.770**	.947**	1					.783**	.891**	1					.744**	.840**	1				
4.RCDS	.714**	.866**	.934**	1				.758**	.763**	.969**	1				.727**	.761**	.878**	1			
5.RODS	.810**	.904**	.915**	.930**	1			.782**	.694**	.878**	.924**	1			.719**	.673**	.791**	.858**	1		
6.RCS	.750**	.904**	.923**	.871**	.928**	1		.715**	.751**	.858**	.934**	.946**	1		.695**	.654**	.770**	.776**	.799**	1	
7.RRPS	.722**	.811**	.850**	.863**	.875**	.921**	1	.710**	.618**	.740**	.758**	.898**	.875**	1	.698**	.673**	.752**	.844**	.863**	.856**	1

**p<0.01

Table 5. Measurement invariance summary fit statistics across gender

	Model	Comparison	x^2	x^2degree of freedom	CFI	TLI	SRMR	RMSEA	RMSEA 90% C.I.	Δx^2	Δdf	ΔCFI	x^2difference test (significance value)
Testing measurement invariance	*Configural invariance*												
	Male baseline model (M_1)	-	1058.210	535	0.916	0.906	0.038	0.039	[0.036,0.043]	-	-	-	-
	Female baseline model (M_2)	-	1413.328	536	0.942	0.935	0.030	0.039	[0.037,0.042]	-	-	-	-
	Testing configural invariance (M_3)	-	2471.538	1071	0.934	0.927	0.033	0.039	[0.037,0.041]	-	-	-	-
	Weak measurement invariance(M_4)		2498.046	1099	0.934	0.929	0.035	0.039	[0.037,0.041]				
	Factor loading invariance	M_4-M_3								26.538	28	-	0.543
	Maker items' factor loading invariance	M_5-M_6	-	-	-	-	-	-	-	6.128	7		0.523
	-without equality restriction on factor loading (M_5)	-	2471.538	1071	0.934	0.927	0.033	0.039	[0.037,0.041]	-	-	-	-
	- with equality restriction on factor loading (M_6)	-	2477.666	1078	0.934	0.927	0.033	0.039	[0.037,0.041]	-	-	-	-
	Strong measurement invariance (invariance of factor loadings and item intercepts) (M_7)		2586.109	1134	0.932	0.928	0.036	0.039	[0.037,0.041]				
		M_7-M_3								114.571	63	0.002	0.000
	Strict measurement invariance (error variance invariance) (M_8)		2866.978	1169	0.920	0.919	0.050	0.041	[0.039,0.043]				
	- no equality restriction on item invariance	M_8-M_3								395.44	98	-	0.000
	- equality restriction on item invariance	M_8-M_7	-	-	-	-	-	-	-	280.869	35	-	0.000
Testing structural invariance	*Factor variance invariance* (M_9)		2607.808	1141	0.931	0.928	0.046	0.039	[0.037,0.041]				
		M_9-M_3								136.270	70	.003	3.653
	Factor covariance invariance (M_{10})		2634.134	1155	0.931	0.928	0.047	0.039	[0.037,0.041]				
		M_{10}-M_3								162.596	84	.003	5.976
	Factor mean invariance	RQS(-0.034,P=0.55), RLS(0.000,P=0.997), RDS(-0.010,P=0.878),RCDS(-0.040,P=0.504), RODS(-0.150,P=0.015), RCS(-0.061,P=0.314), RRPS(-0.178,P=0.003)											

Results of structural invariance show that Δx^2 (df=70) = 136.270, p=3.653 which is not statistically significant, Δ CFI=0.003<0.01. Therefore factor variance invariance across gender. Finding covariance shows that Δx^2 (df=84) = 162.596, p=5.976 which is not statistically significant, Δ CFI=0.003<0.01. Therefore factor covariance invariance across gender. Finally, findings showed that RQS (-0.034, p=0.557), RLS (0.000, p=0.997), RDS (-0.010, p=0.878), RCDS (-0.040, p=0.504), RCS (-0.061, p=0.314) are significantly lower in females than in males, while the factor mean of RODS (-0.150, p=0.015) and RRPS (-0.178, p=0.003) are not significantly different between the gender.

7.3.2 Measurement Invariance across Grade 7 and Grade 8

The model fit similarity for Grade 7 and Grade 8 as such χ^2 (df=535) = 1022.911, p<0.001; CFI=.948; TLI= .942; SRMR= .031, RMSEA= .038; RMSEA 90% C.I. = [.034,.041] for Grade 7; and χ^2 (df=533) = 990.289, p<0.001; CFI= .926; TLI= .917; SRMR= .038, RMSEA= .039; RMSEA 90% C.I. = [.035,.042] for Grade 8. The Grade 7 and Grade 8 baseline models for analyzing configural model fit the data well: χ^2 (df = 1071) = 2013.201, p<0.001; CFI=.939; TLI= .932; SRMR= .034, RMSEA= .038; RMSEA 90% C.I. = [.035,.041] (Table 6).

Weak measurement indicates of testing metric invariance is Δx^2 (df=28) = 53.487, p=0.003 which is statistically

significant. Since one of the factor loading of each factor or 'marker items', including item 1, item 6, item 11, item 16, item 21, item 26, and item 31, are not tested in the LR test so they must be tested invariance of the marker items factor loading. The difference in model x^2 statistics between model with and without equality restriction; Δx^2 (df=7) = 18.942, *p*=0.008 which is statistically significant, indicating that the factor loading of maker items are non-invariance across Grade 7 and Grade 8.

Strong measurement with LR test of scalar invariance shows as such Δx^2 (df=63) = 178.300, *p*=5.915 which is not statistically significant but ΔCFI=0.007 is smaller than 0.01. Thus, both factor loading and item intercept of the RSI are variance across Grade 7 and Grade 8. Strict measurement with LR test of two nested models of error invariance are Δx^2 (df=98) = 382.073, *p*=0.000 of no equality restriction on item invariance which are statistically significant. The corresponding Δx^2 (df=35) = 203.773, *p*=0.000 of equality restriction on item invariance which are statistically significant. Thus, results indicate that item invariances are non-invariant across Grade 7 and Grade 8.

Results of structural invariance show that Δx^2 (df=70) = 188.068, *p*=9.420 which is not statistically significant, Δ CFI=0.007<0.01. Therefore factor variance invariance across Grade 7 and Grade 8. Finding covariance shows that Δx^2 (df=84) = 219.939, *p*=4.052 which is not statistically significant, Δ CFI = 0.009<0.01. Therefore factor covariance invariance across Grade 7 and Grade 8. Finally, findings showed that RQS (0.210, *p*=0.000), RLS (0.075, *p*=0.011), RDS (0.074, *p*=0.022), RCDS (0.128, *p*=0.000), RODS (0.143, *p*=0.000), RCS (0.160, *p*=0.000) are significantly higher in Grade 8 than in Grade 7.

7.3.3 Measurement Invariance across Grade 8 and Grade 9

The model fit similarity for Grade 8 and Grade 9 as such χ^2 (df=533) = 990.289, *p*<0.001; CFI= .926; TLI= .917; SRMR= .038, RMSEA= .039; RMSEA 90% C.I. = [.035,.042] for Grade 8; and χ^2 (df=531) = 880.814, *p*<0.001; CFI= .936; TLI= .929; SRMR= .038, RMSEA= .037; RMSEA 90% C.I. = [.033,.041] for Grade 9. The Grade 8 and Grade 9 baseline models for analyzing configural model fit the data well: χ^2 (df=1064) = 1878.104, *p*<0.001; CFI= .931; TLI= .923; SRMR= .038, RMSEA= .038; RMSEA 90% C.I. = [.035,.041] (Table 7).

Weak measurement indicates of testing metric invariance is Δx^2 (df=28) = 21.196, *p* = 0.817 which is not statistically significant. Since one of the factor loading of each factor or 'marker items', including item 1, item 6, item 11, item 16, item 21, item 26, and item 31, are not tested in the LR test so they must be tested invariance of the marker items factor loading. The difference in model x^2 statistics between model with and without equality restriction; Δx^2 (df=7) = 3.154, *p*=0.870 which is not statistically significant, indicating that the factor loading of maker items are invariance across Grade 8 and Grade 9.

Strong measurement with LR test of scalar invariance shows as such Δx^2 (df=63) = 94.771, *p*=0.006 which is statistically significant but ΔCFI=0.003 is smaller than 0.01. Thus, both factor loading and item intercept of the RSI are non-variance across Grade 8 and Grade 9. Strict measurement with LR test of two nested models of error invariance are Δx^2 (df=98) = 139.157, *p*=0.004 of no equality restriction on item invariance which are statistically significant. The corresponding Δx^2 (df=35) = 44.386, *p*=0.134 of equality restriction on item invariance which are not statistically significant. Thus, results indicate that item variances are non-invariant across groups.

Results of structural invariance show that Δx^2 (df=70) = 97.911, *p*=0.015 which is statistically significant, Δ CFI=0.002<0.01. Therefore factor variance non-invariance across Grade 8 and Grade 9. Finding covariance shows that Δx^2 (df=84) = 119.140, *p*=0.007 which is statistically significant, Δ CFI=0.003<0.01. Therefore factor covariance non-invariance across Grade 8 and Grade 9. Finally, findings showed that RLS (-0.080, *p*=0.014), RDS (-0.093, *p*=0.009), RCDS (-0.095, *p*=0.017), RODS (-0.090, *p*=0.029), RCS (-0.145, *p*=0.000), and RRPS (-0.192, *p*=0.000), are significantly lower in Grade 8 than in Grade 9 while the factor mean of RQS (-0.057, *p*=0.064) are not significantly different between the two samples.

7.3.4 Measurement Invariance across Grade 7 and Grade 9

The ML was used for testing measurement invariance. The model fit similarity for Grade 7 and Grade 9 as such χ^2 (df=535) = 1022.911, *p*<0.001; CFI= .948; TLI= .942; SRMR= .031, RMSEA= .038; RMSEA 90% C.I. = [.034,.041] for Grade 7; and χ^2 (df=531) = 887.814, *p*<0.001; CFI= .936; TLI= .929; SRMR= .038, RMSEA= .037; RMSEA 90% C.I. = [.033,.041] for Grade 9. The Grade 7 and Grade 9 baseline models for analyzing configural model fit the data well: χ^2 (df=1066) = 1910.726, p<0.001; CFI= .944; TLI= .937; SRMR= .035, RMSEA= .037; RMSEA 90% C.I.= [.035,.040] (Table 8).

Weak measurement indicates of testing metric invariance is Δx^2 (df=28) = 28.631, p=0.431 which is not statistically significant. Since one of the factor loading of each factor or 'marker items', including item 1, item 6, item 11, item 16, item 21, item 26, and item 31, are not tested in the LR test so they must be tested invariance of the marker items factor loading. The difference in model x^2 statistics between model with and without equality restriction; Δx^2 (df=7) = 6.7, p=0.461 which is not statistically significant, indicating that the factor loading of maker items are invariance across Grade 7 and Grade 9.

Table 6. Measurement invariance summary fit statistics across grade 7 and grade 8

	Model	Comparison	x^2	x^2degree of freedom	CFI	TLI	SRMR	RMSEA	RMSEA 90% C.I.	Δx^2	Δdf	ΔCFI	x^2difference test (significance value)
Testing measurement invariance	*Configural invariance*												
	Grade 7 baseline model (M_1)	-	1022.911	535	0.948	0.942	0.031	0.038	[0.034,0.041]	-	-	-	-
	Grade 8 baseline model (M_2)	-	990.289	533	0.926	0.917	0.038	0.039	[0.035,0.042]	-	-	-	-
	Testing configural invariance (M_3)	-	2013.201	1068	0.939	0.932	0.034	0.038	[0.035,0.041]	-	-	-	-
	Weak measurement invariance(M_4)		2066.688	1096	0.938	0.932	0.039	0.038	[0.036,0.041]				
	Factor loading invariance	M_4-M_3	-	-	-	-	-	-	-	53.487	28	-	.003
	Maker items' factor loading invariance	M_5-M_6	-	-	-	-	-	-	-	18.942	7	-	.008
	-without equality restriction on factor loading (M_5)	-	2013.201	1068	0.939	0.932	0.034	0.038	[0.035,0.041]	-	-	-	-
	- with equality restriction on factor loading (M_6)	-	2032.143	1075	0.938	0.932	0.036	0.038	[0.036,0.041]	-	-	-	-
	Strong measurement invariance (invariance of factor loadings and item intercepts) (M_7)	-	2191.501	1131	0.932	0.928	0.046	0.039	[0.037,0.042]				
		M_7-M_3	-	-	-	-	-	-	-	178.3	63	.007	5.915
	Strict measurement invariance (error variance invariance) (M_8)	-	2395.274	1166	0.921	0.919	0.049	0.042	[0.039,0.044]	-	-	-	-
	- no equality restriction on item invariance	M_8-M_3	-	-	-	-	-	-	-	382.073	98	-	.000
	- equality restriction on item invariance	M_8-M_7	-	-	-	-	-	-	-	203.773	35	-	.000
Testing structural invariance	*Factor variance invariance* (M_9)	-	2201.269	1138	0.932	0.928	0.053	0.039	[0.037,0.042]	-	-	-	-
		M_9-M_3	-	-	-	-	-	-	-	188.068	70	.007	9.420
	Factor covariance invariance (M_{10})	-	2233.140	1152	0.930	0.928	0.058	0.039	[0.037,0.042]	-	-	-	-
		M_{10}-M_3	-	-	-	-	-	-	-	219.939	84	.009	4.052
	Factor mean invariance	RQS(0.210,P=0.000), RLS(0.075,P=0.011), RDS(0.074,P=0.022),RCDS(0.128,P=0.000), RODS(0.143,P=0.000), RCS(0.160,P=0.000), RRPS(0.168,P=0.000)											

Table 7. Measurement invariance summary fit statistics across grade 8 and grade 9

	Model	Comparison	x^2	x^2degree of freedom	CFI	TLI	SRMR	RMSEA	RMSEA 90% C.I.	Δx^2	Δdf	ΔCFI	x^2difference test (significance value)
Testing measurement invariance	*Configural invariance*												
	Grade 8 baseline model (M_1)	-	990.289	533	0.926	0.917	0.038	0.039	[0.035,0.042]	-	-	-	-
	Grade 9 baseline model (M_2)	-	887.814	531	0.936	0.929	0.038	0.037	[0.033,0.041]	-	-	-	-
	Testing configural invariance (M_3)	-	1878.104	1064	0.931	0.923	0.038	0.038	[0.035,0.041]	-	-	-	-
	Weak measurement invariance(M_4)	-	1899.300	1092	0.931	0.925	0.040	0.037	[0.034,0.040]				
	Factor loading invariance	M_4-M_3	-	-	-	-	-	-	-	21.196	28	-	.817
	Maker items' factor loading invariance	M_5-M_6	-	-	-	-	-	-	-	3.154	7	-	.870
	-without equality restriction on factor loading (M_5)	-	1878.104	1064	0.931	0.923	0.038	0.038	[0.035,0.041]	-	-	-	-
	- with equality restriction on factor loading (M_6)	-	1881.253	1071	0.931	0.924	0.038	0.038	[0.035,0.040]	-	-	-	-
	Strong measurement invariance (invariance of factor loadings and item intercepts) (M_7)	M_7-M_3	1972.875	1127	0.928	0.924	0.044	0.038	[0.035,0.040]	94.771	63	.003	.006
	Strict measurement invariance (error variance invariance) (M_8)	-	2017.261	1162	0.927	0.926	0.047	0.037	[0.034,0.040]	-	-	-	-
	- no equality restriction on item invariance	M_8-M_3	-	-	-	-	-	-	-	139.157	98	-	.004
	- equality restriction on item invariance	M_8-M_7	-	-	-	-	-	-	-	44.386	35	-	.134

	Model	Comparison	x^2	x^2degree of freedom	CFI	TLI	SRMR	RMSEA	RMSEA 90% C.I.	Δx^2	Δdf	ΔCFI	x^2difference test (significance value)
Testing structural invariance	*Factor variance invariance* (M_9)	-	1976.015	1134	0.929	0.925	0.046	0.037	[0.035,0.040]	-	-	-	-
		M_9-M_3	-	-	-	-	-	-	-	97.911	70	.002	.015
	Factor covariance invariance (M_{10})	-	1997.244	1148	0.928	0.925	0.046	0.037	[0.035,0.040]	-	-	-	-
		M_{10}-M_3	-	-	-	-	-	-	-	119.14	84	.003	.007
	Factor mean invariance	RQS(-.057,P=.064), RLS(-.080,P=.014), RDS(-.093,P=.009),RCDS(-.095,P=.017), RODS(-.090,P=.029), RCS(-.145,P=.000), RRPS(-.192,P=.000)											

Table 8. Measurement invariance summary fit statistics across grade 7 and grade 9

	Model	Comparison	x^2	x^2degree of freedom	CFI	TLI	SRMR	RMSEA	RMSEA 90% C.I.	Δx^2	Δdf	ΔCFI	x^2difference test (significance value)
Testing measurement invariance	*Configural invariance*												
	Grade 7 baseline model (M_1)	-	1022.911	535	0.948	0.942	0.031	0.038	[0.034,0.041]	-	-	-	-
	Grade 9 baseline model (M_2)	-	887.814	531	0.936	0.929	0.038	0.037	[0.033,0.041]	-	-	-	-
	Testing configural invariance (M_3)	-	1910.726	1066	0.944	0.937	0.035	0.037	[0.035,0.040]	-	-	-	-
	Weak measurement invariance		1939.357	1094	0.943	0.939	0.037	0.037	[0.034,0.040]				
	Factor loading invariance (M_4)	M_4-M_3	-	-	-	-	-	-	-	28.631	28	-	.431
	Maker items' factor loading invariance	M_5-M_6	-	-	-	-	-	-	-	6.7	7	-	.461
	-without equality restriction on factor loading (M_5)	-	1910.726	1066	0.944	0.937	0.035	0.037	[0.035,0.040]	-	-	-	-
	- with equality restriction on factor loading (M_6)	-	1917.426	1073	0.944	0.937	0.035	0.037	[0.035,0.040]	-	-	-	-
	Strong measurement invariance (invariance of factor loadings and item intercepts) (M_7)	-	2068.080	1129	0.937	0.934	0.040	0.038	[0.036,0.041]	157.354	63	.007	4.948
		M_7-M_3	-	-	-	-	-	-	-	157.354	63	.007	4.948
	Strict measurement invariance (error variance invariance) (M_8)	-	2223.595	1164	0.929	0.928	0.048	0.040	[0.038,0.043]	-	-	-	-
	- no equality restriction on item invariance	M_8-M_3								312.869	98	-	.000
	- equality restriction on item invariance	M_8-M_7	-	-	-	-	-	-	-	155.515	35	-	.000
Testing structural invariance	*Factor variance invariance* (M_9)	M_9-M_3	2072.211	1136	0.937	0.934	0.042	0.038	[0.036,0.041]	161.485	70	.007	3.439
	Factor covariance invariance (M_{10})	M_{10}-M_3	2093.762	1150	0.937	0.935	0.047	0.038	[0.035,0.041]	183.036	84	.004	2.530
	Factor mean invariance	RQS(.140,P=.000), RLS(-.007,P=.828), RDS(-.009,P=.795),RCDS(.034,P=.378), RODS(.055,P=.113), RCS(-.002,P=.950), RRPS(-.012,P=.755)											

Strong measurement with LR test of scalar invariance shows as such Δx^2 (df=63) = 157.354, p=4.948 which is not statistically significant but ΔCFI=0.007 is much smaller than 0.01. Thus, both factor loading and item intercept of the RSI are variance across Grade 7 and Grade 9. Strict measurement with LR test of two nested models of error invariance are Δx^2 (df=98) = 312.869, p=0.000 of no equality restriction on item invariance which are statistically significant. The corresponding Δx^2 (df=35) = 155.515, p=0.000 of equality restriction on item invariance which are not statistically significant. Thus, results indicate that item variances are non-invariant across Grade 7 and Grade 9.

Results of structural invariance show that Δx^2 (df=70) = 161.485, p=3.439 which is not statistically significant, Δ CFI=0.007<0.01. Therefore factor variance invariance across Grade 7 and Grade 9. Finding covariance shows that Δx^2 (df=84) = 183.036, p=2.530 which is not statistically significant, Δ CFI=0.004<0.01. Therefore factor covariance invariance across Grade 7 and Grade 9. Finally, findings showed that RQS (0.140, p=0.000) are significantly higher in Grade 9 than Grade 7. Meanwhile RLS, RDS, RCDS, RODS, RCS, and RRPS are not statistically significant different between the two groups of samples.

8. Discussion

On this line of reasoning, RSs is considered as the most important because students who possessed high RSs are more likely to engage in learning outcomes and achievement at higher level (Zehra, Hassaan, & Mushtaq, 2015). In addition, investigating CV and measurement invariance of RSs across gender and grade provide a further understanding of how RSs can be measured and evaluated.

This study was conducted purposively to examine the validity of the RSI with samples of Grade 7 to Grade 9. The validity of research skills consist of male, female, Grade 7, Grade 8, Grade 9, and total CFA model which are the most well fit for all models. Results from this study provide evidence about the validity of the confirming RSI. Moreover, researchers investigate the measurement invariance across gender and grade. Conclusion is shown in Table 9.

Testing measurement invariance can show metric invariance is mostly invariance except metric invariance across Grade 7 and Grade 8 is non-invariance. In other word, factor loading is generally having same pattern except observed indicator variables measure is different factors across Grade 7 and Grade 8. Secondly, test scalar invariance is mostly invariance except scalar invariance across gender.

Table 9. Conclusion of measurement invariance testing

Group	weak measurement invariance	strong measurement invariance	strict measurement invariance	factor variance invariance	factor covariance invariance
gender	invariance	non-invariance	non-invariance	invariance	invariance
grade7-grade 8	non-invariance	invariance	non-invariance	invariance	invariance
grade8-grade 9	invariance	invariance	non-invariance	non-invariance	non-invariance
grade7-grade 9	invariance	invariance	non-invariance	invariance	invariance

As a result, factor loading and item intercept is generally same pattern except some scalar is different factors across gender. Thirdly, test error variance invariance showed that are non-invariance for all. However, rejecting the hypothesis of strict measurement invariance does not mean RSI are problematic but only implies that the amount of variances of the observed item responses that are unexplained by their underlying factors are non-invariance across group (Wang & Wang, 2012). Therefore RSI is a measurement invariance across Grade 8-Grade 9 and across Grade 7-Grade 9. In contrast, across gender and across Grade 7-Grade 8 is non-invariance of measurement invariance. However, the empirical data are confirmed with theory of research skills. In other word, RSI does the most construct validity.

The structural invariance is testing for structural equivalence, focus on the factor covariance. Some researchers may also wish to test for the equality of the factor variance (Byrne, 2008). However our study is limited to test both factor covariance and factor variance. As conclusion, results revealed that only test invariance across Grade 8 and Grade 9 is non-invariance.

9. Acknowledgements

This work was supported by the Higher Education Research Promotion and National Research University Project of Thailand, Office of the Higher Education Commission, through the Cluster of Research to Enhance the Quality of Basic Education.

References

Asparouhov, T., & Muthen, B. (2014). Multiple-group factor analysis alignment. Structural Equation Modeling: A Multidisciplinary Journal, 21(4), 495-508. doi: 10.1080/10705511.2014.919210.

Bas, A.U., & Yurdabakan, I. (2012). Factor structure of the reactive-proactive aggression questionnaire in Turkish children and gender, grade-level, and socioeconomic status differences in reactive and proactive aggression. Journal of Psycho-educational Assessment, 30(3), 284-297. doi: 10.1177/0734282911428892.

Bentler, P.M. (1992). On the fit of models to covariances and methodology to the Bulletin. Psychological Bulletin, 112(3), 400-404.

Bryne, B.M. (2008). Testing for multigroup equivalence of a measuring instrument: A walk through the process. Comprobando la equivalencia multigrupal de un instrument de medida: pasos del proceso, 20(4), 872-882.

Byrne, B.M. (2012). Structural Equation Modeling with Mplus: Basic Concepts, Applications, and Programming. New York: Routledge.

Canivez, G.L., Neitzel, R., & Martin, B.E. (2005). Construct validity of the Kaufman Brief Intelligence Test, Wechsler Intelligence Scale for Children – third edition, and Adjustment Scales for Children and Adolescents. Journal of Psycho-educational Assessment,

23(1), 15-34. doi: 10.1177/073428290502300102.
Canivez, G.L., & Sproul, K. (2005). Assessing the construct validity of the adjustment scales for children and adolescents. Journal of Psycho-educational Assessment, 23(1), 3-14. doi: 10.1177/073428290502300101.
Czarneski, D. (2013). A First-year course that teachers research skills. PRIMUS, 23(5), 487-497. doi: 10.1080/10511970.2013.764365.
Dimitrov, D.M. (2010). Testing for factorial invariance in the context of construct validation. Measurement & Evaluation in Counseling & Development, 43(2), 121-149. doi: 10.1177/0748175610373459.
Elizabeth, V., & Grant, B.M. (2013). The spirit of research has changed: Reverberations from researcher identities in managerial times. Higher Education Research & Development, 32(1), 122-135. doi: 10.1080/07294360.2012.751362.
Hair, J.F. (2006). Multivariate data analysis. (6th ed.). Upper Saddle River, N.J.: Prentice-Hall International.
Harrell-Williams, L.M., Sorto, M.A., Pierce, R.L., Lesser, L.M., & Murphy, T.J. (2014). Validation of scores from a new measure of pre-service teachers' self-efficacy to teach statistics in the middle grades. Journal of Psycho-educational Assessment, 32(1), 40-50. doi: 10.1177/0734282913486256.
Hu, L.T., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Alternatives. 6(1), Structural Equation Modeling, 1-55. doi: 10.1080/10705519909540118.
Kiley, M., Moyes, T., & Clayton, P. (2009). To develop research skills: Honors programs for the changing research agenda in Australian universities. Innovations in Education and Teaching International, 46(1), 15-25. doi: 10.1080/14703290802646164.
Leat, D., & Reid, A. (2012). Exploring the role of student researchers in the process of curriculum development. Curriculum Journal, 23(2), 189-205. doi: 10.1080/09585176.2012.678691.
Lowe, P.A. (2014). Should test anxiety be measured differently for males and females? Examination of measurement bias across gender on measures of text anxiety for middle and high school, and college students. Journal of Psycho-educational Assessment, 33(3), 238-246. doi: 10.1177/0734282914549428.
MacCallum, R.C., Browne, M.W., & Sugawara, H.M. (1996). Power analysis and determination of sample size for covariance structure modeling. Psychological Methods, 1(2), 130-149.
Marsh, H.W., Hau, K.T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in over generalizing Hu and Bentler's (1999) findings. Structural Equation Modeling, 11(3), 320-341.
Martin, A.J., & Marsh, H.W. (2008). Workplace and academic buoyancy psychometric assessment and construct validity amongst school personnel and students. Journal of psycho-educational assessment, 26(2), 168-184. doi: 10.1177/0734282907313767.
Meerah, T.S.M., Osman, K., Zakaria, E., Ikhsan, Z.H., Krish, P., Lian, D.K.C., & Mahmod, D. (2012). Developing an instrument to measure research skills. Procedia-Social and Behavioral Sciences, 60, 630-636. doi: 10.1016/j.sbspro.2012.09.434.
M.W. Browne, R.C. (1992). Alternative ways of assessing model fit. Sociological methods & Research, 21(2). doi: 10.1177/004912419 2021002005.
Sharma, S., Mukherjee, S., Kumar, A., & Dillon, W.R. (2005). A simulation study to investigate the use of cutoff values for assessing model fit in covariance structure models. Journal of Business Research, 58, 935-943. doi: 10.1016/j.jbusres.2003.10.007
Siegling, A.B., Furnham, A., & Petrides, K.V. (2015). Trait emotional intelligence and personality gender-invariant linkages across different measures of the big five. Journal of Psycho-educational assessment, 33(1), 57-67. doi: 10.1177/0734282914550385.
Srikoon, S., Bunterm, T., Samranjai, J., & Wattanathorn, J. (2014). Research synthesis of research-based learning for education in Thailand. Procedia-Social and Behavioral Sciences, 116, 913-917.
Stappenbelt, B. (2013). The effectiveness of the teaching-research nexus in facilitating student learning. Engneering Education, 8(1), 111-121. doi: 10.11120/ened.2013.00002.
Steiger, J.H. (1990). Structural model evaluation and modification: An interval estimation approach. Multivariate Behavioral Research, 25(2), 173-180. doi: 10.1207/s15327906mbr2502_4.
Stokking*, K., Schaaf, M., Jaspers, J., & Erkens, G. (2004). Teachers' assessment of students' research skills. British Educational Research Journal, 30(1), 93-116. doi: 10.1080/0141192030010016299 83.
Wang, J., & Wang, X. (2012). Structural equation modeling: Applications using Mplus. West Sussex, UK: Higher Education Press.
Wannapiroon, P. (2014). Development of research-based blended learning model to enhance graduate students' research competency and critical thinking skills. Procedia-Social and Behavioral Sciences, 136, 486-490.
Weis, R., & Smenner, L. (2007). Construct validity of the behavior assessment system for children (BASC) self-report of personality evidence from adolescent referred to residential treatment. Journal of Psycho-educational Assessment, 25(2), 111-126. doi: 10.1177/0734282906293762.
Wilkin, C.L. (2014). Enhancing the AIS curriculum: Integration of a research-led, problem-based learning task. Journal of Accounting Education, 32(2), 185-199. doi: 10.1016/j.jaccedu.20.14.04.001.
Willison, J.O.K. (2007). Commonly know, commonly not know, totally unknown: A framework for students becoming researchers. The Higher Education Research and Development, 26(4), 393-409.
Zehra, N., Hassaan, A., & Mushtaq, S. (2015). Research amongst junior and senior medical students: Comparison of knowledge, attitude and practice. Professional Medical Journal, 22(1), 121-126.