A Multilevel Study of Content Coverage and Malaysian Students’ Mathematics Achievement

Aszunarni Ayob¹, Ruhizan M. Yasin¹, Nur Amirah Akmal Kamal Aris²

¹Faculty of Education, Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia
²INTI International University, Nilai, Negeri Sembilan, Malaysia

Received: February 21, 2017
Accepted: April 30, 2017

ABSTRACT

The purpose of this research was to identify the multilevel relationship between student- and school-level content coverage variables and Malaysian students’ Mathematics achievement. Specifically, student-level content coverage consisted of four variables—experience with applied mathematics tasks at school; experience with pure mathematics tasks at school; familiarity with mathematical concepts; and experience with various types of Mathematics problems at school. The multilevel analysis was conducted using Hierarchical Linear Modelling (HLM) 7.0 software. Prior to the multilevel analysis, missing data due to student questionnaire rotation design were multiply-imputed using Predictive Mean Matching estimation via R-package Multivariate Imputation by Chained Equation (MICE). Subsequently, confirmatory factor analysis on the multiply-imputed datasets was conducted using R-package lavaan survey. HLM analysis revealed that except experience with applied mathematics tasks at school that showed negative predictive effect, all other student- and school-level content coverage variables showed positive predictive effects on the Malaysian students’ Mathematics achievement.

KEYWORDS: Content Coverage, Student Achievement, Predictive Mean Matching, Multilevel Analysis, Hierarchical Linear Modelling.

INTRODUCTION

Opportunity to learn measures can help schools and teachers determine whether all students, regardless of their backgrounds, receive sufficient opportunity to learn what they need to learn, failing which students should not be held responsible for their achievements [10, 12, 15, 34, 35]. The importance of measuring opportunity to learn can be witnessed from continually increasing interest among researchers on the opportunity to learn studies over the past five decades [10, 12, 15, 19, 24, 25, 34, 35]. However, due to complex and multifaceted nature of teaching and learning processes, various proxy variables for opportunity to learn have been used in the previous studies [10, 12, 15, 17, 24].

Content coverage is one of the earliest proxy variables for opportunity to learn [5, 6, 12, 24, 31, 28]. In the context of international educational surveys such as First International Mathematics Study (FIMS) and Second International Mathematics Study (SIMS) which were conducted by the International Association for the Evaluation of Educational Achievement (IEA), content coverage was defined as content taught to students [6]. Inspired by significant relationships between content coverage and student achievement in these international studies, several scholars [20] started to include content coverage in their studies in early 1990s. Until now, content coverage as a proxy variable for opportunity to learn is still relevant [17, 24, 33], partly due to its importance within the context of educational system.

Even though content coverage has been recognized as one of the significant predictors of student achievement in previous studies, however, the emphasis is more on the formal curriculum [5, 6, 12, 17, 28, 33]. This study thus extends the previous studies by focusing on content coverage that is beyond the context of the formal curriculum. Precisely, the purpose of this study is to examine, via multilevel modelling, the influence of content coverage in the contexts of both formal curriculum and real-life application, on Malaysian students’ mathematics achievement. More specifically, this study addresses the following research questions:

1. To what extent do the student-level content coverage variables (experience with applied mathematics tasks at school, experience with pure mathematics tasks at school, familiarity with mathematical concepts and experience with various types of Mathematics problems at school) predict the Malaysian students’ mathematics achievement?
2. To what extent do the school-level content coverage variable (mean content coverage) predicts the Malaysian students’ mathematics achievement?

Corresponding Author: Aszunarni Ayob, Faculty of Education, Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia, E-mail: aszunarni@gmail.com
Content Coverage and Its Relationship with Mathematics Achievement

Content coverage is important as it echoes classroom implementation of the intended curriculum. Content coverage reflects what is known as implemented curriculum [17, 20]. Precisely, the intended curriculum refers to specific contents that should be taught to all students within the schooling period. Meanwhile, the implemented curriculum designates actual contents taught to students during classroom teaching and learning processes. Even though teachers are expected to teach all contents as outlined in the intended curriculum, however in reality, the implemented curriculum usually does not match the intended curriculum because of variation in the content coverage itself.

Many studies for the past three decades have showed significant variation in the content coverage among students [4, 9, 21, 29, 30]. This is partly due to inconsistencies in the contents taught by teachers. These inconsistencies are especially evident when teachers teach students with lower learning abilities. Teachers usually teach all contents specified in the intended curriculum to students with higher ability, whereas students with lower ability frequently receive less content coverage [4, 21, 29, 32].

In terms of difficulty level of the content coverage, lower ability students normally are less exposed to contents related to higher order thinking skills or real-life application. Teachers often presume that lower ability students are incapable to understand, and hence solve non-routine tasks as compared to higher ability students [4, 21, 29, 32]. Due to less exposure on the higher order thinking skills and non-routine problems, these lower ability students would often face difficulties while competing with higher ability students, and hence, are more likely to obtain lower scores in the examination.

The effects of content coverage variation on student achievement have been demonstrated in many studies over the past few decades [4, 21, 29, 32]. For example, using data from 1999 TIMSS-R, in [29] studied the extent to which students in different regions and states in the United States have opportunity to learn mathematics contents, and also examined the effects of these opportunities to learn on student achievement. Results showed significant variations in the content coverage across schools and districts. In addition, results also showed significant positive relationship between the content coverage and student achievement (after controlling for socioeconomic status and student pre-achievement).

The positive links between content coverage and student achievement have been revealed in a large volume of published studies [5, 6, 12, 17, 28, 33]. In general, students who are taught and exposed to more mathematics contents have a tendency to score higher in the mathematics assessment. Hence, we hypothesize that:

Hypothesis 1: Content coverage, in terms of (i) experience with applied mathematics tasks at school, (ii) experience with pure mathematics tasks at school, (iii) familiarity with mathematical concepts, and (iv) experience with various types of Mathematics problems at school will have positive relationships with Malaysian students’ Mathematics achievement.

School-Level Content Coverage and Mathematics Achievement

Based on Multilevel Organizational Theory (MOT) as proposed by [16], variables at an upper level may produce direct effects on the outcome. In the context of the educational system, school-level variables may affect student achievement. If the effects of school-level variables do not exist, then the relationships between the student-level variables and student achievement are the same for all schools [8, 13] or in other words, students of the same kind in a particular school would most likely attain the same results. However, in reality, students who share homogeneous characteristics in the same school often attain different achievements, and this may be caused by the direct effect of school-level variables on student achievement [8, 13, 16]. The effects of school-level variables on student achievement have been shown in the many previous studies [2, 7, 13].

In the context of opportunity to learn studies, the effect of school-level content coverage variable on the student achievement however is rarely tested. Therefore, the magnitude and direction of the direct effect of school-level content coverage variable are still questionable. Nevertheless, based on the MOT [16], as well as significant positive relationship between student-level content coverage and student achievement, we hypothesize that:

Hypothesis 2: School-level content coverage (mean content coverage) will have positive relationship with Malaysian students’ Mathematics achievement.

METHODOLOGY

Sample

In this study, we utilized the Malaysian data from the PISA 2012. Specifically, a total of 4247 students from 135 national secondary schools were chosen as a sample. This sample is part of the overall 5197 Malaysian-sampled students from 164 schools of various types (such as fully residential schools, religious schools and others).
Missing Data
The PISA 2012 data were subjected to data loss caused by questionnaire rotation design [24]. Hence, prior to conducting the Hierarchical Linear Modelling (HLM) analysis, we imputed the missing data using Predictive Mean Matching estimation via R-package Multivariate Imputation by Chained Equations (MICE), in which five imputed datasets were created. We then verified the distribution of the imputed datasets using Mardia’s multivariate normality test via R-package Multivariate Normality (MVN).

Measures
The dependent variable for this analysis is mathematics achievement, which is also defined as Mathematics literacy in the PISA 2012. Five plausible values were created in the PISA 2012 to represent mathematics score of each individual student. In this analysis, we used the plausible value option of HLM 7.0, in which five models was run internally for each plausible value and the estimates were then being averaged according to Rubin’s rule [27]. Specifically, since there were five imputed data sets, therefore each model in the analysis was actually repeated 25 times (5 imputed data sets multiplied by 5 plausible values).

Student-Level Measures
The student-level content coverage variables are experience with applied mathematics tasks at school; experience with pure mathematics tasks at school; familiarity with mathematical concepts and experience with various types of Mathematics problems at school.

Prior to forming the student-level content coverage variables, we conducted the Confirmatory Factor Analysis (CFA) using R-package lavaan survey due of its ability to handle the multiply-imputed datasets and also integrate PISA design weight in the analysis. Since the Mardia’s multivariate normality test showed that all imputed datasets did not follow multivariate normal distribution, we therefore employed Maximum Likelihood Robust (MLR) estimation in the CFA. MLR was chosen because previous studies showed that MLR produced a reliable parameter for which standard errors and mean-adjusted chi-square were robust to non-normality [22]. The student-level content coverage variables were then created based on the final CFA results. The description of each variable is as follows:
- Experience with applied mathematics tasks at school (4 items, Composite Reliability = 0.67) measured students’ exposure to applied mathematics tasks at school (e.g. “Calculating how much more expensive a computer would be after adding tax”). Items were rated on 4-point Likert scale ranging from 1 (never) to 4 (frequently).
- Experience with pure mathematics tasks at school (3 items, Composite Reliability = 0.72) measured students’ exposure to pure mathematics tasks at school (e.g. “Solving an equation like 6x^2 + 5 = 29”). Items were rated on 4-point Likert scale ranging from 1 (never) to 4 (frequently).
- Familiarity with mathematical concepts (7 items, Composite Reliability = 0.80) measured students’ perceived familiarity with mathematics concepts (e.g. “Thinking about mathematical concepts: how familiar are you with the following terms: Vectors?”). Items were rated on 5-point Likert scale ranging from 1 (never heard of it) to 5 (know it well, understand the concept).
- Experience with various types of Mathematics problems at school (5 items, Composite Reliability = 0.79) measured students’ exposure to various types of mathematical tasks in mathematics lesson or in tests (e.g. “Algebraic word problem”). Items were rated on 4-Likert scale ranging from 1 (never) to 4 (frequently).

School-Level Measures
The school-level content-coverage variable was obtained by aggregating the student-level content coverage variables. The rwg (j), ICC(1) and ICC(2) values were calculated to justify the aggregation process. The mean values of rwg (j) for all datasets showed a strong or very strong agreement (range from 0.70 to 0.71), which indicate approximately 71% homogeneity of perception within schools for each variable of content coverage. The ICC (1) values for all variables in the five imputed datasets were in the range of 0.78 to 0.79, indicating acceptable reliability [16]. The same pattern is shown by the ICC(2) for all the variables in the five imputed datasets. The rwg(j) and ICC values signified a strong justification for aggregating the student-level content coverage variables to the school-level variable, which was named as mean content coverage.

Hierarchical Linear Modelling (HLM)
In this study, we used the HLM 7.0 software since the PISA 2012 data was hierarchical or nested in nature [24]. Failure to take into account the data structure can produce less accurate estimates. We also applied Full Maximum Likelihood (FML) estimation, in which regression coefficients and variance components were included in the likelihood function. FML warrants the comparability between two different models in terms of fixed effects, and this obviously makes it more appealing than the Restricted Maximum Likelihood estimation, which is another option of estimation in the HLM [14]. We also used grand mean centering in each model in order to reduce multicollinearity.
To test all hypotheses, we ran seven models using step-up model building approach. The first model was the null model (Model 1), which did not contain any independent variables. This null model is equal to one-way random effects ANOVA and can be used to decompose within- and between-school variance components. At the succeeding stage, each content coverage variable was entered separately into the model in order to test hypotheses 1. Next, all significant student-level variables in the Model 2 until Model 5 were entered into Model 6. Finally, Model 7 was built by including all significant student-level content coverage variables in Model 6 as well as the school-level content coverage variable. The HLM analysis was first conducted for each imputed dataset, and subsequently, results from all datasets were pooled and reported.

RESULTS AND DISCUSSION

Descriptive Statistics and Zero-Order Correlations

Univariate analyses in the form of descriptive statistics were conducted for the five imputed datasets using SPSS 22.0 software. Based on Table 1, Malaysian students had vast experience doing pure mathematics tasks (M = 3.26, SD = 0.91) and received a lot of exposure to various types of mathematics tasks during class or during assessment (M = 3.29, SD = 0.76). On the contrary, on average, students still lack of familiarity to mathematics concepts (M = 2.65, SD = 0.94) and reported that they had less experience of conducting applied mathematics tasks (M = 2.49, SD = 0.85). On average, content coverage in the Malaysian National Secondary schools was at moderate level (M = 2.90, SD = 0.72).

Bivariate analysis for student-level content coverage variables was also conducted for the five imputed datasets using SPSS 22.0 software. Table 2 shows the bivariate analysis for student-level content coverage variables. Overall, most of the correlations were significant. The correlations between the student-level content coverage variables were in the range of 0.13 (between familiarity with mathematical concepts and experience with these types of problems at school) up to 0.38 (between experience with applied mathematics tasks at school and experience with pure mathematics tasks at school). Since all correlations were less than 0.8, it can be concluded that there was no issue of multicollinearity.

Table 1: Weighted descriptive statistics for student- and school-level content coverage variables

<table>
<thead>
<tr>
<th>Variables</th>
<th>Mean (M)</th>
<th>Standard Deviation (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Student-level</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experience with applied mathematics tasks at school</td>
<td>2.49</td>
<td>0.85</td>
</tr>
<tr>
<td>Experience with pure mathematics tasks at school</td>
<td>3.36</td>
<td>0.91</td>
</tr>
<tr>
<td>Familiarity with mathematics concepts</td>
<td>2.65</td>
<td>0.94</td>
</tr>
<tr>
<td>Experience with various types of mathematics problems at school</td>
<td>3.29</td>
<td>0.76</td>
</tr>
<tr>
<td><strong>School-level</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean content coverage</td>
<td>2.90</td>
<td>0.72</td>
</tr>
</tbody>
</table>

Table 2: Weighted correlation for student-level content coverage variables (N = 4247)

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Experience with applied mathematics tasks at school</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>2. Experience with pure mathematics tasks at school</td>
<td>0.38**</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>3. Familiarity with mathematical concepts</td>
<td>0.24**</td>
<td>0.18**</td>
<td>-</td>
</tr>
<tr>
<td>4. Experience with various types of problems at school</td>
<td>0.25***</td>
<td>0.37***</td>
<td>0.13***</td>
</tr>
</tbody>
</table>

Content Coverage and Mathematics Achievement

The aim of the study was to examine the effects of student- and school-level content coverage variables on the Malaysian students’ Mathematics achievement. Our analysis was conducted by fitting seven hierarchical linear modeling with null as the first model (Model 1). Model 2 to Model 5 were subsequently built to test hypothesis 1, in which each content coverage variable (e.g. experience with applied mathematics tasks at school), was included into the model one at a time. All significant student-level content coverage variables in Model 2 to Model 5 were then entered into Model 6 to determine the effect of each variable simultaneously. All significant variables in Model 6 were then fitted into Model 7 which incorporated school-level content coverage variable. Table 3 shows the results for Model 1 to Model 7.

The null model (Model 1) indicated that 21% variance in the Mathematics achievement was between-school, whereas 79% variance in the Mathematics achievement was within-school.

Based on Model 2 to Model 7, all content coverage variables were statistically significant predictors of Malaysian students’ mathematics achievement, in which experience with pure mathematics tasks at school had the largest positive predictive effect. On the contrary, experience with applied mathematics tasks at school had a moderate negative predictive effect on Mathematics achievement. Both familiarity with mathematical concepts and experience with various types of problems at school were also significant with the positive effects.
Basically, there were differences in the mean of students’ mathematics achievement between schools, as demonstrated by the significant variance intercept. Nevertheless, no significant variance components for all the student-level content coverage variables indicates that schools did not vary in the relationships between student-level content coverage variables and mathematics achievement.

Specifically, the inclusion of all student-level content coverage variables had successfully explained 21.35% of the between-school variance and 19.25% of the within-school variance. Hence, the hypothesis 1 was partially supported, in which only experience with pure mathematics tasks at school, familiarity with mathematical concepts and experience with various types of mathematics problems at school, had positive relationships with Malaysian students’ mathematics achievements, whilst experience with applied mathematics tasks at school showed the reverse trend.

All statistically significant variables in Model 6 were then entered to Model 7 together with the school-level content coverage variables. Result showed that all student and school-level OTL variables were significant. As expected, except experience with applied mathematics tasks at school, other student-level content coverage variables-experience with pure mathematics tasks at school, familiarity with mathematical concepts and experience with various types of mathematics problems at school had positive relationships with Malaysian students’ mathematics achievements, with experience with pure mathematics tasks at school having the largest predictive effect. At school-level, the mean content coverage had positive predictive effect on mathematics achievement. The inclusion of all student- and school-level content coverage variables had successfully explained an additional 9.4% of the between-school variance, and 0.62% of the within-school variance.

**CONCLUSION**

This study aimed at examining the predictive effects of student- and school-level content coverage variables on the Malaysian students’ mathematics achievement. The results of the study revealed that student-level factors contributed to the most variance explained. In general, there were differences in the mean of student mathematics achievement between the Malaysian national secondary schools.

**Table 3: Fixed effects and variance-covariance estimates for models of the content coverage predictors of Mathematics achievement**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
<th>Model 6</th>
<th>Model 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>406.34***</td>
<td>406.44***</td>
<td>406.45***</td>
<td>406.23***</td>
<td>406.46***</td>
<td>406.39***</td>
<td>406.46***</td>
</tr>
<tr>
<td></td>
<td>(3.40)</td>
<td>(3.34)</td>
<td>(3.04)</td>
<td>(3.31)</td>
<td>(3.35)</td>
<td>(2.98)</td>
<td>(2.80)</td>
</tr>
<tr>
<td>Student-level</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(12.13)</td>
<td>(1.96)</td>
<td>(1.93)</td>
<td>(1.99)</td>
<td>(2.49)</td>
<td>(2.28)</td>
<td>(2.23)</td>
</tr>
<tr>
<td></td>
<td>(1.47)</td>
<td>(1.64)</td>
<td>(1.63)</td>
<td>(1.64)</td>
<td>(1.64)</td>
<td>(1.64)</td>
<td>(1.64)</td>
</tr>
<tr>
<td>Familiarity with mathematical concepts</td>
<td>18.66***</td>
<td>12.35***</td>
<td>12.19**</td>
<td>18.66***</td>
<td>12.35***</td>
<td>12.19**</td>
<td>12.19**</td>
</tr>
<tr>
<td></td>
<td>(1.98)</td>
<td>(1.99)</td>
<td>(1.95)</td>
<td>(1.99)</td>
<td>(1.99)</td>
<td>(1.99)</td>
<td>(1.99)</td>
</tr>
<tr>
<td>Experience with various types of problems at school</td>
<td>24.15***</td>
<td>12.93***</td>
<td>12.53***</td>
<td>24.15***</td>
<td>12.93***</td>
<td>12.53***</td>
<td>12.53***</td>
</tr>
<tr>
<td>School-level</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean content coverage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept variance (τ₀₀)</td>
<td>1144.57***</td>
<td>1107.52***</td>
<td>910.35***</td>
<td>1103.36***</td>
<td>1050.58***</td>
<td>900.21***</td>
<td>792.60***</td>
</tr>
<tr>
<td>Experience with applied mathematics tasks at school</td>
<td>10.91</td>
<td>12.19</td>
<td>21.35</td>
<td>10.91</td>
<td>12.19</td>
<td>21.35</td>
<td>10.91</td>
</tr>
<tr>
<td>Experience with pure mathematics tasks at school</td>
<td>31.93**</td>
<td>16.19</td>
<td>49.33</td>
<td>31.93**</td>
<td>16.19</td>
<td>49.33</td>
<td>31.93**</td>
</tr>
<tr>
<td>Experience with various types of problems at school</td>
<td>61.69</td>
<td>39.93</td>
<td>23.90</td>
<td>61.69</td>
<td>39.93</td>
<td>23.90</td>
<td>61.69</td>
</tr>
<tr>
<td>Level 1 variance (σ²)</td>
<td>4416.12</td>
<td>4335.46</td>
<td>3730.28</td>
<td>4236.50</td>
<td>4084.16</td>
<td>3566.02</td>
<td>3538.72</td>
</tr>
<tr>
<td>Variance in achievement between schools explained (%)</td>
<td>3.24</td>
<td>19.94</td>
<td>8.21</td>
<td>21.35</td>
<td>30.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variance in achievement within schools explained (%)</td>
<td>1.83</td>
<td>15.53</td>
<td>4.07</td>
<td>5.75</td>
<td>19.25</td>
<td>19.87</td>
<td></td>
</tr>
</tbody>
</table>

Number of students = 4247; Number of schools = 135; Standard errors of the estimates are in parentheses; ***p < 0.001, **p < 0.01, *p < 0.05

Except experience with applied mathematics tasks at school, all other content coverage variables—experience with pure mathematics tasks at school, familiarity with mathematical concepts, and experience with
various types of problems at school were positively related to student achievement. Congruent with previous research [21, 29, 32], students with greater experience with pure mathematics had higher mathematics achievement. Similarly, although smaller in regression weights as compared to other variables, familiarity with mathematical concepts and experience with various types of problems at school were positively associated with mathematics achievement. Thus, consistent with previous research [1, 10, 11, 23], students who were more familiar with the mathematical concepts and had more exposure with various types of mathematical problems at school tended to have higher mathematics achievement.

In contrast, experience with applied mathematics at school showed negative relationship with student achievement. This result was rather surprising since previous research [3, 11, 18, 26] showed significant relationship between students’ exposures towards applied mathematical tasks and their Mathematics achievement. Hence, further research is deemed necessary to understand the rationale behind the somewhat contradict finding of this study.

The result also revealed that the Malaysian national secondary schools differed in terms of the mean of students’ mathematics achievements. However, the relationships between student-level content coverage variables and mathematics achievement were the same across schools.

At the school-level, mean content coverage had substantial positive effect on mathematics achievement. This finding was consistent with previous research [3, 11, 18, 26], indicating that students who studied in a school with higher content coverage tended to have higher mathematics achievement. In general, this result indicated that at school level, policy makers should give more attention to content coverage because it has been revealed as a significant predictor of students’ mathematics achievement.

In summary, both student- and school-level content coverage variables influenced Malaysian students’ mathematics achievement. Obviously, experience with pure mathematics tasks at school together with familiarity with mathematical concepts and experience with various types of problems at school were important factors for students’ mathematics achievement. In addition, school-level content coverage was also regarded as crucial factor in the mathematics teaching and learning processes. On the contrary, the negative predictive effect of experience in applied mathematics at school on students’ mathematics achievement was rather astonishing and shall be re-examined further.

In this study, we used aggregated school-level content coverage variable since it was not directly measured in the PISA 2012. Therefore, there is a possibility of committing an atomistic fallacy, which refers to the fallacy of drawing inferences at higher level using data collected at lower level units [14, 16]. Although the aggregation of student-level content coverage variables to school-level content coverage variable is considered appropriate based on the rwg(j), ICC(1) and ICC(2) values, however, the results would be more accurate if the school-level content coverage were directly measured.

REFERENCES