Assessing Mathematical Competence Through Challenging Tasks

Jane Hubbard
Monash University
Jane.Hubbard@monash.edu

James Russo
Monash University
James.Russo@monash.edu

Sharyn Livy
Monash University
Sharyn.Livy@monash.edu

Making accurate judgements and interpretations about student growth and progress in mathematics can be problematic when using open-ended assessments. This study reports on the development of a class-based assessment instrument and marking key designed to assess Year 2 students' mathematics competence to reflect their learning of mathematics through a challenging tasks approach. A qualitative coding process was undertaken to analyse the written responses of 59 Year 2 students resulting in the development of a 7-point marking key to identify levels of progress. The marking key proved effective in supporting the interpretation of the written responses and identifying future learning pathways.

Introduction

Mathematics assessment is a central component within the teaching and learning cycle that can provide insight into student progress and inform future instructional decisions (National Council of Teachers of Mathematics [NCTM], 2014). For teachers to be effective in meeting this intention, it is recommended that assessment practices more accurately reflect students' learning experiences (Wiliam, 2007). An ongoing challenge in mathematics education is finding suitable methods of assessment that authentically reflect student learning and align to the reform orientations evident in contemporary classrooms. Despite a shift in mathematics education that encourages breadth and depth of curriculum knowledge, Dong et al. (2021) identified that many current assessment practices used for young students continue to rely on formal, narrow, skill-based tests. Clarke (2011) described these traditional assessment practices as misrepresenting both mathematics as a discipline as well as the actual student learning that has been experienced. Rather, class-based assessments based on open-ended or rich tasks enable teachers to ascertain what and how students are learning (Yeo, 2011). In a review of mathematics assessment throughout Australasian literature, Serow et al. (2016) identified that further research reporting class-based assessment is required across all levels of schooling. In particular, they recommended that such assessment should strive to provide students with adequate opportunities to comprehensively demonstrate their knowledge and understanding.

Context

The study reported in this paper was conducted as part of a larger project, Exploring Mathematical Sequences of Connected, Cumulative and Challenging Tasks (EMC^3), led by Professor Peter Sullivan and colleagues. Building on previous research of challenging tasks (see Sullivan et al., 2015), EMC^3 aimed to investigate how the sustained use of challenging tasks and the associated pedagogies supports Foundation to Year 2 students' mathematics development (Sullivan et al., 2020). Most of the research reporting on the benefits of learning mathematics through challenging tasks is limited to teacher professional development (e.g., Ingram et al., 2020; Sullivan et al., 2015) or student achievement in the middle to high school years (e.g., Sullivan et al., 2016), with some exceptions (e.g., Russo & Hopkins, 2017). Further investigations about assessing students and interpreting their mathematics competence when learning through challenging tasks in the Early Years will contribute to the literature.

2022. N. Fitzallen, C. Murphy, V. Hatisaru, & N. Maher (Eds.), *Mathematical confluences and journeys* (Proceedings of the 44th Annual Conference of the Mathematics Education Research Group of Australasia, July 3–7), pp. 282–289. Launceston: MERGA.

Literature Review

The interpretation of what and how mathematics should be taught influences teacher assessment choices and therefore, perceptions of student competence (Nortvedt & Buchholtz, 2018). In Australia, the current review of the national curriculum has highlighted the ongoing debate on what constitutes effective approaches for teaching mathematics for students including the Early Years of schooling (Foundation to Year 2). One perspective is that students learn mathematics best when experiencing direct instruction and having time to practise skills and facts before they encounter problem-solving situations (Kirschner et al., 2006). Such approaches are aligned with traditional notions of mathematics achievement and progress is often measured by assessments where students are required to reproduce facts and procedural knowledge (Clarke, 2011). Others believe young students can attain meaningful conceptual knowledge and foundational skills through pedagogies that centralise problem-solving approaches (Chan & Clarke, 2017). Educators aligned with inquiry approaches believe that the opportunities to engage students with problem-solving can strengthen mathematical understanding and support them to adapt their knowledge to a range of contexts (Schoenfeld, 2007). To ensure this approach rigorously improves student outcomes, it is recommended that teachers take an active instructional approach and scaffold student learning with the use of guided questions and rich classroom discussions (Chan & Clarke, 2017).

Regardless of the mathematics teaching orientation with which teachers choose to align, Wiliam (2007) recommends that for assessment practices to be deemed effective and helpful for informing future instruction, instruments and processes should accurately reflect the learning experiences of students. One of the criticisms of inquiry approaches in the Early Years is the lack of substantial evidence reporting its effectiveness for developing sufficient mathematical competence. One reason why student competence may not be readily apparent could be due to the mismatch in pedagogies and assessment practices used in schools (NCTM, 2014). The intention of this study is to investigate assessment practices that can effectively identify student progress while also aligning with student experiences of learning mathematics through a challenging tasks approach.

Developing Mathematics Competence Using Challenging Tasks

Teaching mathematics through problem-solving is one way students' experience of school mathematics can accurately reflect comprehensive notions of mathematics competence. Schoenfeld (2007) believes that through the act of problem-solving, students are afforded opportunities to demonstrate flexible and resourceful thinking; use efficient and productive strategies; and develop persistence and resilience. The thinking and reflection that is required to successfully solve problems also supports students in developing conceptual networks that help them to make sense of basic facts (Baroody et al., 2009).

Problem-solving is considered effective when students are encouraged to engage with mathematics tasks that are considered cognitively demanding (Smith & Stein, 2018). A challenging tasks approach (Sullivan et al., 2015) is one example of how students can experience cognitively demanding mathematics in the pursuit of developing mathematics competence. Challenging tasks, often characterised as being open-ended, encourage students to solve non-routine problems by eliciting their prior knowledge and exploring multiple solutions (Sullivan et al., 2020). Student learning is supported throughout the experience with consistent pedagogies and structures that include clear lesson foci; the use of probing questions; and opportunities to develop reasoning through class discussions. To account for the various levels of student readiness and to increase accessibility for all students, this approach also advocates differentiating the main task through the use of enabling and extending prompts (Sullivan et al., 2020). Pertinent to challenging tasks is accepting that there are multiple interpretations of success with the belief that students demonstrate competence when they can apply and transfer their mathematical knowledge across various contexts (Sullivan et al., 2020). However, with so many manifestations of mathematics competence, it can be difficult for teachers to make appropriate interpretations about student progress as well as determine accurate directions for future learning.

Assessing Mathematics Competence to Align with Challenging Tasks

The literature suggests that class-based assessments should allow for the interpretation of student growth and support teachers to adequately prepare for future learning experiences (Clarke, 2011; Wiliam, 2007). This has seen the development of alternative assessment processes such as rich assessment tasks where student growth is evaluated through marking keys or scoring rubrics (e.g., Downton et al., 2006). In a recent study conducted in the Netherlands, assessment techniques were evaluated to determine their effectiveness (Veldhuis & van den Heuvel-Panhuizen, 2020). The findings indicated that alternative class-based assessments can be productive when teachers possess sufficient levels of pedagogical content knowledge to support the interpretation of results. This reiterates the necessity for teachers to have a clear focus or purpose when administering assessment in the first instance, as well as being able to use assessment data to support improvement of student learning (Wiliam, 2007).

Making judgements about the divergent outcomes that result from open-ended, class-based assessments appears to cause difficulty for teachers (Yeo, 2011). To support teachers in the interpretation of work samples, Tomlinson et al. (2015) suggested that grouping responses according to patterns may be helpful. Such processes are regarded as response coding (Clarke, 2011) or comparative judgement (Jones et al., 2015). In essence, grouping student work according to similar responses or a marked point of difference allows for the scaffolding and then categorising of distinct levels of competence. These processes are most effective when teachers have a clear understanding of the task's intention as well as an ability to make inferences beyond a particular solution (Tomlinson et al., 2015). Similar analytical approaches to interpreting mathematical competence could be helpful when assessing student learning experiences through challenging tasks.

Research Question

The purpose of this paper is to report on the development of a class-based assessment instrument and marking key created to analyse and interpret the mathematics competence of Year 2 students when solving challenging tasks. The instruments were designed to answer the following research question:

> *How can Year 2 students' mathematical competence be analysed and interpreted using an assessment based on a challenging task?*

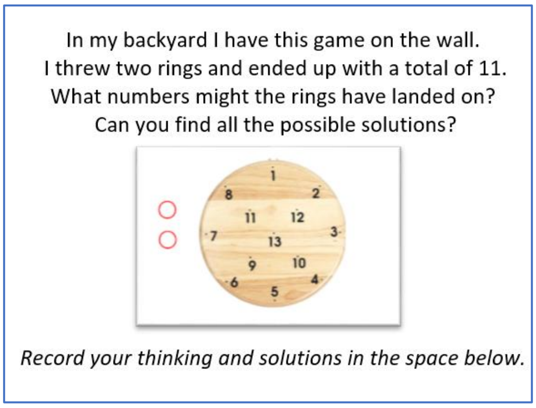
Methodology

The study reported in this paper is part of a larger PhD designed to investigate Year 2 student perceptions of challenging mathematical tasks. The mixed method study included a qualitative design for developing a marking key, whilst quantitative data were used to analyse and make interpretations about students' responses to the assessment task.

Participants, Data Collection Instrument and Administration

Participants were selected from one of the EMC[3] project schools. Three classes of Year 2 students (N = 59) from a Catholic primary school in metropolitan Melbourne participated in this assessment investigation. Students' prior experiences of learning mathematics was mostly through traditional instruction with limited exposure to open-ended tasks that encouraged student reasoning.

The assessment instrument consisted of three items designed to determine the extent to which students could competently find all possible solutions when partitioning two-digit numbers. Figure 1 provides an example of an item from the assessment instrument. The item had a similar mathematics focus to tasks from teacher resource materials for the EMC^3 project.



In my backyard I have this game on the wall.
I threw two rings and ended up with a total of 11.
What numbers might the rings have landed on?
Can you find all the possible solutions?

Record your thinking and solutions in the space below.

Figure 1. Assessment Item 1.

The item shown in Figure 1 can be described as non-routine having multiple possible solutions (for example, 6 & 5; 7 & 4; 8 & 3). Importantly, the context could be considered engaging and realistic for students in Year 2. The subsequent items were slightly more difficult. Item 2 required students to use two rings to find the possible ways to partition 15; and Item 3 asked students to partition 19 using three rings, with one ring already positioned on number seven. While this instrument focused on number partitioning, the open-ended components such as finding all solutions; recording systematically; and communicating mathematical thinking were all considered important elements to successfully demonstrate mathematics competence when solving challenging tasks.

The instrument was administered twice, using the same items each time. The first assessment occurred early in Term 1 and the second assessment occurred approximately nine months later. The specific items were not reviewed with the students between assessments. However, over the study students completed a range of other tasks from the project including suggestions from the Structure of Number Sequence to support their learning of the assessment concepts. It is important to note that during the timeframe of this study, school disruptions occurred in Melbourne as a result of COVID-19 restrictions.

Students were given 30 minutes to attempt to solve all three items. During this time, students were encouraged to read and attempt the items on their own. Students who experienced reading difficulties were able to have the task read to them although no additional guidance or prompts were provided about how to solve the task. Those who finished early were encouraged to review their responses. Such implementation is consistent with the introductory phase of the pedagogical approach used when implementing EMC^3 lessons.

Data Analysis

An initial marking key, created prior to any data collection, consisted of a general 4-point scale anticipating the range of possible responses. During the analysis it became apparent that these initial codes were insufficient in adequately identifying and categorising the range of responses. The coding of student responses was repeated using a combination of processes from the literature such as response coding (Clarke, 2011) and comparative judgement (Jones et al., 2015) to better distinguish and identify the different stages of progress evident within the data. Further details about the coding process are reported in the results

Results

The results are reported in two parts. The first section compares the initial marking key with the final marking key developed through the interpretation of student results. The second section reports both assessments scored according to the final marking key.

Developing the Marking Key

All written responses were read, sorted and coded according to an initial marking key which had a 4-point score range: 0 point (no response/ incorrect response); 1 point (one correct solution); 2 points (more than one correct solution) and 3 points (all correct solutions identified). When different examples of competence were identified throughout the analysis that was not specified within the initial marking key, the coding descriptions were modified and the responses were re-coded. This process was repeated until a final marking key was created that sufficiently represented the different levels of student competence within the data. Before a new code was included, the data were reviewed to determine if any other written responses also fitted in the alternative categories. Codes were changed only when there were at least three examples across all items that could verify the addition of a new category.

The final 7-point marking key (scores 0 to 6) in Table 1 shows the range in skill and knowledge development demonstrated by Year 2 students when participating in a written assessment reflective of a challenging task involving multiple solutions. The progression extends beyond content knowledge and accounts for broader mathematical skills such as comprehension of the context and pattern recognition. That all adjustments to the marking key were derived from the student data itself supports the validity of the categories within the marking key. Furthermore, when the marking key was presented to the EMC^3 project team ($N = 8$) there was consensus that the seven levels represented distinct developments in students' mathematical competence when solving open-ended, challenging tasks. It is also noteworthy that applying the revised marking key (with 7 levels) to the student response data generated a somewhat more reliable 3-item measure of student mathematical competence compared with applying the original marking key (with 4 levels). For example, the Cronbach alpha for the pre-assessment data applying the revised marking key was 0.70, whereas using the original marking key it was 0.62.

One of the more noticeable modifications to the marking key occurred for responses scored 0 and 1. For students who provided an incomplete and seemingly incorrect solution, some demonstrated a level of progress within their response that suggested that they had understood the context of the task and were accessing prior knowledge. For example, Student A who circled two digits on the ring board image that when added, made the target of 11 (e.g., 7 & 4) clearly demonstrated more comprehension than Student B who left it blank or wrote something irrelevant. Therefore, even though Student A made no attempt to formally record a solution (e.g., "7 + 4 = 11"; "7 and 4 is 11") their response was scored 1 point and the response by Student B was scored 0 points. The other prominent change in the marking key focussed on further delineating between responses that included all of the possible solutions. For this level of response, distinguishing between all correct solutions ordered systematically and those that were correct but recorded randomly, or that contained additional 'impossible solutions' (e.g., "12 – 1") became important.

Table 1
Final Marking Key

Score	Description	Elaborations
0	No attempt / completely incorrect attempt	
1	Some evidence of comprehending the task	Circling digits on the ring board that relate to a possible solution
2	One correct solution	
3	More than one correct solution	More than one solution but not all solutions
4	All solutions identified but additional incorrect/irrelevant solutions also included	Including subtraction: 12 – 1; 13 – 2 Including more partitions: 4 + 4 + 3
5	All solutions identified but unsystematic recording	Solutions randomly recorded No order or pattern identifiable
6	All solutions identified and correct, systematic recording and clear communication of solutions	Responses clearly demonstrating patterns, no errors evident

Cohort Data

Using the final marking key as our guide for all items, data were collated according to the frequency of different scores for both the pre and post assessment (Table 2).

Table 2
Year 2 Pre- and Post-assessment Responses (N = 59)

Item	/Score	0	1	2	3	4	5	6	Total
1	Pre	17	2	10	14	2	7	7	59
	Post	0	0	2	19	2	13	23	59
2	Pre	24	2	16	10	4	3	0	59
	Post	2	0	5	26	12	10	4	59
3	Pre	41	10	5	2	0	1	0	59
	Post	13	12	13	17	0	1	3	59

The frequency of scores obtained by students demonstrates various levels of student progress across the three items. Comparisons between pre and post assessments can be made at a particular score level, allowing for judgements to be made about specific progress of mathematics competence. For example, it is evident from the pre assessment that 17 students were unable to answer item 1 in any capacity. In the post assessment, all students were able to provide at least one solution to the task, demonstrating an improvement in students' ability to independently comprehend the task as well as the appropriate mathematics required.

Combining scoring groups can also help make judgements about mathematical competence. For example, in the pre assessment for item 1, around half of students (n = 29; 49%) found only a single solution to the task, which was indicated by a score of 2 or less. In comparison, the post assessment data revealed that only a small minority of students (n = 2; 3%) recorded a score of 2 or less, with almost all students (n = 57; 97%) now identifying multiple solutions to the task indicated by a score of 3 or more. These results show not only did students develop the core skills associated with scores 0-3 but also many were also able to demonstrate higher thinking skills with scores ranging from 4-6. Such insights can support teachers in the interpretation of student progress and necessary directions for future learning. In this instance, the emphasis on future learning might shift from focussing on the need to find more than one solution (following the pre assessment) to demonstrating knowledge that all solutions have been found (following the post assessment).

Discussion and Conclusion

The assessment instruments reported in this study were reflective of students' learning experiences, which according to Wiliam (2007), is necessary for establishing whether learning outcomes are met as intended. The inclusion of three similar assessment items within the instrument simulated the challenging tasks approach through which students had been learning mathematics throughout the year. Moreover, each item required students to demonstrate knowledge and skills beyond a single numerical response. In this particular assessment, students were required to comprehend the context of the task; access prior knowledge to solve the task; demonstrate flexibility with their conceptual understanding by providing multiple solutions; and communicate their solutions clearly. Demonstration of such knowledge and skills reinforce and reflect the mathematics competence intended by the EMC^3 project (Sullivan et al., 2020).

Developing and refining the marking key from the raw data enabled the identification of nuanced differences about how student mathematics competence progresses when learning through challenging tasks. The identification of different scoring codes within the marking key provided a clear scaffold through which interpretations about student learning with challenging tasks could be made. Furthermore, the results demonstrate the versatility in which a marking key can be applied to support teachers in making inferences about mathematics competence on many levels. These processes may alleviate the confusion and uncertainty that comes with interpreting open-ended assessments (Tomlinson et al., 2015) and support teachers in strengthening their knowledge for teaching mathematics (Veldhuis & van den Heuvel-Panhuizen, 2020; Yeo, 2011).

This study aimed to investigate how assessment of student progress can be designed to align with their experiences of learning mathematics through a challenging tasks approach. The marking key derived from student responses captured the diverse range of knowledge and skills that Year 2 students demonstrated when they solve non-routine, cognitively demanding tasks. While limitations of this study include the small sample size across one school setting, the processes described here may provide a template to initiate further research into assessment of challenging tasks in the Early Years.

Acknowledgements. EMC^3 was funded by the Australian Research Council, Catholic Education Diocese of Parramatta and Melbourne Archdiocese Catholic Schools (LP 180100611).

References

Baroody, A. J., Bajwa, N. P., & Eiland, M. (2009). Why can't Johnny remember the basic facts? *Developmental Disabilities Education Reviews*, *15*, 69–79. https://doi.org/10.1002/ddrr.45

Chan, M. C. E., & Clarke, D. (2017). Structured affordances in the use of open-ended tasks to facilitate collaborative problem-solving. *ZDM Mathematics Education*, *49*, 951–963. http://doi.org/10.1007/s11858-017-0876-2

Clarke, D. J. (2011). Open-ended tasks and assessment: The nettle or the rose? In B. Kaur & K. Wong (Eds.), *Assessment in the mathematics classroom: Yearbook 2011 association of mathematics educators* (pp. 131–163). World Scientific. https://doi.org/10.1142/9789814360999_0007

Downton, A., Knight, R., Clarke, D., & Lewis, G. (2006). *Mathematics assessment for learning: Rich tasks & work samples*. Mathematics Teaching and Learning Centre.

Dong, Y., Clements, D. H., Day-Hess, C. A., Samara, J., & Dumas, D. (2021). Measuring early childhood mathematical cognition: Validating and equating two forms of the research-based early mathematics assessment. *Journal of Psychoeducational Assessment*, *39*(8), 983–998. https://doi.org/10.1177/07342829211037195

Ingram, N., Holmes, M., Linsell, C., Livy, S., McCormick, M., & Sullivan, P. (2020). Exploring an innovative approach to teaching mathematics through the use of challenging tasks: A New Zealand perspective. *Mathematics Education Research Journal, 32*(3), 497–522. https://doi.org/10.1007/s13394-019-00266-1

Jones, I., Swan, M., & Pollitt, A. (2015). Assessing mathematical problem-solving using comparative judgement. *International Journal of Science and Mathematics Education, 13*, 151–177. https://doi.org/10.1007/s10763-013-9497-6

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist 41*(2), 75–86.

National Council of Teachers of Mathematics. (2014). *Principles to actions: Ensuring mathematical success for all.* NCTM.

Nortvedt, G. A., & Buchholtz, N. (2018). Assessment in mathematics education: Responding to issues regarding methodology, policy, and equity. *ZDM Mathematics Education, 50*, 555–570. https://doi.org/10.1007/s11858-018-0963-z

Russo, J., & Hopkins, S. (2017). Student reflections on learning with challenging tasks: 'I think the worksheets were just for practice, and the challenges were for maths'. *Mathematics Education Research Journal, 29*(3), 283–311. https://doi.org/10.1007/s13394-017-0197-3

Schoenfeld, A. H. (2007). What is mathematical proficiency and how can it be assessed? In A. H. Schoenfeld (Ed.), *Assessing mathematical proficiency* (pp. 59–74). Cambridge University Press. https://doi.org/10.1017/CBO9780511755378

Serow, P., Callingham, R., & Tout, D. (2016). Assessment of mathematics learning: What are we doing? In K. Makar, S. Dole, J. Visnovska, M. Goos, A. Bennison & K. Fry (Eds.), *Research in mathematics education in Australasia, 2012–2015* (pp. 235–254). Springer. http://doi.org/10.1007/978-981-10-1419-2_12

Smith, M. S., & Stein, M. K. (2018). *5 practices for orchestrating productive mathematical discussions (2nd ed).* National Council of Teachers of Mathematics.

Sullivan, P., Askew, M., Cheeseman, J., Clarke, D., Mornane, A., Roche, A., & Walker, N. (2015). Supporting teachers in structuring mathematics lessons involving challenging tasks. *Journal of Mathematics Teacher Education, 18*(2), 123–140. https://doi.org/10.1007/s10857-014-9279-2

Sullivan, P., Bobis, J., Downton, A., Hughes, S., Livy, S., McCormick, M., & Russo, J. (2020). Ways that relentless consistency and task variation contribute to teacher and student mathematics learning. In A. Coles (Ed.), *For the learning of mathematics monograph 1. Proceedings of a symposium on learning in honour of Laurinda Brown* (pp. 32–37). FLM Publishing Association.

Sullivan, P., Borcek, C., Walker, N., & Rennie, M. (2016). Exploring a structure for mathematics lessons that initiate learning by activating cognition on challenging tasks. *Journal of Mathematical Behaviour, 41*, 159–170. http://dx.doi.org/10.1016/j.jmathb.2015.12.002

Tomlinson, C. A., Moon, T., & Imbeau, M. B. (2015). *Assessment and student success in a differentiated classroom.* ASCD. https://files.ascd.org/staticfiles/ascd/pdf/siteASCD/publications/assessment-and-di-whitepaper.pdf

Veldhuis, M., & van den Heuvel-Panhuizen, M. (2020). Supporting primary school teachers' classroom assessment in mathematics education: Effects on student achievement. *Mathematics Education Research Journal, 32*, 449–471. https://doi.org/10.1007/s13394-019-00270-5

Wiliam, D. (2007). Keeping learning on track: Classroom assessment and the regulation of learning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 1053–1098). Information Age Publishing.

Yeo, K. K. J. (2011). Implementing alternative assessment in the lower primary mathematics classroom. In B. Kaur & K. Wong (Eds.), *Assessment in the mathematics classroom: Yearbook 2011, Association of Mathematics Educators* (pp. 113–130). World Scientific. https://doi.org/10.1142/9789814360999_0006