International Journal of Academic Research in Business and Social Sciences
2017, Vol. 7, No. 4
ISSN: 2222-6990

A Confirmatory Factor Analysis of Content Coverage Measure Using Multiply Imputed Datasets

Aszunarni Ayob
Faculty of Education, Universiti Kebangsaan Malaysia
Email: aszunarni@gmail.com

Ruhizan M. Yasin
Faculty of Education, Universiti Kebangsaan Malaysia
Email: ruhizan@ukm.edu.my

Hasnah Shuhaimi
Faculty of Information Management, Universiti Teknologi MARA (UiTM) Selangor, Malaysia
Email: hasnahshuhaimi@gmail.com

Saiful Farik Mat Yatin
Faculty of Information Management, Universiti Teknologi MARA (UiTM) Selangor, Malaysia
Email: farik@salam.uitm.edu.my

DOI: 10.6007/IJARBSS/v7-i4/2969 URL: http://dx.doi.org/10.6007/IJARBSS/v7-i4/2969

Abstract
This paper re-examined the validity of the content coverage measure, which is a proxy for opportunity to learn (OTL), by employing a confirmatory factor analysis (CFA) on five multiply-imputed datasets. Data for this study were drawn from the PISA 2012 Malaysian sample. Specifically, we used a sample of 4247 students from 135 Malaysian national secondary schools. The PISA 2012 content coverage measure comprised four constructs, namely *experience with applied mathematics tasks at school*, *experience with pure mathematics tasks at school*, *familiarity with mathematical concepts* and *experience with various types of problems at school*. Prior to conducting the CFA, missing data resulted from student questionnaire rotation design were multiply-imputed using predictive mean matching (PMM) estimation via R-package Multivariate Imputation by Chained Equations (MICE). Subsequently, we conducted the CFA using R-package lavaan.survey that incorporates multiply-imputed data and survey weights as well as non-normality of data through its Maximum Likelihood Robust (MLR) estimation. After a few cycles of theory-guided model specification involving deletion of several items with low factor loadings, examination of various fit indices, and inspections of Composite Reliability (CR) and Average Variance Extracted (AVE) values, results showed that the final congeneric CFA model for the content coverage measure provided good fit to the data.
Keywords: *confirmatory factor analysis, content coverage, multiple imputation, predictive mean matching, maximum likelihood robust*

1. Introduction

Over the past few decades, there have been an unceasing increase in studies on opportunity to learn (OTL) due to its significant effect on student achievement (Gearhart et al., 1999; Herman & Klein, 1997; Jaafar, 2006; Kurz, Elliott, Wehby, & Smithson, 2010; Martinez, Bailey, Kerr, Huang, & Beaurgard, 2010; McDonnell, 1995; OECD, 2012; Reeves, Carnoy, & Addy, 2013; Schmidt, Alexander, & Zoido, 2014; Stevens, 1996, 1993; Wang, 1998; Winfield, 1993). However, since teaching and learning processes are rather complex and multifaceted, various variables have been used as measures for the OTL (Gearhart et al., 1999; Herman & Klein, 1997; Jaafar, 2006; Kurz et al., 2010; Martinez et al., 2010; OECD, 2012; Reeves et al., 2013; Stevens, 1993, 1996; Wang, 1998; Winfield, 1993).

Specifically, in the international studies such as First International Mathematics Study (FIMS), Second International Mathematics Study (SIMS) and Trend in International Mathematics and Science Study (TIMSS), OTL has been measured using *content coverage.* Basically, in these international studies that involve various countries with different educational systems, content coverage has been used as a tool for assessing validity of the findings across countries (Chang, 1984; International Association for the Evaluation of Educational & Achievement, 2011).

The importance of content coverage as OTL measure is clearly evident when it has been continuously emphasized in the Programme for International Student Assessment (PISA), which was conducted by the Organisation for Economic Co-operation and Development (OECD). Precisely, in PISA the 2012, the content coverage was defined as mathematical content that were taught to students, and its internal consistency was assessed using Cronbach's alpha (OECD, 2012). In addition, the OECD also used correlations between certain constructs as to ensure cross-country validity of the content coverage measure (OECD, 2012).

Despite numerous efforts taken by the OECD to ensure validity of the content coverage measure, however little attention has been paid to the effect of employing multiply imputed datasets on the construct validity. Hence, further investigations are critically needed to cross-validate the construct validity of the content coverage measure by using the multiply imputed datasets, especially when the PISA 2012 was subjected to large amount of missing data due to student questionnaire rotation design (OECD 2012). Thus, the aim of this paper was to re-examine, via a confirmatory factor analysis (CFA), the validity of the constructs for the PISA 2012 content coverage measure, using multiply imputed datasets.

2. Method

2.1 Sample

The sample for this study was drawn from the PISA 2012 Malaysian data, consisting of 4247 students from 135 Malaysian national secondary schools. This sample represented 82% of the overall sample of 5197 Malaysian students from 164 schools. In the PISA 2012, all samples were used in the development of content coverage measure. On the contrary, in this study, we only utilized a subsample consisting of students from the Malaysian national secondary schools due to diverse characteristics between Malaysian national secondary schools and other types of Malaysian schools such as residential schools, technical schools and private schools.

3. Measures

Table 1 shows the constructs and items for the PISA 2012 content coverage measure (the list of items for all constructs can be obtained from OECD (2012)). All items were scored on four-point scale.

Table 1: Constructs and items for content coverage measure in PISA 2012

Construct	Description	Item
EXAPPLAM	*Experience with applied mathematics tasks at school*	ST61Q01, ST61Q02, ST61Q03, ST61Q04, ST61Q06, ST61Q08
EXPUREM	*Experience with pure mathematics tasks at school*	ST61Q05, ST61Q07, ST61Q09
FAMCON	*Familiarity with mathematical concepts*	ST62Q01, ST62Q02, ST62Q03, ST62Q06, ST62Q07, ST62Q08, ST62Q09, ST62Q10, ST62Q12, ST62Q15, ST62Q16, ST62Q17, ST62Q19
EXASSIGNM	*Experience with various types of problems at school*	ST73Q01, ST73Q02, ST74Q01, ST74Q02, ST75Q01, ST75Q02, ST76Q01, ST76Q02

Source: OECD, 2012

3.1 Data Imputation, Test of Multivariate Normality and Selection of Software for Confirmatory Factor Analysis

In the PISA 2012, the OECD employed rotated student context questionnaires to ensure inclusion of large coverage of context questions without burdening students with longer time to complete the questionnaires. The rotated questionnaire consisted of questions in the "common part", which were answered by all students, and questions in the "rotated part", which were answered by two thirds of the student sample. Hence, approximately 33% of the PISA 2012 data for student questionnaire were subjected to data loss (OECD, 2012). Precisely, for all items selected in this study, 35.05% of the data were missing and they were assumed to be Missing Completely at Random (MCAR) (Kaplan and Su 2016).

We therefore carefully imputed the data by using multiple imputation method that properly accounts for the uncertainty of missing data. Moreover, the multiple imputation method produces unbiased results if the missing data mechanism is MCAR (Kaplan and Su 2016). For the estimation of parameters, we employed Predictive Mean Matching (PMM), which is a semi-parametric imputation approach. The data imputation process was conducted using R-package Multivariate Imputation by Chained Equation (MICE), in which five imputed datasets were generated.

We then verified the distributions of the five imputed datasets via Mardia's Multivariate Normality test using R-package Multivariate Normality (MVN). Subsequently, we conducted the CFA by using R-package lavaan.survey that conveniently handles multiply imputed datasets, PISA survey weights as well as non-normality of data through its Maximum Likelihood Robust (MLR) estimation (Oberski, 2014).

3.2 PISA Sampling Design

Due to complex sampling design employed in the PISA 2012, the OECD used 80 replicate weights generated by balanced-repeated replication (BRR) using Fay (1989)'s method with ρ=0.5 (OECD, 2012). Hence, in order to take into account the sampling design, we employed designed-based approach (Oberski, 2014; Pornprasertmanit, Lee, & Preacher, 2014) by including these replicate weights in the CFA using a specific R-package lavaan.survey function (**svrepdesign**). The incorporation of PISA sampling design in the CFA analysis is deemed necessary because failure to take into account the sampling design might produce biased and inconsistent estimates (Oberski, 2014).

3.3 Confirmatory Factor Analysis

CFA is a branch of structural equation modelling (SEM). Unlike exploratory factor analysis (EFA) that is normally used for exploring the possible underlying constructs for a set of observed items, the CFA is a theoretically driven because it aims at testing hypothesized relationships between the observed items and the constructs based on specified underpinning theory. In this study, the CFA was used to determine the factor structure and dimensionality of the PISA 2012 content coverage measure. Specifically, the CFA was carried out in order to study the hypothesized relationship between the constructs for content coverage measure (i.e. *experience with applied mathematics* (EXPPLAM), *experience with pure mathematics* (EXPUREM), etc.) and their respective items.

For the CFA, we employed MLR estimation because previous studies showed that MLR produced reliable parameters for which standard errors and mean-adjusted chi-square were robust to non-normality (Brown, 2006; Oberski, 2014). The chi-square from the MLR is known as the scaled Satorra-Bentler chi-squared, with non-significant value indicates model fit (Brown, 2006; Oberski, 2014). Statistical calculations in the CFA were executed five times because this study involved five imputed datasets, and subsequently following Rubin (1976), the final estimators were obtained based on the pooled results.

To determine model fit, besides the chi-square value, we also used several other goodness-of-fit indices, such as standardized root mean square residual (SMSR), root mean square error of approximation (RMSEA), Comparative fit index (CFI) and Tucker-Lewis Index (TLI) (Brown, 2006; Hair, Black, Babin, & Anderson, 2010). Specifically, in this study, the Satorra-Bentler chi-square was expected to be significant due to large sample size (Hair et.al, 2010), and the following criteria were used to assess model fit: SMSR ≤ 0.08, RMSEA < 0.07, CFI > 0.90 and TLI > 0.90 (Hair et al., 2010).

In addition, we also assessed the validity of the constructs for content coverage measure based on the standardized factor loadings with 0.3 as a cut-off value (Sellin & Keeves, 1997). We also verified the validity of the constructs for content coverage measure based on the Average Variance Extracted (AVE), with 0.5 as a cut-off value (Hair et al., 2010). The AVE indicates the percentage of variance interpreted by the latent factors. Besides assessing the validity of the constructs for content coverage measure, we also assessed the reliability of the constructs through their Composite Reliability (CR) values. Explicitly, the CR value indicates internal consistency of a construct, with 0.6 as a cut-off value (Hair et al., 2010). However, in the process of building the CFA model, any decision made regarding model specification were guided by underlying theories in this study, as strongly emphasized by Hair et al. (2010).

4. Results

4.1 Multivariate Normality

Table 2 shows the results of Mardia's Multivariate Normality tests, in which the multivariate skewness and kurtosis for all five imputed datasets were significant, indicating that all imputed datasets did not follow multivariate normal distribution. Therefore, we used MLR estimation in the CFA in order to take into account the non-normality of data.

Table 2: Mardia's Multivariate Normality Test

Imputed Data	Skewness			Kurtosis		
	Value	χ^2 test	*p*-value	Value	χ^2 test	*p*-value
1	64.93	45962.80	0.00	3912.82	89.06	0.00
2	69.09	48906.22	0.00	3130.27	91.87	0.00
3	66.41	47010.61	0.00	3014.06	89.60	0.00
4	69.09	48906.22	0.00	3130.27	91.87	0.00
5	66.16	46828.34	0.00	3014.10	89.62	0.00

4.2 CFA for Opportunity to Learn

Since the content coverage measure was hypothesized to consist of four constructs, therefore we began the CFA by building a four-factor model. After a few series of theory-guided model specification, results showed that some of the items had to be eliminated due to either having low factor loadings, or low in both CR and AVE values. The eliminated items were as follows:

EXAPPLAM: ST61Q01, ST61Q06

EXPUREM: No item was eliminated

FAMCON: ST62Q01, ST62Q02, ST62Q03, ST62Q06,ST62Q12, ST62Q16, ST62Q17, ST62Q04, ST62Q11

EXASSIGNM: ST75Q02, ST76Q01 and ST76Q02

4.3 Assessing Model Fit and Construct Validity

We then assessed model fit by comparing two models, namely MLIDSW model that incorporated Maximum Likelihood estimation, multiply imputed datasets and PISA survey weights, and MLRIDSW model that incorporated Maximum Likelihood Robust, multiply imputed datasets and PISA survey weights.

As shown in Table 2, both non-normality of data and the PISA survey weights had considerable effects on the chi-square test of model fit. Specifically, the scaling corrections (or average generalized design effects) taking into account both non-normality and the PISA survey weights was 1.20.

Table 2 also shows the goodness-of-fit indices used to determine model fit as well as the CR and AVE values for each construct. As expected, the final CFA model for the content coverage measure had a significant Satorra-Bentler chi-square $\left(\mathrm{SB}\ \chi^2(146) = 1707.54, p < 0.01\right)$ due to large sample size (Hair et al., 2010). However, the values of all goodness-of-fit indices were beyond their respective cut-off values (Hair et al., 2010), indicating that the final CFA model for the content coverage measure provided good fit to the data (RMSEA = 0.05 (90% confidence interval = 0.048 to 0.052, CFI = 0.92; and TLI = 0.91; SMSR = 0.04).

Subsequently, we determined the validity of the final CFA model for the content coverage measure based on the standardized factor loadings, CR and AVE. Based on Table 4, all items had factor loadings exceeding 0.5, reflecting that all items were indicators for each construct (Sellin & Keeves, 1997). The CR values were in the range of 0.72 to 0.92, which were more than the cut-off value of 0.6 (Hair et al., 2010), while the AVE values were between 0.37 and 0.78. Specifically, the AVE values for EXAPPLAM, FAMCON and EXASSIGNM were 0.40, 0.37 and 0.44 respectively, which were less than the cut-off value of 0.5 (Hair et al., 2010). However, following Hair et al. (2010), we retained EXAPPLAM, FAMCON and EXASSIGNM in the final CFA model for the content coverage measure because they were believed to provide theoretical rationale that explained the content coverage measure (OECD, 2012).

Table 3: Assessment of model fit, reliability and validity for content coverage measure

Measure	Model	Chi-square	*Scaling correction	CFI	TLI	RMSEA	SMSR	Construct	Code Item	Factor Loading (> 0.3)	AVE (>0.5)	CR (> 0.6)
Content Coverage	MLIDSW	1681.931^{***}		0.94	0.93	0.05	0.03	EXAPPLAM	ST61Q02	0.62	0.40	0.72
	MLRIDSW	1405.873^{***}	1.20	0.93	0.92	0.05	0.03		ST61Q03	0.71		
									ST61Q04	0.62		
									ST61Q08	0.57		
								EXPUREM	ST61Q05	0.88	0.78	0.91
									ST61Q07	0.88		
									ST61Q09	0.86		
								FAMCON	ST62Q07	0.61	0.37	0.80
									ST62Q08	0.71		
									ST62Q09	0.64		
									ST62Q10	0.63		
									ST62Q15	0.58		
									ST62Q19	0.52		
									ST62Q13	0.55		
								EXASSIGNM	ST73Q01	0.60	0.44	0.79
									ST73Q02	0.58		
									ST74Q01	0.79		
									ST74Q02	0.76		
									ST75Q01	0.54		

MLIDSW – Maximum likelihood with multiply imputed data and PISA sampling weights; MLRIDSW – Maximum likelihood robust with multiply imputed data and PISA sampling weights; CR – Composite Reliability; AVE – Average Variance Extracted; *** p<0.001.

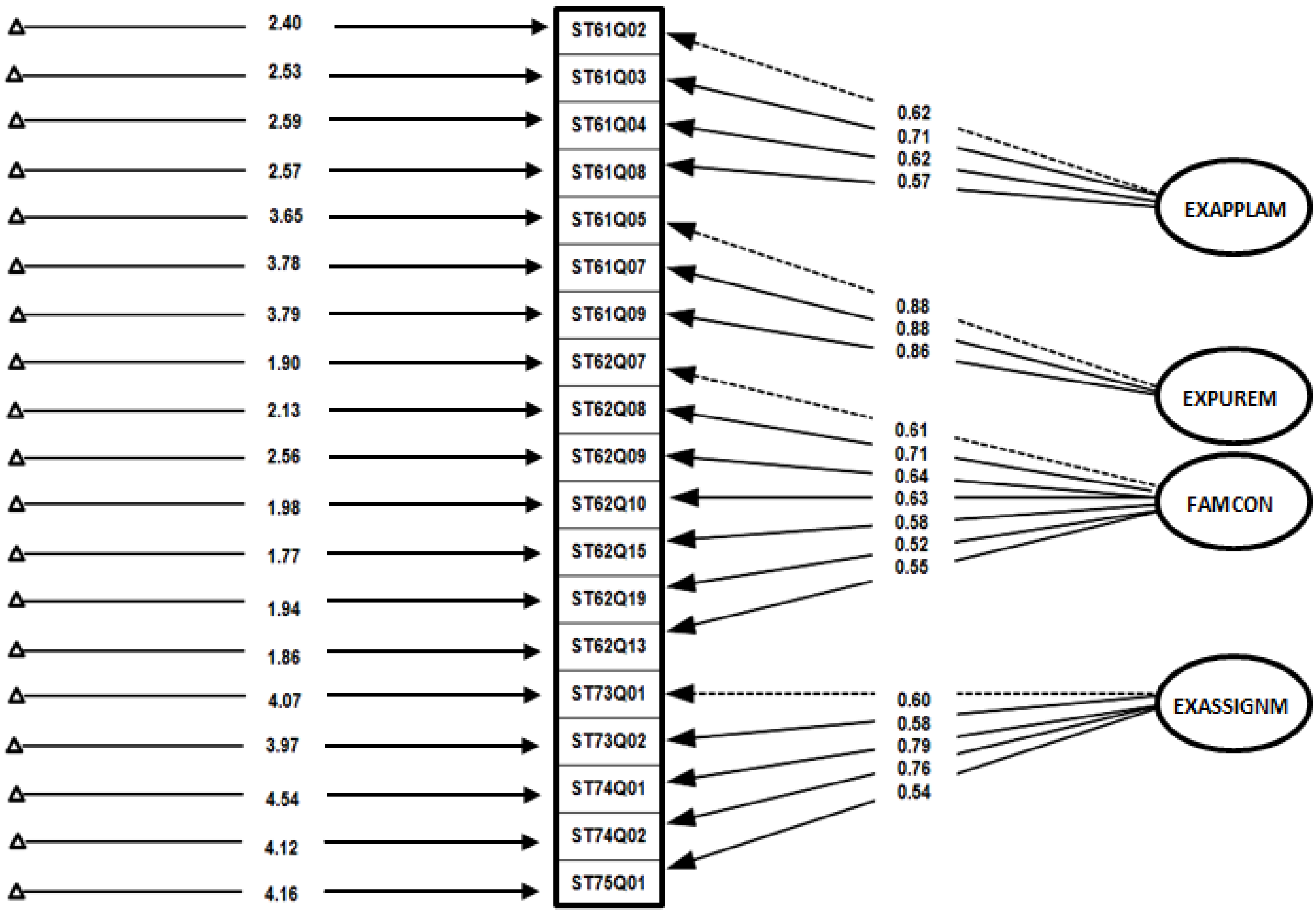

Figure 1: Confirmatory factor analysis for content coverage

5. Conclusion

In this study, we re-examined the construct validity of the PISA 2012 content coverage measure via a CFA by using sample of the Malaysian national secondary school students. Due to large missing data, we multiply imputed the original data prior to conducting the CFA. Specifically, we imputed five datasets via MI approach and PMM estimation using R-package MICE.

The PISA 2012 content coverage measure comprised four constructs. We validated the constructs for the content coverage measure by building a congeneric CFA model using MLR estimation in R-package lavaan.survey. After conducting a few cycles of theory-guided model specification involving elimination of several items with low loadings, examination of various goodness-of-fit indices and inspections of CR and AVE values, we found that the final CFA models for the content coverage measure provided good fit to the data.

Based on the final CFA model, even though the factor structure of the content coverage measure was still the same as in the original PISA 2012 results (OECD, 2012), however some of the items had to be eliminated due to having very low standardized factor loadings. Since the square of a factor loading indicates the percent of variance in a particular item explained by the construct, therefore such low factor loadings might signify problems in the factor structure of the construct (Brown, 2006; Hair et al., 2010). Thus, by including items with low standardized factor loadings in the scale development of PISA 2012 content coverage measure, the Cronbach's Alpha values and correlations between selected constructs estimated by OECD might be biased.

We therefore suggest that before conducting further statistical analysis involving the PISA 2012 content coverage measure, the original data shall be multiply imputed due to large missing data; the construct validity of the measures shall be re-examined by employing CFA on the multiply imputed datasets; and amendment to the measure should be made accordingly**.**

6. Limitation

This study only used MI via PMM approach. Further studies might use other imputation and estimation techniques such as Bayesian Bootstrap Predictive Mean Matching. In addition, due to nested nature of the PISA 2012 data, we used design-based approach by incorporating the survey weights into the single-level CFA. Future studies might employ model-based approach, such as multilevel CFA (MCFA) which includes higher level units into the models (Pornprasertmanit et al., 2014).

Acknowledgement

After the main body of paper please insert acknowledgement of all those (personals or institutions) that have helped in conducting this study.

Corresponding Author
*Aszunarni Ayob, Faculty of Education, Universiti Kebangsaan Malaysia, Bangi, Malaysia, aszunarni@gmail.com

References

Adams, R. J., Lietz, P., & Berezner, A. (2013). On the use of rotated context questionnaires in conjunction with multilevel item response models. *Large-Scale Assessments in Education*, *1*(5), 1–27.

Brown, T. A. (2006). Confirmatory factor analysis for applied research. In D. A. Kenny (Ed.), *Methodology in the social sciences*. New York, USA: The Guilford Press.

Chang, L. C. (1984). *The effects of teacher and student perceptions of opportunity to learn on achievement in beginning algebra in five countries*. University of Illinois.

Gearhart, M., Saxe, G. B. G. B., Seltzer, M., Schlackman, J., Ching, C. C. C., Nasir, N., Rhine, S., Sloan, T. F. (1999). Opportunities to learn fractions in elementary mathematics classrooms. *Journal for Research in Mathematics Education*, *30*(3), 286–315.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. . (2010). *Multivariate analysis* (7th ed.). Pearson Prentice Hall.

Herman, J. L., & Klein, D. C. D. (1997). *Assessing opportunity to learn: a California example* (Vol. 453). Los Angeles, CA.

International Association for the Evaluation of Educational, & Achievement. (2011). *TIMSS 2011 Assessment Frameworks*. *Trends in International Mathematics and Science Study 2011*.

Jaafar, S. B. (2006). An alternative approach to measuring opportunity-to-learn in high school classes. *Alberta Journal of Educational Research*, *52*(2), 107–126.

Kaplan, D., & Su, D. (2016). On matrix sampling and imputation of context questionnaires with implications for the generation of plausible values in large-scale assessments. *Journal of Educational and Behavioral Statistics*, *41*(1), 57–80.

Kurz, A., Elliott, S. N., Wehby, J. H., & Smithson, J. L. (2010). Alignment of the intended, planned, and enacted curriculum in general and special education and its relation to student achievement. *The Journal of Special Education*.

Martinez, J. F., Bailey, A. L., Kerr, D., Huang, B. H., & Beaurgard, S. (2010). *Measuring opportunity to learn and academic language exposure for english language learners in elementary science classrooms*. Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

McDonnell, L. M. (1995). Opportunity to learn as a research concept and a policy instrument. *Educational Evaluation and Policy Analysis*.

Oberski, D. L. (2014). lavaan.survey: An R package for complex survey analysis of structural equation models. *Journal of Statistical Software*, *57*(1), 1–27.

OECD. (2012). *PISA 2012 technical report*. Paris, Perancis: OECD Publishing.

Pornprasertmanit, S., Lee, J., & Preacher, K. J. (2014). Ignoring clustering in confirmatory factor analysis: some consequences for model fit and standardized parameter estimates. *Multivariate Behavioral Research*, *49*, 518–543.

Reeves, C., Carnoy, M., & Addy, N. (2013). Comparing opportunity to learn and student achievement gains in southern african primary schools: a new approach. *International Journal of Educational Development*, *33*(5), 426–435.

Rubin, D. B. (1976). Inference and missing data. *Biometrika*, *63*(3), 581–292.

Schmidt, W. H., Alexander, T. J., & Zoido, P. (2014). *Schooling matters: opportunity to learn in PISA 2012* (No. 95). Michigan State, USA: OECD and Michigan State University.

Sellin, N., & Keeves, J. P. (1997). Path analysis with latent variables. In *Educational Research, Methdology And Measurement: An International Handbook*. Oxford: Pergamon Press.

Stevens, F. I. (1993). Applying an opportunity-to-learn conceptual framework to the investigation of the effects secondary analyses of multiple-case-study data summary. *The Journal of Negro Education*, *62*(3), 232–248.

Stevens, F. I. (1996). The need to expand the opportunity to learn conceptual framework: should students, parents and school resources be included? In *The Annual Meeting of the American Educational Research Association*. New York: American Educational Research Association.

Wang, J. (1998). Opportunity to learn: the impacts and policy implications. *Educational Evaluation and Policy Analysis*.

Winfield, L. F. (1993). Investigating test content and curriculum content overlap to assess opportunity to learn. *The Journal of Negro Education*, *62*(3), 288–310.