Symposium: Big Ideas in School Mathematics

Chair: Yew Hoong Leong
Nanyang Technological University
yewhoong.leong@nie.edu.sg

The "Big Ideas in School Mathematics" (BISM) is a Research Project funded by the Ministry of Education, Singapore, and administered through the Office of Educational Research, National Institute of Education, Nanyang Technological University. The project began in 2020 and its aim is to investigate various areas in relation to teaching towards mathematical Big Ideas in Singapore schools. The study has currency in so far as "Big Ideas" were introduced in the latest Syllabus Revision by the Ministry of Education. There are three sub-studies in the project: the first is on the development of instruments to measure knowledge of BISM among primary- and secondary-level students and teachers; the second is on professional development work for secondary-level teachers on BISM; the third is similar to the second but for primary-level teachers. The papers in this symposium report information and findings on all these sub-studies.

Overview of the Symposium Papers and Presenters

Presenters: Associate Professor Leong Yew Hoong (Chair), Associate Professor Toh Tin Lam (Paper 1), Mr Mohamed Jahabar Jahangeer (Paper 2), Assistant Professor Choy Ban Heng (Paper 3), Professor Berinderjeet Kaur (Paper 4)

Paper 1: Overview of the research project on Big Ideas in School Mathematics

Authors: Toh Tin Lam, Tay Eng Guan, Berinderjeet Kaur, Leong Yew Hoong, Tong Cherng Luen

This paper provides a brief overview of the entire research project and the component sub-studies.

Paper 2: Assessment of Big Ideas in School Mathematics: Exploring an Aggregated Approach

Authors: Mohamed Jahabar Jahangeer, Toh Tin Lam, Tay Eng Guan, Tong Cherng Luen

This paper reports on developments under Sub-study 1. An item from the student BISM instrument will be discussed. It argues for the use of an "aggregated approach" in considering the scores of the student responses.

Paper 3: From Inert Knowledge to Usable Knowledge: Noticing Affordances in Tasks Used for Teaching Towards Big Ideas About Proportionality

Authors: Choy Ban Heng, Yeo Boon Wooi Joseph, Leong Yew Hoong

This paper reports on developments under Sub-study 2. Part of the professional development under this project involved teachers designing their own instructional materials to foreground a targeted Big Idea. Snippets of tasks in these instructional materials will be discussed.

Paper 4: Primary School Teachers Solving Mathematical Tasks Involving the Big Idea of Equivalence

Authors: Berinderjeet Kaur, Tong Cherng Luen, Mohamed Jahabar Jahangeer

This paper reports on developments under Sub-study 3. An item from the teacher BISM instrument will be discussed. Some data on teachers' responses to the item will be shared. There are thus implications to teacher professional development on the Big Idea of Equivalence.

2023. In B. Reid-O'Connor, E. Prieto-Rodriguez, K. Holmes, & A. Hughes (Eds.), *Weaving mathematics education research from all perspectives. Symposium in the proceedings of the 45th annual conference of the Mathematics Education Research Group of Australasia* (pp. 39–55). Newcastle: MERGA.

Assessment of Big Ideas in School Mathematics: Exploring an Aggregated Approach

Mohamed Jahabar Jahangeer
Nanyang Technological University
jahabar.jahangeer@nie.edu.sg

Tin Lam Toh
Nanyang Technological University
tinlam.toh@nie.edu.sg

Eng Guan Tay
Nanyang Technological University
engguan.tay@nie.edu.sg

Cherng Luen Tong
Nanyang Technological University
cherngluen.tong@nie.edu.sg

In this paper we report our development of instruments to measure Big Ideas in school mathematics. In tackling the issue of assessment fatigue among students, we present an aggregated approach to measure students' knowledge of Big Ideas.

There has been little research done on how the knowledge and understanding of Big Ideas can be assessed. In one of the rare examples we could find, Niemi et al. (2006) suggested three main types of assessments to measure Big Ideas in mathematics: basic computation tasks, partially-worked problems (with or without explanations), and explanation tasks. Charles' (2005) definition of a Big Idea in Mathematics as "a statement of an idea that is central to the learning of mathematics, one that links numerous mathematical understandings into a coherent whole" (p.10) implies the need to contrast a task across more than one topic to be able to tease out the use of a Big Idea in the task. We followed this basic principle in constructing an instrument to assess Big Idea 'ability'. An example of an item, consisting of five parts, on Equivalence is shown in Figure 1. We have piloted some of the items which we have constructed. The dimensionality of these items are reported in Jahangeer et al. (2023), a separate paper in this conference. An important consequence from a Rasch analysis was that we could only use Part 5 as a reliable measure of Big Idea 'ability' since within an item, Parts 1 to 3 violate the item independence requirement of a Rasch scale.

Assessment Fatigue

Assessment has always been an integral part of teaching and learning. Analysis of assessment performance is used for a variety of purposes including placement, selection and certification. In many countries, standardised and high stakes assessments are put in place at milestone grades to determine placement and selection of students to the next course of their education. Well-designed assessment tools and analysis can provide accurate information regarding student learning.

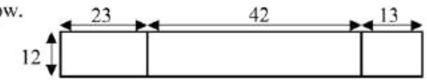
Inaccuracies or deviations from what students have mastered could have been contributed by the students themselves. In particular, the cognitive demand required on students may contribute to them experiencing cognitive fatigue, which naturally affects their overall performance. According to Ackerman and Kanfer (2009), "[a]nticipations of subjective fatigue may lead some individuals to avoid tasks altogether" (p. 176). The duration of an assessment may result in unwilling students not committed to performing to the best of their abilities, affecting the validity of the responses. Thus, a balance between the reliability and validity of the assessment and the duration of assessment without causing a negative anticipation of cognitive fatigue, is an area of worthwhile concern for educators and researchers.

Returning to our attempt to assess Big Idea 'ability', the same consideration of duration of assessment in relation to test validity and reliability arises. Each item of ours necessarily consists of parts to enable a Big Idea to surface across different topics. However, just two items would require at least 30 minutes. A valid Rasch scale would require at least six items to cover a significant range of ability. We derived this based on Andrich's work which, when describing the invariance of appropriate comparison on measures using the Rasch model, used a six-item questionnaire for an example (Andrich, 1988, p. 22).

Part 1: The diagram below shows three rectangles joined together to form a bigger rectangle.

You may use the diagram to fill in the blanks below.

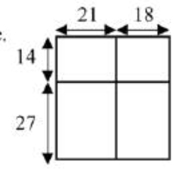
23 x 12 + 42 x 12 + 13 x 12 = ? x 12



Part 2: The diagram below shows four rectangles joined together to form a bigger rectangle.

Fill in the missing numbers below.

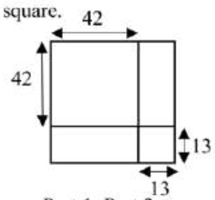
14 x 21 + Δ x 18 + 27 x 21 + 27 x 18 = ♥ x 39



Part 3: The diagram below shows 2 squares and 2 rectangles joined together to form a bigger square.

Fill in the missing numbers below.

55 x 55 = 42 x 42 + 13 x 13 + 13 x Δ x ♥



Part 4: Which of these following statements best describes the common mathematical idea across Part 1, Part 2 and Part 3?

a) I used diagrams for the parts

b) I used Equivalence for the parts

c) I used Guess and Check for the parts

d) I used Proportionality for the parts

e) Others (Please elaborate)

Part 5: The shaded area of the figure below can be used to show a mathematical statement. Which of the following statements matches the shaded part of the figure?

a) (1 + 6) + (2 + 5) + (3 + 4) … + (6 + 1) = 6 x 7

b) 1 + 2 + 3 + … + 6 = (6 x 7) ÷ 2

c) 1 + 2 + 3 + … + 7 = (7 x 8) ÷ 2

d) 1 + 2 + 3 + … + 8 = (8 x 9) ÷ 2

e) 1 + 2 + 3 + … + 7 = (7 x 8)

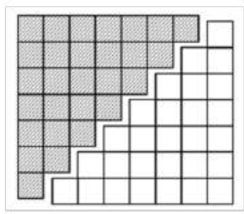


Figure 1. An item comprising 5 parts.

An Aggregated Solution

We base our solution to the conundrum on the methodology and raison d'etre of sampling, i.e., to understand a population, there is no need for every student to complete the entire instrument. For example, the Programme for International Student Assessment (PISA) carried out international standardized testing every three years across various domains. Each domain consists of items which are subdivided into smaller blocks and each student involved in the assessment will be given a booklet made up of a few blocks. PISA made use of 'plausible values' to determine a student's performance and to give a population score instead of an individual score. The successful computation of plausible values, however, requires a deeper knowledge of mathematics which is not accessible to educators, generally. We are proposing a simpler structure that is mathematically easier and can be implemented by educators in schools.

We propose an aggregated structure which involves the creation of a 'Super-Student' (SS). Each SS is made up of four students of similar ability in Mathematics. A random grouping of students to form a SS will likely confound the results—a strong student in the grouping could have solved a difficult item and a weaker student in the grouping could not solve an easy item assigned. Thus, the SS would be invalid due to the misfit in responses. Although no study has been done to assess the correlation between mathematics ability and big idea thinking, Schoenfeld (2019) mentioned in his study that high performing individuals are able to see and use Big Ideas in problem solving. We thus make a reasonable assumption that students of similar ability may have the same level of Big Idea thinking. In this light, we propose to constitute an SS, with all four students having identical ability (ideally but impossible in practice), by rank ordering students based on their past semestral assessment marks as a proxy of their mathematical ability. Going down the list, every four students are grouped into an SS and given a new SS ID. For example, in a school of 320 students, the top four students will constitute SS01, the next four SS02 and the last four students in the ordered list will be SS80. Triangulation can be carried out with teachers to validate that the students grouped together are indeed of similar ability. To differentiate the students within each SS, a suffix is added, e.g., SS01a, SS01b, SS01c and SS01d for the four students that constitute SS01. This is done to facilitate the correct distribution of the items.

We envisage a final instrument for a Big Idea consisting of eight items (each with five parts). The eight items are split into eight testlets, T1 to T8 as shown in Table 1. Each testlet is only made up of two items and each student attempts only one of the testlets. Table 1 shows how the testlets are distributed to the students as well as to each SS. Since each testlet has only two complete items, it can be administered easily within a much shorter duration and will reduce cognitive fatigue.

Table 1

Matrix Distribution of Items to Two SS Comprising a Total of Eight Students

	T1	T2	T3	T4	T5	T6	T7	T8
I1	SS01a							SS02d
I2	SS01a	SS02a						
I3		SS02a	SS01b					
I4			SS01b	SS02b				
I5				SS02b	SS01c			
I6					SS01c	SS02c		
I7						SS02c	SS01d	
I8							SS01d	SS02d

As a result of the SS structure and distribution of testlets, each SS will have taken the entire set of items while each student only attempts two items. Thus, the duration required to complete the test is only 25% of the time required to complete all the eight items. The score collated will be for each SS instead of for every student in the school. This SS structure can be used not only in obtaining an aggregated score for assessing group performance on an instrument, but it can also be used for validating an instrument during its initial item creation stage.

Conclusion

While assessments are important to monitor learning, too many high stakes assessments will reduce available time for teaching, erode the joy of learning and cause a high level of worry and stress about exams and results. However, assessment remains crucial to monitor if learning has taken place and is an important feedback mechanism to improve teaching as well as learning. In place of high stakes assessments, an aggregated structure as proposed may gather sufficient information regarding learning without increasing student cognitive load nor take up too much precious curriculum time. This may be a worthwhile contribution towards the joy of learning.

One of the main issues that arise from the SS structure is the validity of the SS itself. How similar are the four students within each SS? With no prior research done on the relationship between math ability and Big Idea thinking, it is difficult to validate the structure we have proposed. At this juncture, we have piloted the items and the SS structure is due to be tested and analysed later. We intend to explore and analyse the performance of the SS using two different approaches.

The first approach is to study the misfit of SS scores using Rasch analysis. In the development of the instrument, the items would be calibrated and validated using Rasch model. Using the same Rasch model analysis, we will be able to do a fit analysis by looking at person (SS) misfit information, if any. In the event of any person misfit cases, we hope that the misfit is due to the individual students doing the two items erratically, and not caused by the different students within the SS, e.g., the misfit is due to SS01a getting items with higher difficulty correctly while SS01c answering items with lower difficulty incorrectly. The second approach is by comparing a super-student score against the scores of each of the four students forming the super-student structure based on plausible values created for each student. The technique to calculate the plausible values can be found in Von Davier et al. (2009). We will collect our data from July 2023 and report the results thereafter.

Acknowledgements

The study reported here has been funded by the Singapore Ministry of Education under the research project OER31/19BK which is managed by the National Institute of Education, Singapore.

References

Ackerman, P. L., & Kanfer, R. (2009). Test length and cognitive fatigue: An empirical examination of effects on performance and test-taker reactions. *Journal of Experimental Psychology. Applied, 15*(2), 163–181. https://doi.org/10.1037/a0015719.

Andrich, D. (1998). *Rasch models for measurement* (Vol. 68). Sage.

Jahangeer, M. J. et al. (in press). Exploring the dimensionality of big ideas in school mathematics. *Proceedings of the 45th annual conference of the Mathematics Education Research Group of Australasia*. Newcastle: MERGA.

Niemi, D., Vallone, J., & Vendlinski, T. (2006). *The power of big ideas in mathematics education: Development and pilot testing of POWERSOURCE assessments.* CSE Report 697, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Schoenfeld, A. H. (2019). Reframing teacher knowledge: A research and development agenda. *ZDM*, 1–18.