Chapter 4
Booklet Design for a Longitudinal Study: Measuring Progress in New-Immigrant Children's Mathematics Achievement

Pei-Jung Hsieh

Abstract Because of the increasing number of new-immigrant children in Taiwan, we conducted a large-scale assessment of mathematics achievement at the Grade 4 level in 2012 and performed a follow-up of a sample of panel members in 2014. The objective of the study was to develop a valid instrument in order to measure and investigate the average new-immigrant student's progress between Grade 4 and Grade 6 in mathematical literacy. We compiled 78 selected-response items into a set of 13 booklets. All the test items were selected from the Taiwan Assessment of Student Achievement (TASA 2012), which is the largest nationally representative and continuing assessment in Taiwan. Each item is designed to measure one of the four mathematics content areas, which are number and measurement, geometry, statistics and probability, and algebra. We propose three rules for identifying and selecting appropriate items. First, the number of anchor items is 39 (50 %), providing the basis for equating scores on different grades. Second, we designated as high priorities Grade 4 items with a difficulty parameter larger than −0.5 and Grade 6 items with a parameter between −1 and 0.5. Third, the proportions of each content area should be similar to those listed in the General Guidelines of Grades 1–9 Curriculum for Mathematics. We applied a three-parameter logistic model and used PARSCALE to estimate the item parameters. After these procedures, each test block contained six items, and each booklet comprised four blocks based on the balanced incomplete block design. The mean difficulty of the items in each block ranged from 0.053 to 0.153. Consequently, the mean difficulty of the items in each booklet ranged from 0.083 to 0.128. The distributions of items across the content areas were 64.10 %, 19.23 %, 5.13 %, and 11.54 %, mostly corresponding to the curriculum framework. This study demonstrated that a detailed consideration of the percentage of anchor items, the range of item difficulties, and the distribution of content areas can be useful for constructing a measurement tool in a longitudinal study.

Keywords Balanced incomplete block design · Achievement growth · Panel study

P.-J. Hsieh (✉)
National Academy for Educational Research, New Taipei City, Taiwan
e-mail: pjh@mail.naer.edu.tw

Q. Zhang and H. Yang (eds.), *Pacific Rim Objective Measurement Symposium (PROMS) 2014 Conference Proceedings*,
DOI 10.1007/978-3-662-47490-7_4

4.1 Introduction

With the continuous decrease in the fertility rate of Taiwanese citizens, the proportion of new-immigrant children among primary school students continues to increase every year. In the 2004–2005 school-year, there were 1,883,533 pupils, and this number dropped to 1,297,120 in the 2013–2014 school year, a decrease of 31.13 %. However, the number of new-immigrant pupils increased from 40,907 in the 2004–2005 academic year to 157,431 in the 2013–2014 academic year, an increase from 2.17 to 12.14 %. The highest number of these children had mothers with Chinese nationality, followed by mothers with Vietnamese and Indonesian nationalities (Ministry of Education 2014). Following school admission, additional attention should be paid to education-related topics such as academic performance, parent–teacher communication, and adjustment to school by children born in transnational marriage families.

The National Academy for Educational Research conducted a large-scale investigation for fourth and sixth-grade students in 2012. The results showed that children from new-immigrant families in the fourth and sixth grades performed significantly worse in mathematics than children from nonimmigrant families. However, the difference in performance seemed to decline with age. In addition, growth in academic performance could vary among children based on the differences in their family socioeconomic status (Wang et al. 2012). Besides the correlation between family socioeconomic status and performance in mathematics, which could change with time, the relationship between psychological variables and academic achievement is a subject of concern in the educational sector. To further understand the effects of the environmental context of new immigrant children and time on their performance in mathematics and to find more convincing evidence, we resampled the 2012 fourth-grade subjects to conduct a long-term longitudinal study on the same population of new-immigrant children in 2014.

To understand the growth trend in new-immigrant children's performance in mathematics, to investigate the casual pathways of the children's performance in mathematics, and to analyze the effect of education policies on performance in mathematics, this study developed reliable and valid evaluation tools to accurately measure the growth in the mathematics performance of new immigrant children in fourth to sixth grades.

4.2 Content Standards

We developed a mathematics framework based on the General Guidelines of Grades 1–9 Curriculum for Elementary and Junior High School Education. A mathematics curriculum equips students with an understanding of the basic concepts of figures, shapes, and quantities as well as the ability to calculate and organize and to apply such knowledge and skills in daily life. It also enables

comprehending the principles of reasoning and problem solving, the ability to elaborate clearly on mathematics-related concepts, and making appropriate connections among materials and contents between this and other learning areas. The mathematics curriculum for Grades 1–9 is divided into four stages: Stage 1 begins in Grade 1 and ends in Grade 3; Stage 2 begins in Grade 4 and ends in Grade 5; Stage 3 begins in Grade 6 and ends in Grade 7; and Stage 4 begins in Grade 8 and ends in Grade 9 (Ministry of Education 2006).

4.3 Test Item Bank

The mathematics items used in the 2012 large-scale assessment of new-immigrant children were constructed by the TASA, which is the largest nationally representative and continuing assessor of Taiwanese students' knowledge and skill-sets in five subject areas. Assessments are conducted periodically in mathematics, Mandarin, English, science, and social science. TASA assessments began in 2005. (For a detailed description of the TASA assessment plan, see Table 4.1)

TASA classifies mathematics assessment questions into two dimensions: content area and mathematical complexity. Each question is designed to measure one of four mathematics content areas: (a) number and measurement, (b) geometry, (c) statistics and probability, and (d) algebra. Moreover, items are classified according to three types of mathematical abilities: conceptual understanding, procedural knowledge, and problem solving.

The distribution of items among the various mathematical content areas and mathematical complexities reflects the relative proportion of the mathematics curriculum. In 2012, there were 65 selected-response test items and 13 constructed-response test items for Grade 4 as well as Grade 6 students. Table 4.2 lists the distribution of selected-response items, and Table 4.3 shows the

Table 4.1 TASA assessment schedule

Year/grade	Mathematics	Mandarin	English	Science	Social science
2014	11	11	11	11	11
2013	8	8	8	8	8
2012	4, 6	4, 6	6	4, 6	6
2011	11	11	11	11	11
2010	8	8	8	8	8
2009	4, 6	4, 6	6	4, 6	6
2007	4, 6, 8, 11	4, 6, 8, 11	4, 6, 8, 11	4, 6, 8, 11	6, 8, 11
2006	4, 6, 8, 11	4, 6, 8, 11	4, 6, 8, 11	4, 6, 8, 11	6, 8, 11
2005	6	6	6		

Table 4.2 Distribution of selected-response items for Grade 4 in 2012

Content areas	Types of mathematical abilities			Number of items
	Conceptual understanding	Procedural knowledge	Problem solving	
Number and measurement	17	12	14	43
Geometry	4	4	4	12
Statistics and probability	1	1	2	4
Algebra	2	2	2	6
	24	19	22	65

Table 4.3 Distribution of constructed-response items for Grade 4 in 2012

Content areas	Types of mathematical abilities			Number of items
	Conceptual understanding	Procedural knowledge	Problem solving	
Number and measurement	0	1	7	8
Geometry	0	1	2	3
Statistics and probability	0	0	1	1
Algebra	0	0	1	1
	0	2	11	13

Table 4.4 Distribution of selected-response items for Grade 6 in 2012

Content areas	Types of mathematical abilities			Number of items
	Conceptual understanding	Procedural knowledge	Problem solving	
Number and measurement	14	10	16	40
Geometry	5	3	5	13
Statistics and probability	0	0	3	3
Algebra	2	2	5	9
	21	15	29	65

distribution of constructed-response test items for Grade 4 in 2012. In addition, Table 4.4 details the distribution of selected-response items, and Table 4.5 displays the distribution of constructed-response test items for Grade 6 in 2012. In order to link the test scale scores between Grade 4 and Grade 6, we embedded 17 common selected-response items into both item pools.

TASA and our 2012 large-scale assessment of new-immigrant children use a balanced incomplete block (BIB) design to assign blocks or groups of selected-response cognitive items to student booklets. A BIB design satisfies four

Table 4.5 Distribution of constructed-response items for Grade 6 in 2012

Content areas	Types of mathematical abilities			Number of items
	Conceptual understanding	Procedural knowledge	Problem solving	
Number and measurement	0	0	6	6
Geometry	0	1	1	2
Statistics and probability	0	1	1	2
Algebra	0	1	2	3
	0	3	10	13

Table 4.6 TASA balanced incomplete block booklet design

Booklet version	Position			
	Position 1 cognitive block	Position 2 cognitive block	Position 3 cognitive block	Position 4 cognitive block
S1	M3	M2	M1	M10
S2	M4	M5	M10	M6
S3	M7	M10	M8	M9
S4	M10	M11	M12	M13
S5	M5	M1	M9	M11
S6	M2	M6	M11	M7
S7	M11	M3	M4	M8
S8	M1	M4	M7	M12
S9	M8	M12	M5	M2
S10	M12	M9	M6	M3
S11	M13	M7	M3	M5
S12	M9	M13	M2	M4
S13	M6	M8	M13	M1

conditions: every treatment in the booklet design is covered at most only once in a booklet; every treatment in the booklet design appears with equal frequency across all booklets; every booklet has an identical length, containing the same number of clusters; and every pair of treatments in the booklet design occurs together in the booklets with equal frequency (Frey et al. 2009). We assigned 65 selected-response items into a set of 13 booklets, and each booklet comprised four blocks (i.e., every student response to 20 items). The BIB booklet design (Table 4.6) enabled TASA and our 2012 large-scale assessment of new-immigrant children to sample a sufficient number of students to obtain precise results for each test.

Table 4.7 Distribution of selected-response items for new-immigrant children in 2014

Content areas	Types of mathematical abilities			Number of items
	Conceptual understanding	Procedural knowledge	Problem solving	
Number and measurement	10 + 10	8 + 3	11 + 8	29 + 21
Geometry	1 + 4	2 + 2	2 + 4	5 + 10
Statistics and probability	1 + 0	0 + 0	1 + 2	2 + 2
Algebra	0 + 2	1 + 1	2 + 3	3 + 6
	12 + 16	11 + 6	16 + 17	39 + 39

Note The number before the + sign represents the number of Grade 4 items, whereas the number after the + sign represents the number of Grade 6 items

4.4 Method

Because of budget limitations, we chose only selected-response items from 2012 for the 2014 booklets. We propose three rules for identifying and selecting appropriate items. First, the percentage of anchor items must be 50 %, providing the basis for equating scores on different grades. Second, the Item Response Theory model of TASA is constrained to have a mean ability of zero, and thus, the TASA sample represents the whole population of Taiwanese students with a higher mathematics ability compared with new-immigrant children. Therefore, the Grade 4 items with a difficulty parameter larger than −0.5 and the Grade 6 items with a parameter between −1 and 0.5 are considered high priority. Third, the proportion of each content area should be similar to that in the General Guidelines of Grades 1–9 Curriculum for Mathematics. PARSCALE 4 is used to estimate the item parameters.

4.5 Results

Totally, we selected 78 selected-response items. The distribution of items is listed in Table 4.7. The item distribution across the content areas was 64.10, 19.23, 5.13, and 11.54 %, which is relatively similar to the original distribution of the Grade 4 and Grade 6 items. Following the 2012 BIB design, we compiled all of the items into a set of 13 booklets, and each booklet comprised four blocks (i.e., every new-immigrant student response to 24 selected-response items in 40 min). The mean difficulty in the items on each block ranged from 0.053 to 0.153, and consequently, in each booklet it ranged from 0.083 to 0.128 (Table 4.8).

Table 4.8 The mean difficulty of the items on each booklet

Booklet version	Mean difficulty on each block				Mean difficulty on each booklet
	Position 1 cognitive block	Position 2 cognitive block	Position 3 cognitive block	Position 4 cognitive block	
S1	M3 0.153	M2 0.100	M1 0.152	M10 0.093	0.125
S2	M4 0.092	M5 0.128	M10 0.093	M6 0.053	0.092
S3	M7 0.092	M10 0.093	M8 0.097	M9 0.143	0.106
S4	M10 0.093	M11 0.087	M12 0.132	M13 0.117	0.107
S5	M5 0.128	M1 0.152	M9 0.143	M11 0.087	0.128
S6	M2 0.100	M6 0.053	M11 0.087	M7 0.092	0.083
S7	M11 0.087	M3 0.153	M4 0.092	M8 0.097	0.107
S8	M1 0.152	M4 0.092	M7 0.092	M12 0.132	0.117
S9	M8 0.097	M12 0.132	M5 0.128	M2 0.100	0.114
S10	M12 0.132	M9 0.143	M6 0.053	M3 0.153	0.120
S11	M13 0.117	M7 0.092	M3 0.153	M5 0.128	0.123
S12	M9 0.143	M13 0.117	M2 0.100	M4 0.092	0.113
S13	M6 0.053	M8 0.097	M13 0.117	M1 0.152	0.105

4.6 Conclusion

This study demonstrated that a detailed consideration for the percentage of anchor items, the range of item difficulties, and the distribution of content areas can be useful for constructing a measurement tool in a longitudinal study.

Acknowledgments This research was supported by the Ministry of Science and Technology (MOST 102-2511-S-656-005) and the National Academy for Educational Research.

References

Frey, A., Hartig, J., & Rupp, A. A. (2009). An ncme instructional module on booklet designs in large-scale assessments of student achievement: Theory and practice. *Educational Measurement: Issues and Practice, 28*(3), 39–53. doi:10.1111/j.1745-3992.2009.00154.x.

Ministry of Education. (2006). *General guidelines of grades 1-9 curriculum for elementary and junior high school education*. Taipei, Taiwan: Ministry of Education.

Ministry of Education. (2014). *Overview of new immigrant's students (2013-2014 school year)*. Taipei, Taiwan: Ministry of Education.

Wang, R.-J., Kuo, K.-B., & Cheng, C.-M. (2012). *Factors influencing the academic performance of new-immigrant children in Taiwan: Individuals, families, and schools*. Taipei, Taiwan: National Academy for Educational Research.