Growth in Mathematics Cognitive and Content Domains: A 6-Year Longitudinal Study

Magdalena Mo Ching Mok,
Jinxin Zhu and Dennis M. McInerney
The Hong Kong Institute of Education

Anthony Or
Education Bureau, Hong Kong

Keywords: mathematics, content domains, cognitive domains, growth, cross-lagged model
Running Head: Growth in Mathematics Cognitive and Content Domains

Abstract

This study aimed to model mathematics growth of students from Primary 3 to Secondary 3 from two approaches, namely, growth in content domains, and in cognitive domains. The sample consisted of 866 Hong Kong students who were followed longitudinally from Primary 3 to Secondary 3. The data were originally collected by the Education Bureau of Hong Kong. Plausible values of achievements at Primary 3, Primary 6, and Secondary 3 in three cognitive domains (Knowing, Applying, and Reasoning) were obtained using between-item multidimensional Rasch partial credit modelling. A modified autoregressive cross-lagged design was used to predict later achievement in the cognitive domains by prior achievements in these domains. Analysis was repeated for growth in content domains (Number, Shape and Space, Measures, Data Handling, and Algebra). Analysis showed that students' later mathematics achievement was strongly predicted by their previous achievement in the Knowing and the Number domains.

Introduction

This study aimed to explore students' longitudinal growth in mathematics from primary to secondary year levels. Numerous research studies have shown that students' previous mathematics achievement predicts their later achievement (Aunio, 2010; DiPerna, 2005; Geary, 2011; Jordan, 2009; Lefevre, 2010; Locuniak, 2008; Pagani, 2010; Passolunghi & Lanfranchi, 2012; Romano, 2010). For instance, children's proficiencies in numbers at kindergarten was reported to positively predict their later mathematical achievement in the first stage of primary schooling (Aunio, 2010; Geary, 2011; Jordan, 2009; Lefevre, 2010; Locuniak, 2008; Pagani, 2010; Passolunghi & Lanfranchi, 2012; Romano, 2010). Early number proficiency was found (Locuniak, 2008; Pagani, 2010; Romano, 2010) to be the strongest predictor among the other general predictors.

Previous studies found that mathematics achievement at a later stage was strongly predicted by primary students' prior achievement in the subject (DiPerna, 2005). Mathematics skills at school entry had the greatest predictive power, among the three key elements of school readiness, namely, school-entry academic skills, attention, and socio-emotional skills (Duncan, 2007). In England, a two-year longitudinal study of children's early development showed that five-year-old children's total scores of mathematics (comprising eight content domains: concepts of comparison, classification, one-to-one correspondence, seriation, using number-words, structured counting, 'resultive' counting and applying general knowledge of numbers in real-life situations) at the end of their reception year were predictive of later achievement at seven years (Aubrey, 2003). An extension of the research reconfirmed the results with similar predictive value of mathematical scores of the same group of children at eleven years old (Aubrey, Godfrey, & Dahl, 2006).

Another two-year longitudinal study in Belgium showed that procedural counting and conceptual counting knowledge in Grade one were predictors of numerical facility and arithmetical achievement, respectively, in Grade three, and procedural counting knowledge in Grade three predicted numerical facility in Grade five, but the prediction of that between Grades two and four was not sustained (Desoete, 2009).

In America, a five-year longitudinal study, tracking the students from kindergarten to Grade five once a year, showed that students' quantitative competencies uniquely predicted their mathematics learning and achievement growth, and the contributions were above and beyond that of domain general abilities (intelligence, central executive, phonological loop, visual-spatial sketch pad and processing speed) (Geary, 2011).

Similar to primary school mathematics, achievement in mathematics at secondary levels was also predicted by students' prior achievement in the subject. Indeed, previous achievement was found to be the strongest predictor among other predictors including, attitudes toward mathematics, outcome expectancy, value of studying mathematics, engagement in mathematics, classroom context of mathematics lessons, parents' education level, expectation of parents, and parental school involvement (Hemmings, Grootenboer, & Kay 2011; Kyttälä & Björn, 2010; Reynolds, 1991; Yates, 2000). A review by Marsh and Martin (2011) showed, inter alia., that prior achievement was the more important predictor of later achievement after controlling for self-concept.

In comparison to studying mathematics development from the perspective of growth in content domains, research concerning growth in mathematics cognitive domains is relatively rare. Such questions as, "What is the role of competencies in reasoning at primary grades in developing competencies in reasoning later in secondary grades?" are hardly ever addressed in research. Development of mathematics cognitive domains from primary to secondary grade levels remains a mystery. Further, with only a few exceptions (e.g. Wilkins & Ma, 2010), these early studies used total scores as repeated measures in tracking mathematics growth. This situation is far from satisfactory. Investigation into mathematics growth from a contents perspective is curriculum-based and total scores carry little information for teachers on how to enhance their instruction. Given that mathematics curricula are different for different education systems, transferability of findings on mathematics growth in certain content areas (e.g. Algebra) across systems may be limited by the proximity of curriculum between the target system and the system where the research was originally undertaken. Research into mathematics growth in cognitive domains offers an alternative approach.

The present study focused attention on student's growth in mathematics using students' achievement in three cognitive domains, namely, knowing, applying and reasoning when they were at Primary 3, as predictors of their later achievement in these cognitive domains when they were at Primary 6, and at Secondary 3. In parallel, this study also investigated mathematics growth over six years of the same group of students using their achievements at Primary 3 in four content domains, namely, Number, Shape and Space, Measures, and Data Handling as predictors for their mathematics achievement at Primary 6 in these content domains and in Algebra. Achievements at Primary 3 were also used to predict achievements in Number, Shape and Space, Data Handling and Algebra at Secondary 3. In other words, this study explored growth via two paths, one using students' prior achievement in cognitive domains to predict their later achievement in cognitive domains, and another using students' prior achievement in content domains to predict their later achievement in content domains. It would be ideal theoretically if items could be classified in terms of a cognitive by content domain matrix and prediction made based on this 3 × 5 matrix classification but in reality, there were too few items in most of the cells to make valid and stable estimations from a statistical perspective. This study confined itself to investigating growth separately within each of the two classifications, namely, by cognitive domains and by content domains.

The conceptual framework for this study is underpinned by a modified version of the auto regression cross-lagged model (Bollen & Curran, 2004), which hypothesized the longitudinal relationships from Primary 3 to Secondary 3 between students' achievement in the Knowing, Applying, and Reasoning cognitive domains (Figure 1) and between their achievements in Number, Shape and Space, Measures, Algebra, and Data Handling (Figure 2). The core of the auto regression cross-lagged model is that a student's mathematics achievement at a later time (e.g. time 2) is explained by the student's mathematics achievement at an earlier time (time 1). This study modified the auto regression cross-lagged model by allowing achievement at an even later stage (e.g. time 3) to be explained not only by achievement at an immediate past testing occasion (time 2), but also by a much earlier previous achievement (time 1). Further, repeated measures of more than one dimension (e.g. three cognitive dimensions) are analysed simultaneously instead of looking into only one aspect of growth.

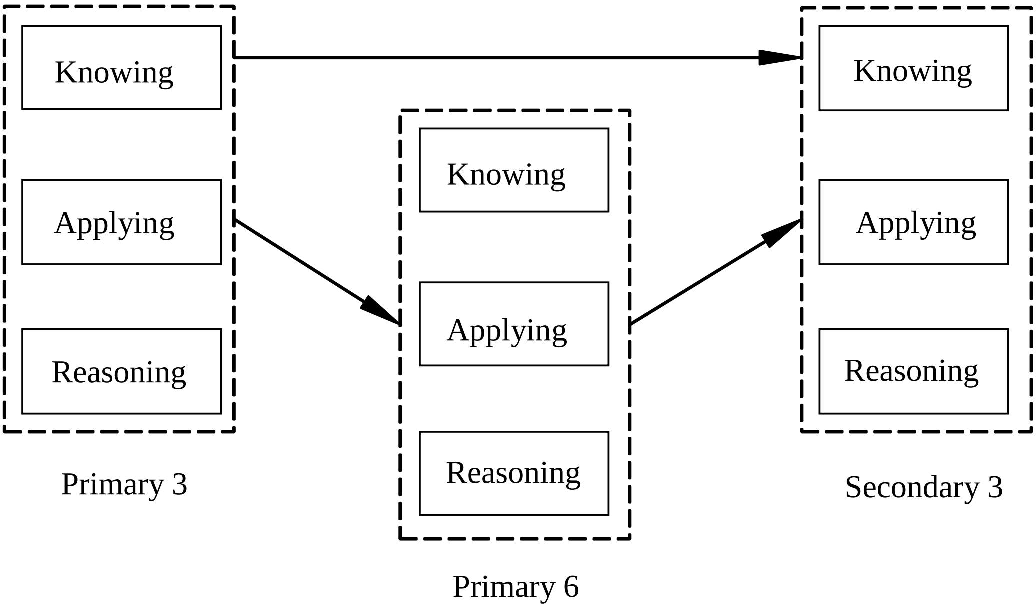


Figure 1. Conceptual model on mathematics cognitive growth at Primary 3, Primary 6 and Secondary 3

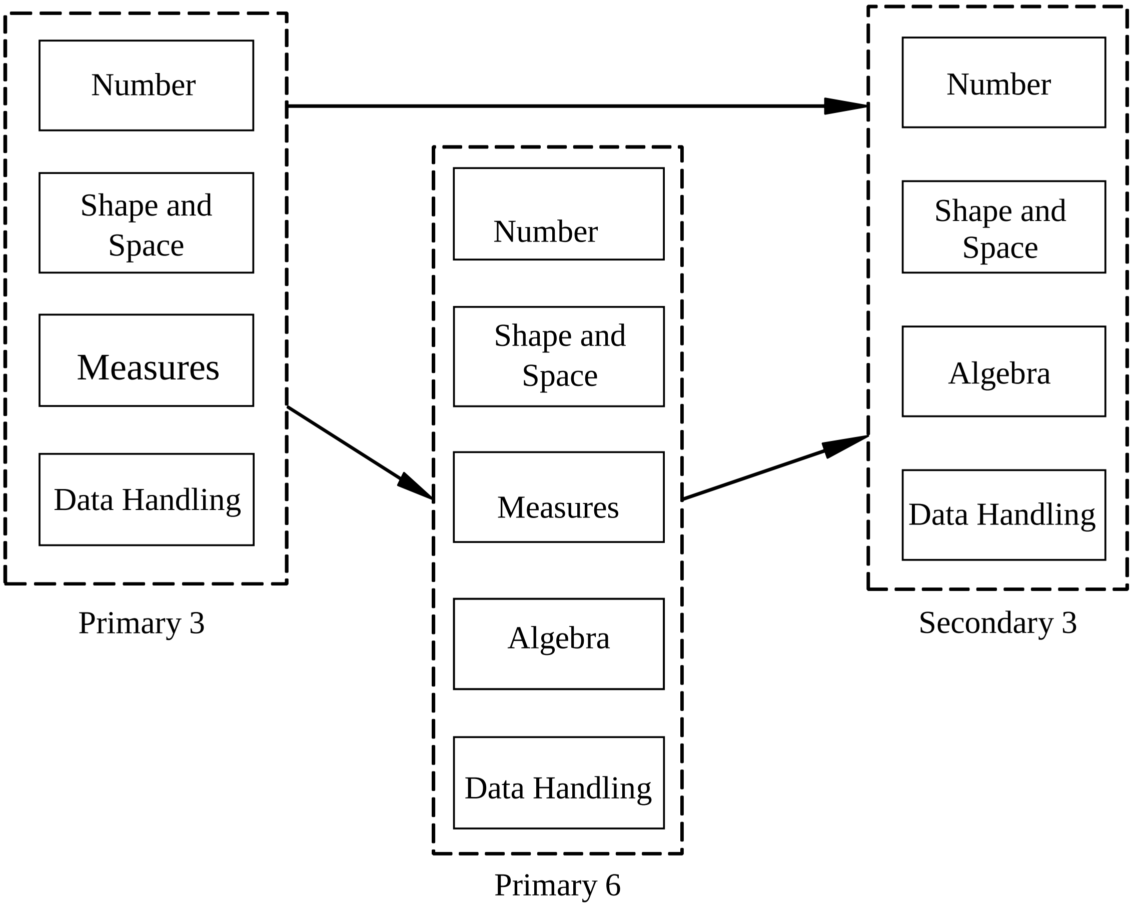


Figure 2. Conceptual model on mathematics content-domain growth at Primary 3, Primary 6 and Secondary 3

Background to Mathematics Learning in Hong Kong

Hong Kong parents and the society place strong values on the studying of mathematics. Mathematics is in the curriculum of kindergarten, primary, and secondary schools. Performance of mathematics at the Hong Kong Diploma of Education determines a student's opportunity for university education. The government has specific guidelines on the mathematics curriculum as well as instruction hours at both primary and secondary levels (Curriculum Development Council, 2002).

The Hong Kong mathematics curriculum is designed to developed mathematics self-confidence in students. Focusing on the application of mathematics in work and daily life, the long term goal of the mathematics curriculum is to prepare students for lifelong learning, and to develop in students positive attitudes toward mathematics learning (CDC, 2002, 9-10). Learning and teaching in the subject of mathematics in Hong Kong is structured around a curriculum framework that organises content knowledge and skills into strands, or content domains (Curriculum Development Council 2002, p. 10). The mathematics curriculum for Primary1 to 6 is organised around five content domains, namely, Number, Shape and Space, Measures, Data Handling and Algebra. These topics are taught throughout primary school years, and the associated concepts are revisited, and progressively broadened and deepened at later school years. Number, Shape and Space, Measures and Data Handling are taught at every year level between Primary1 and Primary 6. The exception is Algebra, which is taught only from Primary 4 onwards. These five domains were combined into three domains for mathematics curriculum at Secondary year levels, namely, (a) Number and Algebra, (b) Measures, Shape and Space, and (c) Data Handling (Curriculum Development Council Committee, 2002).

It should be noted that the current study focused on junior Secondary school years (i.e., Secondary 1 to Secondary 3), and attention was paid to the content domains of Number, Shape and Space, Data Handling, and Algebra. At junior secondary school years, the content domain of measures was only given a small weight and only one item was identified in the Secondary 3 TSA assessment for the cohort in this study. Consequently, this item was not included in the analysis for the current study. Further, the content domain of Number was given heavy emphasis in the curriculum. Thirty-six out of 68 items belong to the

Number domain in the assessment items. Consequently, the items for Number and for Algebra were separated into two distinct content domains in the analysis for this study rather than combined as specified in the Mathematics curriculum. In addition, some items in the TSA assessment used in this study could be classified into more than one content domain. This issue was handled in the analysis using a Within-item Multidimensional Partial Credit Rasch model which permitted cross-loading of items onto more than one domain (Adams, Wilson, & Wang, 1997).

Methods

Sample

Secondary data for this study were originally collected by the Hong Kong government at the Territory-wide System Assessment (TSA) from a random and representative sample of 866 Primary 3 students three times: in 2004, 2007 (students now in Primary 6), and 2010 (students now in Secondary 3). The longitudinal data collected from each student were their responses to individual assessment items at the TSA in these three years. More details about the sample were provided in the next section.

Territory-wide System Assessment (TSA)

The Territory-wide System Assessment (TSA), along with Student Assessment (SA), is one of the two components of the Basic Competence Assessment designed by the Hong Kong Government as a reform initiative focusing on the promotion of assessment for learning literacy (Education Commission, 2000). Until 2013, the TSA is an assessment administered annually by the government for all schools in Hong Kong based on the Chinese Language, English Language and Mathematics curricula at Primary 3 (i.e., end of Key Stage 1), Primary 6 (i.e. end of Key Stage 2), and Secondary 3 (i.e., end of Key Stage 3). The goal is to collect system-wide assessment data on student performance in these three subjects in order to facilitate policy review and policy formulation at the system- and school-levels. Further, system- and school-level reports are provided to schools with detailed information at item- and sub-paper levels on the strengths and weaknesses of students against basic competencies in the three subjects with an aim to inform teaching and learning.

The TSA is criterion-referenced and response patterns include multiple choice, matching, drag-and-drop, as well as constructed response. Items were scored either dichotomously (0/1 for wrong/right answers) or with partial credits (0/1/2 for totally wrong/partially correct/fully correct answers) according to the scoring schemes (Hong Kong Examinations and Assessment Authority, 2014; Mok, 2010).

Given the wide-range of curriculum topics in each subject and at each Key Stage to be covered by the TSA, a matrix sampling method is used. The method consists of random distribution to candidates of four assessment booklets, which are linked by a set of common anchor items for each curriculum subject. In other words, there were 4 × 4 × 4 = 64 combination of booklets across the three TSA test occasions between Primary 3 and Secondary 3 for each of Mathematics, Chinese Language and English Language assessments. In this study, 866 students were identified to have sat the TSA conducted in 2003, 2010, and 2013 with the same test booklets (Booklets 4, 4 and 1 respectively) in these three test administrations. Item-level responses at the three test occasions of these 866 students were used in the current study. The sample thus represented a truly random and representative sample of 1/64 (1.5625%) of the population.

Ethics

Data collection was conducted strictly abiding to ethical principles. Confidentiality of schools and students were maintained throughout by using codes known only to senior personnel of the Hong Kong Examinations and Assessment System. Schools and candidates were invited to give informed consent before the TSA that the assessment data were collected for policy research purposes. No background information (e.g. socio-economic status), other than gender and school membership, of the candidates were collected.

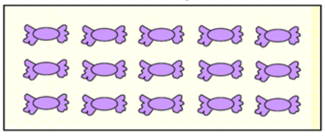
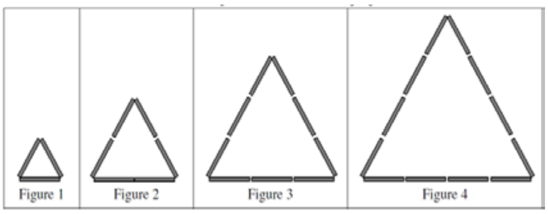
Measures

Assessment items in the study were content analysed and categorised according to two classification systems: (a) according to cognitive domains as defined by the TIMSS 2001 classification framework (Mullis, Martin, Ruddock, O’Sullivan, & Preuschoff, 2009); and (b) according to content domains defined by the curriculum guides of the Hong Kong government (Curriculum

Development Council, 2002). Two sets of scales were formed from items that made up the respective domains.

In the TIMSS 2001 framework (Mullis, Martin, Ruddock, O'Sullivan, & Preuschoff, 2009), mathematics items were classified according to intellectual demands of students into three mathematical cognitive domains, namely, Knowing, Applying, and Reasoning. Knowing in mathematics refers to the skills of recalling, recognizing, computing, retrieving, measuring, and classifying/ordering; Applying refers to the skills of selecting, representing, modeling, implementing and solving routine mathematical problems; Reasoning in mathematics refers to the skills of analyzing, generalizing/specializing, integrating/synthesizing, justifying, and solving non-routine problems. Scales were formed by gathering items in the same content domain using between-item Multidimensional Partial Credit Rasch model (Adams, Wilson, & Wang, 1997). Table 1 shows example items in each of the three cognitive domains.

Table 1. Example items in Knowing, Applying, and Reasoning cognitive domains

Cognitive Domain	Example Item
Knowing	$53 \times 409 =$
Applying	There were 15 chocolates in a box. Mimi took $\frac{1}{3}$. Fanny took $\frac{1}{5}$. Who took more? How many chocolates did she take?  Answer: took more. She took chocolates.
Reasoning	Michael used some sticks of the same length to form the following figures:  According to the above pattern, how many sticks should Michael use in the 5th figure?

Source of items: Retrieved on 30 October 2013 from
http://www.bca.hkeaa.edu.hk/web/Common/res/2007priPaper/P6Math/2007_TSA_6ME4.pdf

The number of items across the Knowing, Applying, and Reasoning cognitive domains are not equal, reflecting the distribution of these cognitive skills in the mathematics curriculum. Assessment at Primary 3 had 28 Knowing items, 18 Applying items, and 15 Reasoning items. Assessment at Primary 6 had had 36 Knowing items, 19 Applying items, and 19 Reasoning items. Assessment at Secondary 3 had 30 Knowing items, 20 Applying items, and 18 Reasoning items. Analysis found that scales formed by items in these domains had good psychometric properties (Table 2). The scales were internally consistent (Cronbach's Alphas are between 0.780 and 0.929) and multidimensional Rating Scale Rasch analysis showed that the scales had good model-data fit. Item Infit (i.e., weighted) Mean Square (MNSQ) statistics (Wu, Adams, Wilson, & Haldane, 2007) were between 0.71 and 1.39, and only 18out of 203 items had Outfit statistics greater than 1.5.

Table 2. Psychometric Property of the Scales (Categorized by Cognitive Domain)

Scale	No. of Items	Cronbach's Alpha	Rasch Item Infit (Range)	Rasch Item Outfit (Range)	No. of items with Outfit >1.5
P3 Booklet 4 (61 Items)					
P3 Knowing	28	0.830	0.89-1.18	0.78-1.37	0
P3 Applying	18	0.812	0.89-1.18	0.73-1.46	0
P3 Reasoning	15	0.780	0.82-1.33	0.54-1.73	2
P6 Booklet 4 (74 Items)					
P6 Knowing	36	0.885	0.84-1.22	0.70-1.43	0
P6 Applying	19	0.895	0.86-1.34	0.79-1.74	2
P6 Reasoning	19	0.886	0.73-1.36	0.61-2.68	3
S3 Booklet 1 (68 Items)					
S3 Knowing	30	0.914	0.71-1.39	0.52-1.79	5
S3 Applying	20	0.929	0.70-1.37	0.58-1.70	3
S3 Reasoning	18	0.921	0.77-1.38	0.68-1.92	3

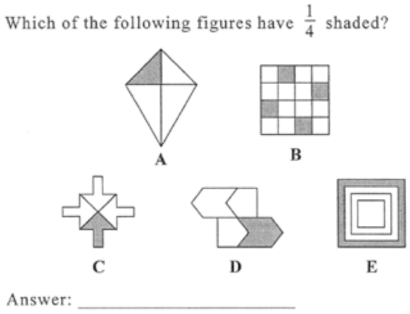
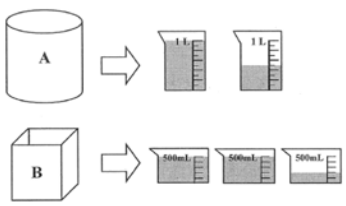
In parallel, assessment items were content analysed and classified into mathematical content domains of Number, Shape and Space, Measures, and Data Handling for Primary 3; Number, Shape and Space, Measures, Data

Handling, and Algebra for Primary 6; and into Number, Shape and Space, Data Handling, and Algebra for Secondary 3. Scales were formed by gathering items in the same content domain. Notably, four items involved concepts and skills of more than one content domain and this was handled in the analysis using within-item Multidimensional Partial Credit Rasch model (Adams, Wilson, & Wang, 1997).

The number of items across the content domains of Number, Measures, Shape and Space, Data Handling, and Algebra is not equal, reflecting the distribution of these content skills in the Hong Kong mathematics curriculum. Over half of the items of assessment at each year level were Number items. There were 17 Measures items at Primary 3, 22 at Primary 6, and 1 item at Secondary 3, which, however, was excluded from the study because of reliability issue. There were 15 Shape and Space items at Primary 3, 11 at Primary 6, and 23 at Secondary 3. Data Handling was not an important domain at all year levels, and was represented by 8, 6, 7 items at Primary 3, Primary 6 and Secondary 3 respectively. Algebra was not in Key Stage I curriculum and so there was no Algebra item at Primary 3 assessment, but there were 3 and 20 Algebra items at Primary 6 and Secondary 3 respectively (Table 2). Example items in the mathematics content domains are presented in Table 3.

Table 3. Example items in content domains

Content Domain	Example Item
Number	(Item 7, P6) Which of the following expressions is most suitable for estimating the value of $2.1+5\frac{1}{9}\div\frac{1}{3}$? A. $2+5\div 3$ B. $2+5\times 3$ C. $2+6\div 3$ D. $2+6\times 3$

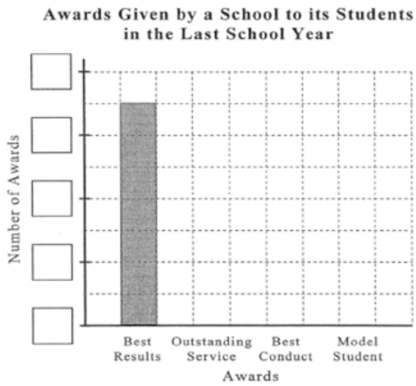
Shape and Space	(Item 20, P6) 
Measures	(Item 24, P6)  Containers A and B are completely filled with water. All the water in the containers is then poured into two different types of beakers (see the diagram above). The capacity of container ___ is larger as it can hold ___ mL more water than the other container.
Algebra	(Item 38, P6) Solve the equation: $\frac{y}{3} - 2\frac{1}{2} = 3$ $y =$______

<table>
<tr>
<td>Data Handling</td>
<td>
(Item 42, P6)

The table below shows the number of awards given by a school to its students in the last school year.

<table>
<tr><td>Awards</td><td>Best Results</td><td>Outstanding Service</td><td>Best Conduct</td><td>Model Student</td></tr>
<tr><td>Number</td><td>70</td><td>50</td><td>40</td><td>60</td></tr>
</table>

Using the above data, complete the following bar chart and fill in the boxes with the correct numbers.



</td>
</tr>
</table>

Source of items: Retrieved on 30 October 2013 from
http://www.bca.hkeaa.edu.hk/web/Common/res/2007priPaper/P6Math/2007_TSA_6ME4.pdf

The psychometric properties of scales formed from assembling items according to content domains were not as strong as psychometric properties of scales formed from assembling items according to their cognitive domains. This was most likely due to some content domain scales were rather short. Analysis found that most of the content domain scales had good psychometric properties (Table 4) but the Data Handling scales at Primary 3 (comprising 8 items) and at Primary 6 (comprising 6 items) had Cronbach's Alpha 0.595and 0.531 only. The other scales were internally consistency (Cronbach's Alphas between 0.615 and 0.905). Within-item Multidimensional Partial Credit Rasch analysis showed that the scales had good model-data fit (Item Infit statistics were between 0.73 and 1.33), but 9 items in the Measures scales had Outfit statistics greater than 1.5, especially at Primary 6. The last result indicated that data from students at the extreme ends of the proficiency scales might not fit the Rasch model well.

Table 4. Psychometric Property of the Scales (Categorized by Content Domain)

Content Domain	No. of Items	Cronbach's Alpha	Rasch Infit (Range)	Rasch Outfit (Range)	No. of Items with Outfit outside 0.5-1.5
P3 Booklet 4 (61 items)					
P3 Number	32	0.813	0.81-1.28	0.64-1.60	2
P3 Measures	17	0.793	0.87-1.28	0.82-1.60	2
P3 Shape and Space	15	0.777	0.94-1.25	0.73-1.47	0
P3 Data Handling	8	0.595	0.88-1.25	0.47-1.49	1
P6 Booklet 4 (74 items)					
P6 Number	38	0.913	0.73-1.33	0.65-2.35	4
P6 Measures	22	0.784	0.81-1.33	0.67-2.54	7
P6 Shape and Space	11	0.730	0.91-1.08	0.82-1.13	0
P6 Data Handling	6	0.531	0.96-1.17	0.82-1.43	0
P6 Algebra	3	0.615	0.95-1.16	0.89-1.20	0
S3 Booklet 1 (68 items)					
S3 Number	36	0.905	0.79-1.25	0.71-1.46	0
S3Shape and Space	23	0.902	0.71-1.41	0.62-2.72	4
S3 Data Handling	7	0.769	0.77-1.15	0.52-2.20	1
S3 Algebra	20	0.868	0.69-1.70	0.51-4.80	7

* Some items involved concepts and skills of more than one domain

Analysis

This study followed a two-step analysis approach. First, Multidimensional Partial Credit Rasch model (Adams, Wilson, & Wang, 1997; Masters, 1982) was used to calibrate the items and students on the same measurement scale using the ConQuest software (Version 2) (Wu, et al., 2007). Multidimensional, rather than unidimensional, Rasch model was used in order to enhance measurement precision and to capitalize on the relationship between students' scores in different cognitive domains (Adams, Wilson, & Wang, 1997; Mok & Xu, 2013). This step of the analysis generated five sets of plausible values of the latent variables (Wu, 2005) for each variable for each student at each year level. The second step of analysis used the sets of plausible values to fit path model using the Mplus software (version 6) (Muthén & Muthén, 1998-2010). The analysis was conducted for each set of plausible values and then the results across all sets were averaged and reported. Further, conditions had been given to the use of multilevel modelling methods in view of the nested data structure (students nested within schools). Multilevel modelling was not used in this study because the matrix sampling method resulted in only a few (often one and the majority less than five) students from each school being included in the current sample.

Results and Discussion

Analysis showed that the hypothesized conceptual models tested in this study were supported by the data. For all the models tested, the goodness of fit indices Comparative Fit Index (CFI) and Tucker Lewis Index (TLI) were greater than 0.96, and the Root Mean Square Error of Approximation (RMSEA) and Standardized Root Mean Square Residual (SRMR) were less than 0.05 and 0.09 respectively. The Chi-squared values for the fitted models were both substantially lower than the baseline model, and the Chi-squared values were not statistically significant (probability greater than 0.05) (Byrne, 2012). Details of results on growth in mathematics from Primary 3 to Secondary 3 are detailed below.

Growth in Mathematics Cognitive Domains

In the path model of cognitive domains, scores in the domains at earlier year levels were used as predictors for scores in the domains at later year levels. Results showed that Knowing in Primary 3 predicted all cognitive domains in both Primary 6 and Secondary 3. The finding provides strong evidence for the importance of early mathematics instruction in knowing. Knowing in Primary 6 also predicted all cognitive domains in Secondary 3. Each of Applying and Reasoning in Primary 3 predicted all cognitive domains in Primary 6 but not in Secondary 3. Applying at Primary 6 predicted Applying at Secondary 3 but none of the other cognitive domains at Secondary 3. Reasoning at Primary 6 did not predict any other cognitive domains in Secondary 3. The predictive power of Primary 6 Reasoning was only through its strong correlations with Knowing and Applying cognitive domains (Pearson Product Moment Correlation Coefficients were 0.844 and 0.844 respectively). These results are presented in Table 5 and Figure 3. The model has excellent fit: CFI = 1.000, TLI = 1.000, RMSEA = 0.000, SRMR = 0.001, Chi Square = 0.074 (d.f. = 1, P = 0.785). (Need to check). It is interesting to speculate on whether this result may be related to different conceptions of 'reasoning' as students grow cognitively. Reasoning may be conceptualized (and defined) differently for items at different year levels. This is a matter for further investigation.

Table 5. Path Coefficient and Standard Errors (Cognitive Domains)

STDYX	Path Coeff.	S.E.	Est./S.E.	Prob.
P6 Knowing ON				
P3 Knowing	0.258	0.075	3.456	p <0.05
P3 Applying	0.244	0.077	3.182	p <0.05
P3 Reasoning	0.168	0.042	4.034	p <0.05
P6 Applying ON				
P3 Knowing	0.254	0.075	3.405	p <0.05
P3 Applying	0.237	0.077	3.089	p <0.05
P3 Reasoning	0.183	0.041	4.425	p <0.05
P6 Reasoning ON				
P3 Knowing	0.263	0.076	3.476	p <0.05
P3 Applying	0.214	0.078	2.749	p <0.05
P3 Reasoning	0.179	0.042	4.292	p <0.05
S3 Knowing ON				
P6 Knowing	0.288	0.084	3.445	p <0.05
P6 Applying	0.182	0.097	1.888	NS
P6 Reasoning	0.099	0.069	1.431	NS
P3 Knowing	0.104	0.049	2.148	p <0.05
P3 Applying	0.005	0.033	0.146	NS
P3 Reasoning	0.065	0.039	1.677	NS
S3 Applying ON				
P6 Knowing	0.306	0.082	3.718	p <0.05
P6 Applying	0.188	0.095	1.979	p <0.05
P6 Reasoning	0.091	0.068	1.328	NS
P3 Knowing	0.104	0.040	2.587	p <0.05
P3 Reasoning	0.067	0.038	1.745	NS

S3 Reasoning　ON				
P6 Knowing	0.308	0.083	3.695	p <0.05
P6 Applying	0.179	0.096	1.861	NS
P6 Reasoning	0.091	0.069	1.312	NS
P3 Knowing	0.088	0.044	1.986	p <0.05
P3 Applying	0.006	0.021	0.305	NS
P3 Reasoning	0.072	0.039	1.861	NS
P6 Knowing　WITH				
P6 Applying	0.920	0.005	176.885	p <0.05
P6 Reasoning	0.844	0.010	86.122	p <0.05
P6 Applying　WITH				
P6 Reasoning	0.884	0.007	119.514	p <0.05
S3 Applying　WITH				
S3 Knowing	0.897	0.007	131.941	p <0.05
S3 Reasoning　WITH				
S3 Knowing	0.848	0.009	90.234	p <0.05
S3 Applying	0.959	0.003	319.533	p <0.05

Note: NS stands for "Not Significant at 5%"

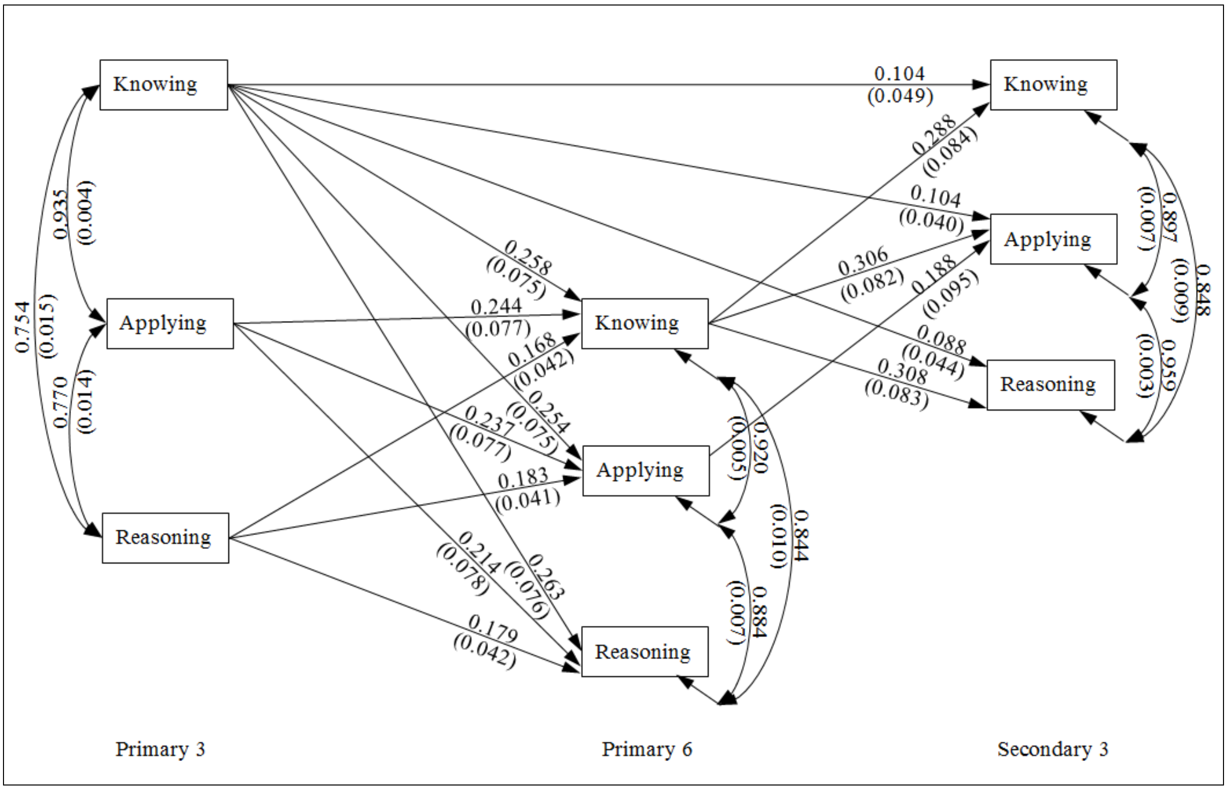


Figure 3. Standardized result (STDYX) of Path model in Cognitive Domain

Note: Only paths coefficient significant at α = 0.05 are shown. CFI: 1.000, TLI: 1.000, RMSEA: 0.000, SRMR: 0.001, Chi Square: 0.074 (d.f. = 1, P = 0.785)

As presented in Table 6, the R-squared values of scores at Primary 6 in the Knowing, Applying, and Reasoning cognitive domains were 0.457, 0.474, and 0.460 respectively. This means between 45.7% and 47.4% of variances of Primary 6scores in Knowing, Applying, and Reasoning were explained by students' scores at Primary 3 in these cognitive domains. Results also showed that the R-squared values at Secondary 3 in the Knowing, Applying, and Reasoning cognitive domains were0.403, 0.406, and 0.385 respectively. This means between 38.5% and 40.3% of variances of Secondary 3 scores in Knowing, Applying, and Reasoning were explained by students' scores at Primary 3 and at Primary 6 in these cognitive domains (Table 6).

Table 6. R-Squared Values of Predicted Cognitive Domains at P6 and S3

Achievement in Cognitive Domains	R-Squared	S.E.
P6 Knowing	0.457	0.025
P6 Applying	0.474	0.025
P6 Reasoning	0.460	0.025
S3 Knowing	0.403	0.026
S3 Applying	0.406	0.026
S3 Reasoning	0.385	0.026

The total effects on Knowing (0.258), Applying (0.254), and on Reasoning (0.263) at Primary 6byscores in the Knowing domain at Primary 3 was greater than the effects of the other cognitive domains at Primary 3.The total effects on Knowing (0.244), Applying (0.237), and Reasoning (0.214) at Primary 6 by scores in the Applying domain at Primary 3 were slightly lower but still strong. The total effects on Knowing (0.250), Applying (0.255), and on Reasoning (0.236) at Secondary 3ofscores in the Knowing domain at Primary 3 was substantially greater than effects of the other cognitive domains at Primary 3. Similarly, the total effects on Knowing (0.288), Applying (0.306), and on Reasoning (0.308) at Secondary 3 of achievement in the Knowing domain at Primary 6 was substantially greater than effects of the other cognitive domains at Primary 3 (Table 7b). These results highlighted the significant importance in building strong foundations at primary years on students' proficiency in factual recall, undertaking mathematical computation, measurement and classification (i.e. Knowing) to later mathematics achievements in cognitive knowledge and skills.

The total effects on Primary 6 Knowing (0.168), Applying (0.183), and Reasoning (0.179) of achievement at Primary 3 Reasoning was comparatively weaker, although statistically significant, than the total effects of the other Primary 3 cognitive domains (Table 7a). The total effect on Secondary 3 Knowing (0.164), Applying (0.168), and Reasoning (0.172) of achievement at Primary 3 Reasoning was comparable to total effects of Primary 3 Applying and both were weaker than total effects of Primary 3 Knowing. Further, total effect on Secondary 3 Knowing (0.099), Applying (0.091), and Reasoning (0.091) of achievement at Primary 6 Reasoning was weak and not statistically significant (Table 7b). These results imply that there is no direct contribution of proficiency in mathematics reasoning skills at primary levels (Primary3 and 6) to their later mathematics achievement at secondary level (Secondary 3). This is an interesting finding which might mean either 'reasoning' is not well measured (or a wrong classification for what the selected items infer) and does not capture what 'reasoning' is to younger children, or that 'reasoning' needs to be further understood and developed at the younger year levels to ensure it relates to mathematics achievement at later year levels.

Table 7a. Total Effect on Cognitive Domains at P6

	Dependent Variables		
	Knowing	Applying	Reasoning
P3 Knowing	0.258 (0.075)	0.254 (0.075)	0.263 (0.076)
P3 Applying	0.244 (0.077)	0.237 (0.077)	0.214 (0.078)
P3 Reasoning	0.168 (0.042)	0.183 (0.041)	0.179 (0.042)

Notes: Standardized (STDYX) estimated parameters were reported. All estimates significant at $\alpha = 0.05$.

Table 7b. Total Effect on Secondary 3 Cognitive Domains

	Dependent Variables		
	Knowing	Applying	Reasoning
P3 Knowing	0.250 (0.062)	0.255 (0.057)	0.236 (0.059)
P3 Applying	0.141 (0.054)	0.139 (0.044)	0.144 (0.049)
P3 Reasoning	0.164 (0.044)	0.168 (0.024)	0.172 (0.045)
P6 Knowing	0.288 (0.084)	0.306 (0.082)	0.308 (0.083)
P6 Applying	0.182 (0.097)	0.188 (0.095)	0.179 (0.096)
P6 Reasoning	0.099 (0.069)	0.091 (0.068)	0.091 (0.069)

Notes: Standardized (STDYX) estimated parameters were reported. All estimates significant at $\alpha = 0.05$.

Growth in Mathematics Content Domains

In the path model of content domains, achievements in the domains at earlier year levels were used to predict achievements at later year levels. Results are presented in Table 8 and in Figure 4. The path model had good fit to the data: Both CFI and TLI took the value of 1.000; RMSEA was 0.000, and SRMR was 0.002, and Chi Squared value was 2.146 (d.f. = 5, P = 0.829). Results of the analysis showed that achievement in the Number domain at Primary 3 was the strongest predictor of achievements at all content domains at Primary 6. Primary 3 achievement in Number domain significantly predicted achievements in all content domains at Primary 6, but there was no significant direct effect on any content domain at Secondary 3. Achievement in the Shape and Space at Primary 3significantly predicted achievement in Number, Shape and Space, and Algebra at Primary 6, but not at Secondary 3. Achievement in Measures at Primary 3 significantly predicted Number, Shape and Space, and Measures at Primary 6, but not at Secondary 3. Achievement at Number domain at Primary 6 was the only significant predictor of achievements at all content domains at Secondary 3. Achievements of the other domain at Primary 6 affected Secondary 3 achievements only indirectly via their correlations with the Number domain achievement.

Table 8. Path Coefficient and Standard Errors (Cognitive Domains)

STDYX	Path Coeff.	S.E.	Est./S.E.	Prob.
P6 Number ON				
P3 Number	0.338	0.040	8.530	p <0.05
P3 Measures	0.183	0.046	3.965	p <0.05
P3 Shape and Space	0.150	0.046	3.257	p <0.05
P3 Data Handling	0.020	0.033	0.618	NS
P6 Shape and Space　ON				
P3 Number	0.276	0.043	6.453	p <0.05
P3 Measures	0.167	0.049	3.398	p <0.05
P3 Shape and Space	0.144	0.049	2.935	p <0.05
P3 Data Handling	0.024	0.035	0.676	NS
P6 Measures　ON				
P3 Number	0.305	0.041	7.362	p <0.05
P3 Measures	0.197	0.048	4.095	p <0.05
P3 Shape and Space	0.118	0.048	2.456	p <0.05
P3 Data Handling	0.016	0.035	0.468	NS
P6 Algebra ON				
P3 Number	0.246	0.045	5.462	p <0.05
P3 Measures	0.097	0.052	1.851	NS
P3 Shape and Space	0.147	0.052	2.838	p <0.05
P3 Data Handling	0.047	0.037	1.259	NS
P6 Data Handling　ON				
P3 Number	0.192	0.048	3.979	p <0.05
P3 Measures	0.061	0.055	1.108	NS
P3 Shape and Space	0.092	0.055	1.659	NS
P3 Data Handling	0.038	0.040	0.970	NS
S3 Number ON				
P6 Number	0.391	0.064	6.071	p <0.05
P6 Measures	0.067	0.051	1.318	NS
P6 Shape and Space	0.086	0.044	1.982	p <0.05
P6 Algebra	0.032	0.041	0.794	NS
P3 Number	0.081	0.040	2.045	p <0.05

P3 Measures	0.043	0.045	0.955	NS
P3 Shape and Space	0.060	0.044	1.379	NS
P3 Data Handling	0.015	0.031	0.481	NS
S3 Shape and Space ON				
P6 Number	0.399	0.065	6.151	p <0.05
P6 Measures	0.059	0.051	1.148	NS
P6 Shape and Space	0.078	0.044	1.748	NS
P6 Data Handling	0.007	0.011	0.660	NS
P6 Algebra	0.025	0.042	0.606	NS
P3 Number	0.080	0.040	1.975	p <0.05
P3 Measures	0.041	0.045	0.928	NS
P3 Shape and Space	0.067	0.044	1.532	NS
P3 Data Handling	0.017	0.031	0.548	NS
S3 Algebra ON				
P6 Number	0.413	0.067	6.152	p <0.05
P6 Measures	0.045	0.053	0.852	NS
P6 Shape and Space	0.061	0.046	1.325	NS
P6 Data Handling	-0.001	0.019	-0.031	NS
P6 Algebra	0.022	0.043	0.498	NS
P3 Number	0.093	0.041	2.237	p <0.05
P3 Measures	0.019	0.046	0.403	NS
P3 Shape and Space	0.064	0.045	1.405	NS
P3 Data Handling	0.017	0.033	0.515	NS
S3 Data Handling ON				
P6 Number	0.329	0.071	4.613	p <0.05
P6 Measures	0.051	0.056	0.914	NS
P6 Shape and Space	0.093	0.049	1.902	NS
P6 Data Handling	0.005	0.023	0.200	NS
P6 Algebra	0.009	0.046	0.204	NS
P3 Number	0.076	0.044	1.743	NS
P3 Measures	0.028	0.049	0.582	NS
P3 Shape and Space	0.070	0.048	1.458	NS
P3 Data Handling	0.022	0.034	0.653	NS

P3 Number WITH				
P3 Shape and Space	0.689	0.018	38.730	p <0.05
P3 Measures	0.724	0.016	44.667	p <0.05
P3 Data Handling	0.477	0.026	18.338	p <0.05
P3 Shape and Space WITH				
P3 Measures	0.774	0.014	56.926	p <0.05
P3 Data Handling	0.573	0.023	25.363	p <0.05
P3 Measures WITH				
P3 Data Handling	0.511	0.025	20.286	p <0.05
P6 Measures WITH				
P6 Shape and Space	0.662	0.019	34.832	p <0.05
P6 Data Handling	0.320	0.030	10.520	p <0.05
P6 Algebra	0.426	0.028	15.338	p <0.05
P6 Shape and Space WITH				
P6 Data Handling	0.448	0.027	16.607	p <0.05
P6 Algebra	0.508	0.025	20.151	p <0.05
P6 Data Handling WITH				
P6 Algebra	0.447	0.027	16.419	p <0.05
S3 Number WITH				
S3 Algebra	0.812	0.011	71.263	p <0.05
S3 Shape and Space	0.944	0.004	262.333	p <0.05
S3 Data Handling	0.751	0.015	50.770	p <0.05
S3 Algebra WITH				
S3 Shape and Space	0.811	0.011	71.140	p <0.05
S3 Data Handling	0.656	0.019	33.804	p <0.05
S3 Shape and Space WITH				
S3 Data Handling	0.785	0.013	61.313	p <0.05
P6 Number WITH				
P6 Measures	0.757	0.014	52.583	p <0.05
P6 Shape and Space	0.689	0.018	38.708	p <0.05
P6 Data Handling	0.325	0.030	10.704	p <0.05
P6 Algebra	0.694	0.018	38.989	p <0.05

Notes: Standardized (STDYX) estimated parameters are reported. NS stands for not significant at $\alpha = 0.05$

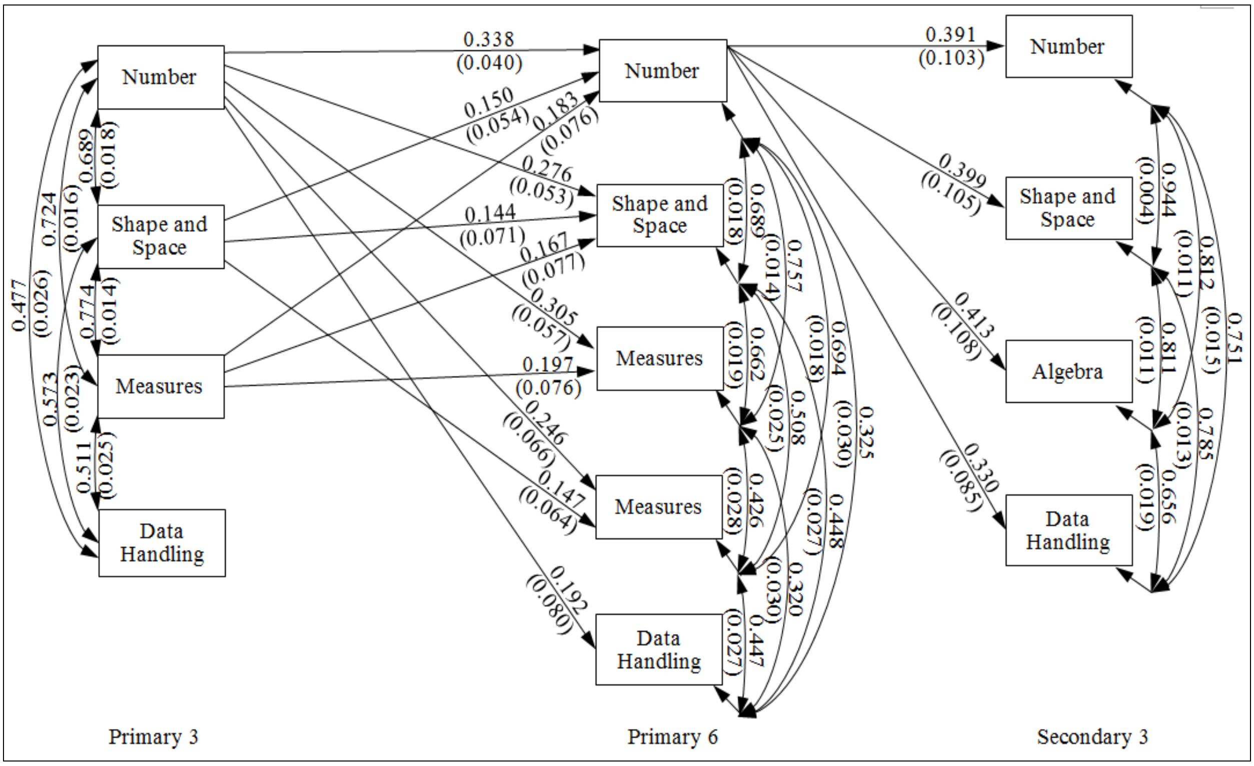


Figure 4. The standardized result (STDYX) of Path model in Content Domain

Note: Only paths coefficient significant at α = 0.05 are shown. CFI: 1.000, TLI: 1.000, RMSEA: 0.000, SRMR: 0.000, Chi Square: 0.136 (d.f. = 1, P = 0.713)

As presented in Table 9, the R-squared values of mathematics achievement in different content domains at Primary 6were all statistically significant. The R-squared values for Number, Measures, Shape and Space, Data Handling, and Algebra at Primary 6were 0.387, 0.331, 0.300, 0.118, and 0.225 respectively. This means between 11.8% and 38.7% of variances in Number, Measures, Shape and Space, Data Handling, and Algebra at Primary 6 were explained either directly or indirectly by students' achievement at Primary 3 in these content domains. Results also showed that the R-squared values for Number, Shape and Space, Data Handling, and Algebra at Secondary 3 were 0.452, 0.446, 0.344, and 0.410 respectively. This means between 34.4% and 45.2% of variances in Number, Shape and Space, Data Handling, and Algebra at Secondary 3 were explained either directly or indirectly by students' achievement at Primary 6 and Primary 3 in the content domains.

Table 9. R-Squared Values of Content Domains at P6 and S3

Content Domains	R-Square	S.E.
P6 Number	0.387	0.028
P6 Measures	0.331	0.033
P6 Shape and Space	0.300	0.030
P6 Data Handling	0.118	0.026
P6 Algebra	0.225	0.031
S3 Number	0.452	0.030
S3 Shape and Space	0.446	0.031
S3 Data Handling	0.344	0.034
S3 Algebra	0.410	0.032

Presented in Table 10a are the total effects of Mathematics content domains at Primary 3 on the domains at Primary 6. Achievement at the Number domain at Primary 3 had the strongest total effect on all content domains at Primary 6. Primary 3 Shape and Space had significant total effect on Primary 6 Number, Shape and Space, and Algebra. Primary 3 Measures had significant total effect on Primary 6 Number, Shape and Space and Measures. There was no effect of Primary 3 Data Handling on any of the content domains at Primary 6.

The total effects of Mathematics content domains at Primary 3 and at Primary 6 on the domains at Secondary 3 are presented in Table 10b. Although there was no direct effect of Primary 3 achievements in Number, Shape and Space, and Measures on students' achievements at Secondary 3 in Number, Shape and Space, Algebra, and Data Handling content domains, there was significant total effect, ranging from 0.117 to 0.267, possibly channeled through achievements in Primary 6 Number, which had total effect in the order of 0.4 on the content domains at Secondary 3. These findings suggest the importance of an orderly sequence of teaching in Number. The lack of direct effects from Shape, Measure, Algebra and Data Handling at Primary 6 to achievement at Secondary 3 needs further investigation. It might be that the nature of the curriculum changes dramatically from Primary 6 to Secondary 3 and therefore the predictive salience of these content domains from earlier year levels (Primary 3 and 6) is weakened. This suggests that there may be a disjoint in the spiral curriculum teaching approach advocated in Hong Kong.

Table 10a. Total Effect on Mathematics Content Domains (Primary 6)

	Number	Shape and Space	Measures	Algebra	Data Handling
P3 Number	0.338	0.276	0.305	0.246	0.192
P3 Shape and Space	0.150	0.144	NS	0.147	NS
P3 Measures	0.183	0.167	0.197	NS	NS
P3 Data Handling	NS	NS	NS	NS	NS

Notes: Standardized (STDYX) estimated parameters are reported. NS stands for not significant at $\alpha = 0.05$.

Table 10b. Total Effect on Mathematics Content Domains (Secondary 3)

	Number	Shape and Space	Algebra	Data Handling
P3 Number	0.266	0.262	0.267	0.231
P3 Shape and Space	0.143	0.148	0.141	0.139
P3 Measures	0.145	0.142	0.118	0.117
P3 Data Handling	NS	NS	NS	NS
P6 Number	0.391	0.399	0.413	0.329
P6 Shape and Space	NS	NS	NS	NS
P6 Measures	NS	NS	NS	NS
P6 Algebra	NS	NS	NS	NS
P6 Data Handling	NS	NS	NS	NS

Notes: Standardized (STDYX) estimated parameters are reported. NS stands for not significant at $\alpha = 0.05$.

Conclusion

The purpose of this study was to model long-term mathematics growth over 6 years from Primary 3 to Secondary 3. The sample comprised 866 students whose mathematics achievements were tracked longitudinally. Two perspectives were used, namely, mathematics growth in cognitive domains and in content domains. Scales based on students' responses to individual items at the Territory-wide Systems Assessment at Primary 3, Primary 6 and Secondary 3 were formed by assessment items using Multidimensional Partial Credit Rasch model by classifying items according to cognitive domains of Knowing, Applying, and Reasoning, and in parallel according to content domains of Number, Shape and Space, Measures, Algebra, and Data Handling. A modified autoregressive cross-lagged design was used whereby students' mathematics achievement in each domain at Primary 6 was predicted by their Primary 3 achievements, and achievement at Secondary 3 was predicted by their achievements at both Primary 3 and Primary 6.

Findings of this study must be interpreted in the context of the limitations of this study. Due to ethical principles observed by the Hong Kong government, no student background information including their gender, age, and class membership, was collected. It was therefore not possible to explain students' mathematics growth in terms of these background variables. Further, given that matrix sampling was used in the assessment design involving four booklets at each testing occasion, the number of students per school sharing the same three booklets over the three occasions could be very small (on average between two and three students per school). The small sample size per school prohibited multilevel modeling to be used. In addition, the matrix sampling meant that it was not possible to combine the two categorisations (cognitive x content) of mathematics domains in the analysis to investigate the their interaction effect, making it impossible to address such important question as "What is the predictive power of achievement in Knowing Numbers at Primary 3 for Primary 6 mathematics achievement?"

Notwithstanding the limitations, there are many strengths to this study, including the simultaneous treatment of multiple domains which enables the cross-lagged relationships across three occasions to be explored. Further, understanding students' mathematic learning from multiple perspectives can make an essential contribution to instructional designs.

Results found that, inter alia, achievement in Knowing at Primary 3 was the most powerful predictor for achievement in all mathematics cognitive domains at Primary 6, and at Secondary 3. Further, achievement in Number at Primary 3 was the most powerful predictor for achievement in all content domains at Primary 6. In turn, achievement in Number at Primary 6 was the only direct predictor for achievement in all other content domains at Secondary 3. These findings highlighted the importance to later mathematics achievement of solid foundations in memory and familiarity of mathematics basic facts (e.g. Table of Multipliers), rules (e.g. the four fundamental operations of arithmetic), operations (e.g. the use of brackets in solving equations) at early year levels.

It is also interesting to note that in comparing the path models across time for both cognitive domains and content domains more direct and indirect paths were significant for the cognitive domains analyses. All three cognitive domains at Primary 3 predicted the cognitive domains at Primary 6, while both Knowing and Applying at Primary 6 predicted the three cognitive domains at Secondary 3. In contrast, for the content domains, Number, Shape, Measure and Data Handling, only three had direct effects at Primary 6, and this was reduced to one (Number) from Primary 6 to Secondary 3. These findings perhaps suggest that there is a stronger consistency in the mathematical cognitive development of children across time than there is in the development of content knowledge. As suggested above, these findings may reflect that the cognitive load demanded of children as material becomes more difficult results in enhanced cognitive ability over time. However, conversely, the development of content knowledge over time may not be directly related to earlier content knowledge because of structural issues in syllabi continuity. This is an issue for further study, and, in particular with an independent measure of mathematical achievement to which both cognitive dimensions and content dimensions can be related.

This study did not find critical importance in early achievement in reasoning skills (e.g. solving unfamiliar problems using mathematics) to later mathematics achievement. With the caveat that this result might very well due to under development of students' mental ability at young age of Primary 3 and 6, instructional designs should avoid over-emphasis on introducing reasoning skills at early primary year levels.

References

Adams, R. J., Wilson, M. R., & Wang, W. C. (1997). The multidimensional random coefficients multinomial logit model. *Applied Psychological Measurement, 21*, 1-23.

Aubrey, C., Godfrey, R., & Dahl, S. (2006). Early mathematics development and later achievement: Further evidence. *Mathematics Education Research Journal, 18*(1), 27-46.

Aubrey, C. (2003). The development of children's early numeracy through key stage 1. *British Educational Research Journal, 29*(6), 821-840. doi:10.1080/0141192032000137321

Aunio, P. (2010). Predicting children's mathematical performance in grade one by early numeracy. *Learning and Individual Differences, 20*(5), 427-435. doi:10.1016/j.lindif.2010.06.003

Bobis, J., & Gould, P. (1999). 'The mathematical achievement of children in the count me in too program'. In J. Truran, & K. Truran (Eds.), *Making the difference* (pp. 84-90). Adelaide: Mathematics Education Research Group of Australasia.

Bollen, K. A., & Curran, P. J. (2004). Autoregressive latent trajectory (ALT) models a synthesis of two traditions. *Sociological Methods and Research, 32*(3), 336-383.

Byrne, B. M. (2012). *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. New York: Routledge Academic.

Curriculum Development Council. (2002). *Mathematics Education Key Learning Area Curriculum Guide (Primary 1 – Secondary 3).* Hong Kong: The Printing Department.

Desoete, A. (2009). Classification, seriation, and counting in grades 1, 2, and 3 as two-year longitudinal predictors for low achieving in numerical facility and arithmetical achievement? *Journal of Psychoeducational Assessment, 27*(3), 252-264. doi:10.1177/0734282908330588

DiPerna, J. C. (2005). A model of academic enablers and mathematics achievement in the elementary grades. *Journal of School Psychology, 43*(5), 379-392. doi:10.1016/j.jsp.2005.09.002

Duncan, G. J. (2007). School readiness and later achievement. *Developmental Psychology, 43*(6), 1428-1446. doi:10.1037/0012-1649.43.6.1428

Education Commission. (2000). *Education blueprint for the 21st century: Learning for life, Learning through life–Reform proposals for education system in Hong Kong*. Hong Kong: Printing Department.

Geary, D. C. (2011). Cognitive predictors of achievement growth in mathematics: A 5-year longitudinal study. *Developmental Psychology, 47*(6), 1539-1552. doi:10.1037/a0025510

Hemmings, B., Grootenboer, P., & Kay, R. (2011). Predicting mathematics achievement: The influence of prior achievement and attitudes. *International Journal of Science and Mathematics Education, 9*(3), 691–705.

Hong Kong Examinations and Assessment Authority (2014). TSA: Introduction. Retrieved 31 Nov 2014 from http://www.bca.hkeaa.edu.hk/web/TSA/en/Introduction.html

Jordan, N. C. (2009). Early math matters: Kindergarten Number competence and later mathematics outcomes. *Developmental Psychology, 45*(3), 850-867. doi:10.1037/a0014939

Kyttälä, M., & Björn, P. M. (2010). Prior mathematics achievement , cognitive appraisals and anxiety as predictors of Finnish students ' later mathematics performance and career orientation. *Educational Psychology : An International Journal of Experimental*, *30*(4), 431–448.

LeFevre, J. A., Fast, L., Skwarchuk, S. L., Smith-Chant, B. L., Bisanz, J., Kamawar, D., Penner-Wilger, M. (2010). Pathways to mathematics: Longitudinal predictors of performance. *Child Development*, *81*(6), 1753-1767.

Locuniak, M. N. (2008). Using kindergarten Number sense to predict calculation fluency in second grade. *Journal of Learning Disabilities, 41*(5), 451-459. doi:10.1177/0022219408321126

Marsh, H.W., & Martin, A.J. (2011). Academic self-concept and academic achievement: Relations and causal ordering. *British Journal of Educational Psychology, 81*(1), 59-77.

Masters, G.N. (1982). A Rasch model for partial credit scoring. *Psychometrika, 47,* 149-174.

Mok, M. M. C. & Xu, K. (2013). Using Multidimensional Rasch to enhance measurement precision: Initial results from simulation and empirical studies. *Journal of Applied Measurement, 14*(1), 27-43.

Mullis, I. V. S., Martin, M. O., Ruddock, G. J., O'Sullivan, C. Y., & Preuschoff, C. (2009). TIMSS 2011 Mathematics Framework. In *TIMSS 2011 Assessment Frameworks* (pp. 19–46). TIMSS & PIRLS International Study Center Lynch School of Education, Boston College.

Muthén, L. K., & Muthén, B. O. (1998-2012). *Mplus statistical analysis with latent variables: User's Guide*. Los Angeles, CA: Muthén & Muthén.

Pagani, L. S. (2010). School readiness and later achievement: A french canadian replication and extension. *Developmental Psychology, 46*(5), 984-994. doi:10.1037/a0018881

Passolunghi, M. C., & Lanfranchi, S. (2012). Domain-specific and domain-general precursors of mathematical achievement: A longitudinal study from kindergarten to first grade. *British Journal of Educational Psychology, 82*(1), 42-63. doi:10.1111/j.2044-8279.2011.02039.x

Pearn, C. (1998). 'Is there a need for a mathematics intervention program in grades 3 and 4?' In M. Goos, & E. Warren (Eds.), *Teaching mathematics in new times*(pp. 444-451). MERGA, Brisbane,: Proceedings of 21st Mathematics Education Research Group of Australasia.

Reynolds, A. J. (1991). The middle schooling process: Influences on science and mathematics achievement from the longitudinal study of American youth. *Adolescence*, *26*(101), 133–158.

Romano, E. (2010). School readiness and later achievement: Replication and extension using a nationwide Canadian survey. *Developmental Psychology, 46*(5), 995-1007. doi:10.1037/a0018880

Wu, M. (2005). The role of plausible values in large-scale surveys. *Studies in Educational Evaluation, 31*, 114-128.

Wu, M. L., Adams, R. J., Wilson, M. R., & Haldane, S. (2007). ConQuest (Version 2.0) [Computer Software]. Camberwell, Australia: ACER.

Yates, S. M. (2000). Task involvement and ego orientation in mathematics achievement: A three year follow-up. *Issues in Educational Research*, *10*(1), 77–91.

Corresponding author's email：　mmcmok@ied.edu.hk