Aligning Victorian Curriculum achievement standards with PAT Maths and Reading scales

Contents

Tables and figures

Introduction

The ACER Progressive Achievement Tests in Reading and Mathematics, respectively known as PAT Reading and PAT Maths, are a set of assessments that allow teachers to accurately and efficiently measure students' abilities, to diagnose gaps, strengths and weaknesses in student learning, and monitor student progress over time. In Victoria, more than 90% of schools have access to PAT Maths and PAT Reading, either via individual school licences or system-wide agreements. Over time, ACER has received requests from educators, particularly in Victoria, about how achievement measured on the PAT Maths and Reading scales might correspond with the Victorian Curriculum achievement standard levels in Mathematics and English (Reading and Viewing), respectively.

In 2023, ACER conducted a bookmark standard-setting workshop to assess whether it is possible to align PAT scale scores in these domains with the achievement standards for a subset of Victorian F−10 curriculum levels. The bookmark exercise was designed to determine where cut scores for each achievement standard level would lie. The exercise was restricted to Victorian F−10 achievement levels 1 to 6 for each of Mathematics and English.

Items from PAT Maths and PAT Reading assessments were used to construct ordered item booklets, where the panels of expert raters would locate each bookmark. Items were selected to cover the entire PAT scales for each learning area, and as much of the construct content as possible.

Standard setting process

Objectives and method

This study was designed to determine where each achievement standard, Levels 1–6, begins and ends on the PAT Reading and PAT Maths scales.

PAT assessments report scale scores for students that reflect student achievement (ability estimate) in the learning area. Each PAT test item also has a scale score that reflects the difficulty of the item (item difficulty estimate) relative to the scale. In PAT Reading and PAT Maths assessments, students and items are mapped onto the same Reading or Maths scale, respectively. The levels of knowledge and skills required to correctly respond to an item increase as students make progress on the scale.

The Victorian Curriculum achievement standards provide descriptions of substantively different levels of progress along a continuum of learning as defined by the Victorian Curriculum F–10. The achievement standards describe what students can do and understand when they have attained that level.

Because PAT items have well-established scale score difficulties, the bookmark standard-setting procedure was used to find cut scores on PAT scales for six achievement standard levels in each of the two domains.

Hambleton (1998) and Cizek (2006) have described a generic set of steps to establish standards, regardless of the approach or procedure adopted. This study was conducted in several stages according to an adapted version of Cizek's (2006) description, as follows:

1. **Participants:** Select a large and representative panel of expert raters.

2. **Preparation**: Define the procedure to be followed and prepare professional development materials for panel members.

3. **Target/borderline student:** Prepare descriptions of and discuss with panellists the referent candidate for each cut score.

4. **Training:** Provide professional development to the panellists on the method and their tasks.

5. **Collect data:** Compile judgements from the expert panel, summarise outcomes and provide feedback.

6. **Discuss:** Facilitate a discussion amongst the expert panel based on feedback from the first judgement exercise.

7. **Revision:** Provide experts the opportunity to review or re-confirm their judgements following discussion with their peers.

8. **Impact:** Provide opportunity to review and revise judgements, if necessary, in light of available data on the impact of their judgements.

9. **Evaluate:** Provide outcomes to participants and confirm they are satisfied and have confidence in the process.

10. **Documentation:** Document the process.

Participants

Expert panellists were sought from ACER's extensive connection with teachers and schools through the PAT assessment program, and through other ACER-run marking programs that require the input of qualified and experienced teachers and educators. Invitations were extended to PAT professional learning workshop participants from the preceding 12 months, as well as experienced ACER markers with teaching experience. Ideally, selected participants were educators with a good knowledge of primary level Mathematics or Reading and a sound understanding of what students can typically do at given year levels.

Participants, henceforth referred to as 'raters', were divided into four groups, two per domain, depending on their backgrounds and teaching experience. Relevant participant experience ranged from several years to decades of teaching experience across different school sectors. One group in each domain considered the cut scores for standards at achievement standard Levels 1 to 4 (the 'lower level group'), while the other group in each domain considered the cut scores for achievement standard Levels 3 to 6 (the 'upper level group'). Both groups independently considered the cut scores for Levels 3 and 4.

Training

In the first part of the training, participants were introduced to the key standard-setting concepts and to the PAT assessment scales, which indicate both the achievement of students who undertake the assessments and the difficulty of test items, or questions, with more difficult items at the upper end of the scale and easier items at the lower end. The Victorian Curriculum achievement standards provide descriptions of substantively different levels of achievement along a continuum of learning.

During a standard-setting session raters use an ordered item booklet, where items from the PAT assessments are provided in their order of difficulty, with the easiest items at the beginning of the list and most difficult items at the end of the list. Participants were introduced to the concept of the borderline, or target, student, a student sitting at the borderline between two achievement levels. It was also emphasised that the bookmark was to be placed at the borderline between the level considered (for example, Level 1) and the next (Level 2), meaning the score would represent the cut point where a student had mastered the level under consideration.

The raters were introduced to the ACER Signum application and provided with guidance on how to use it to record their bookmarks and track their thinking around bookmark placement by adding comments where necessary. Participants were encouraged to make use of the Angoff style yes/no judgements for each item in the set they were looking at and then use those individual item level decisions to make an on-balance judgement as to the placement of the bookmark.

The second part of the training focussed on how to determine where each participant believed a bookmark should be placed. This training involved taking participants through the conceptual and practical components of the exercise. After explaining that a target student is one that has just mastered all the concepts described at that achievement level, presenters unpacked the central question that participants needed to ask of each item in this process:

"If this item were administered in a test to the target student, how likely is it that they would be able to answer the item correctly?"

Finally, raters were instructed on the role of the achievement standards and curriculum content descriptions in the mapping exercise. It was impressed upon participants that, while achievement standards were the focus of the exercise, the curriculum content descriptions should be used to support the mapping, particularly where the achievement standards were broadly defined or required clarity. It was pointed out that there is a difference between achievement standards and curriculum content descriptions that participants should keep in mind during the exercise, namely that achievement standards describe expected development by the end of the year and curriculum content descriptions refer to skills and concepts for development throughout the year.

Data collection and discussion

Data collection took place over two rounds. In the first round, raters used ACER Signum to assess each item in the ordered item booklet against the skills and knowledge detailed in achievement standards at each level they were assigned. ACER staff provided guidance regarding use of ACER Signum, any questions of theory or uncertainty about the bookmarking process, and any technical difficulties. For each bookmark level, the raters considered whether the borderline, or target, student at that level would be more likely than not to respond correctly to the item under consideration. Raters were instructed that when the pattern of responses changed from mostly yes to mostly no for each item, that was the region in which they should place their bookmark. Generally speaking, this would be the first item where the pattern changed, but sometimes where that pattern was inconsistent, the best place falls within a series of items where the pattern begins to change. In this sense, raters were required to make an on-balance judgement as to where the bookmark would be set.

The data from this first round were collated and the initial bookmark placements analysed for presentation prior to the second round of judgements, undertaken at an in-person workshop designed to facilitate discussion amongst each group on the placement of the bookmarks and their assessment around the skills and knowledge required to respond to each item with reference to the achievement standard level that was being considered.

Impact and evaluation

As part of the process of evaluating the position of the bookmarks, the panels of raters were provided with some evidence of the impact of the location of their bookmarks when considering the typical PAT achievement of students at relevant stages of their schooling. The PAT Maths and PAT Reading Australian norms were used to show the distribution of PAT achievement relative to the bookmark standards that had been set during the workshop. The Australian norms were updated in 2022 using results from PAT Maths and PAT Reading tests completed by students across Australia in 2019. The norms provide a reference for the distribution of achievement at each year level nationally.
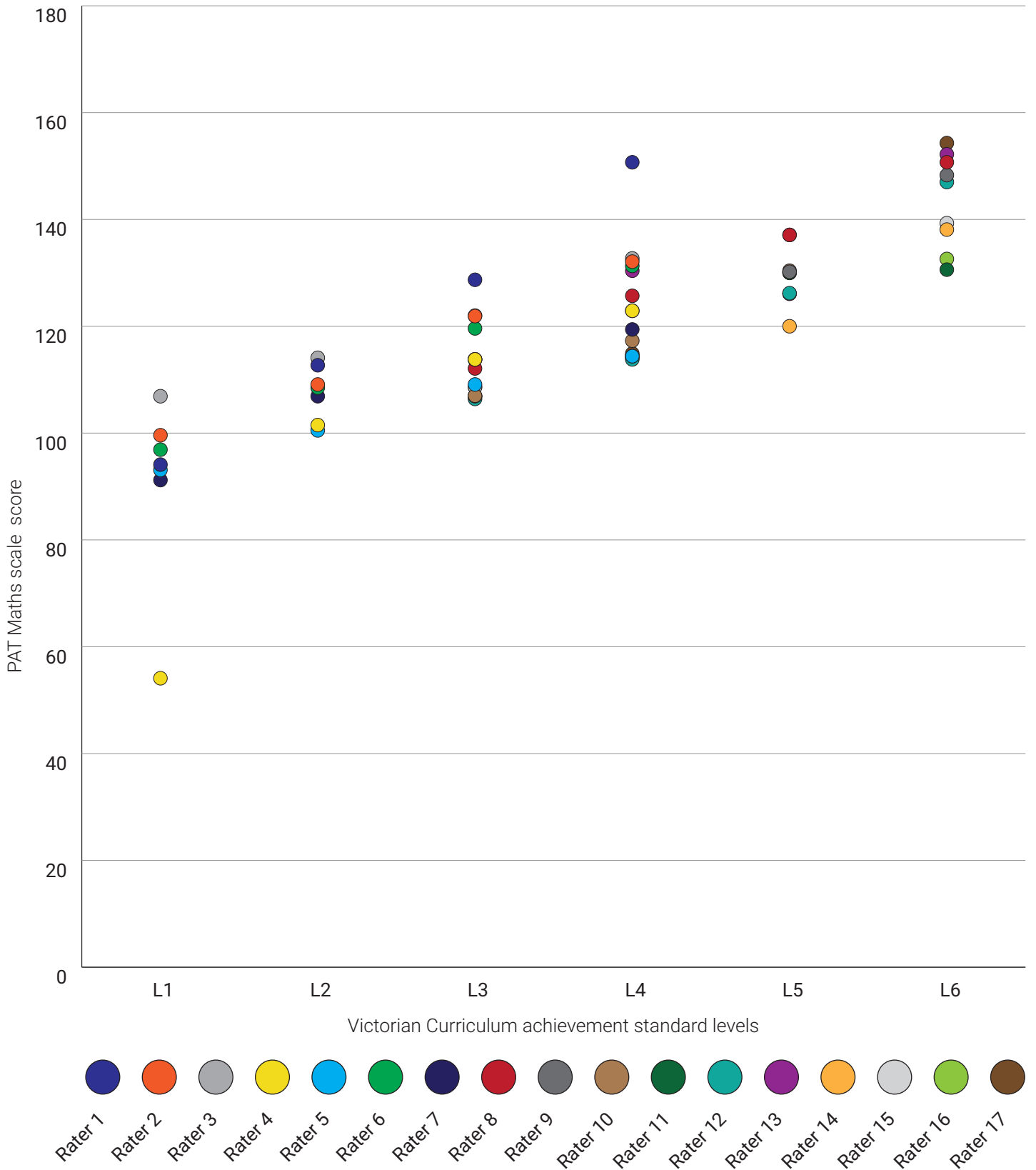
Mathematics

Results

Results in Maths are presented in two parts: the results after the independent judgement round (Round 1) and the results after the discussion round (Round 2). The cut scores on the PAT Maths scale and the standard errors are shown in Table 1. This table details the cut score and errors derived from the judgements of each group separately, and then those combined where the groups overlapped for Levels 3 and 4. Figure 1 shows the cut scores assigned by each rater in Round 1.

Table 1 Mathematics cut scores and standard errors after Round 1

| | Cut scores | | | | | |
| | Combined groups | | Lower level group | | Upper level group | |
Achievement standard	Mean	Std error	Mean	Std error	Mean	Std error
Level 1	90.8	6.4	90.8	6.4		
Level 2	107.6	1.9	107.6	1.9		
Level 3	112.9	1.6	118.4	2.5	109.1	0.9
Level 4	122.7	2.5	129.1	4.5	118.2	1.9
Level 5	129.3	1.6			129.3	1.6
Level 6	143.7	2.9			143.7	2.9

Figure 1 Mathematics bookmark placements after Round 1

Aligning Victorian Curriculum achievement standards with PAT Maths and Reading scales

At Level 1 there was one outlier judgement for the first cut point. This was the very first item in the set and was placed in error. This judgement was corrected in subsequent rounds. Another outlier judgement was observed from the lower level group for the Level 4 bookmark. This placement was significantly higher than the placement of others in both groups.

Raters in the group assigning bookmarks for Levels 3 to 6 started at approximately 9 scale score points lower for Level 3 than the group starting with the lower levels found, and their Level 4 cut score was almost 11 scale score points lower. Agreement between raters at these levels was closer for the upper level group (see Figure 2). The judgements were ultimately combined and the average of all raters in both groups taken to determine the cut score for Levels 3 and 4.

These results and some preliminary impact information were presented to the groups together. After this point, the groups were separated and invited to participate in a discussion round where they examined the placement of their bookmarks and were allowed to amend their judgements in light of the collegiate discussion. The results after those amendments are shown in Table 2.

Table 2 *Mathematics cut scores and standard errors after Round 2*

|	Cut scores					
	Combined groups		Lower level group		Upper level group	
Achievement standard	Mean	Std error	Mean	Std error	Mean	Std error
Level 1	97.5	0.5	97.5	0.5		
Level 2	107.6	0.3	107.6	0.3		
Level 3	112.9	1.2	117.3	1.8	109.9	0.5
Level 4	120.4	1.9	127.3	2.7	115.6	1.1
Level 5	127.4	1.2			127.4	1.2
Level 6	134.9	1.9			134.9	1.9

Figure 2 *Mathematics bookmark placements after Round 2*
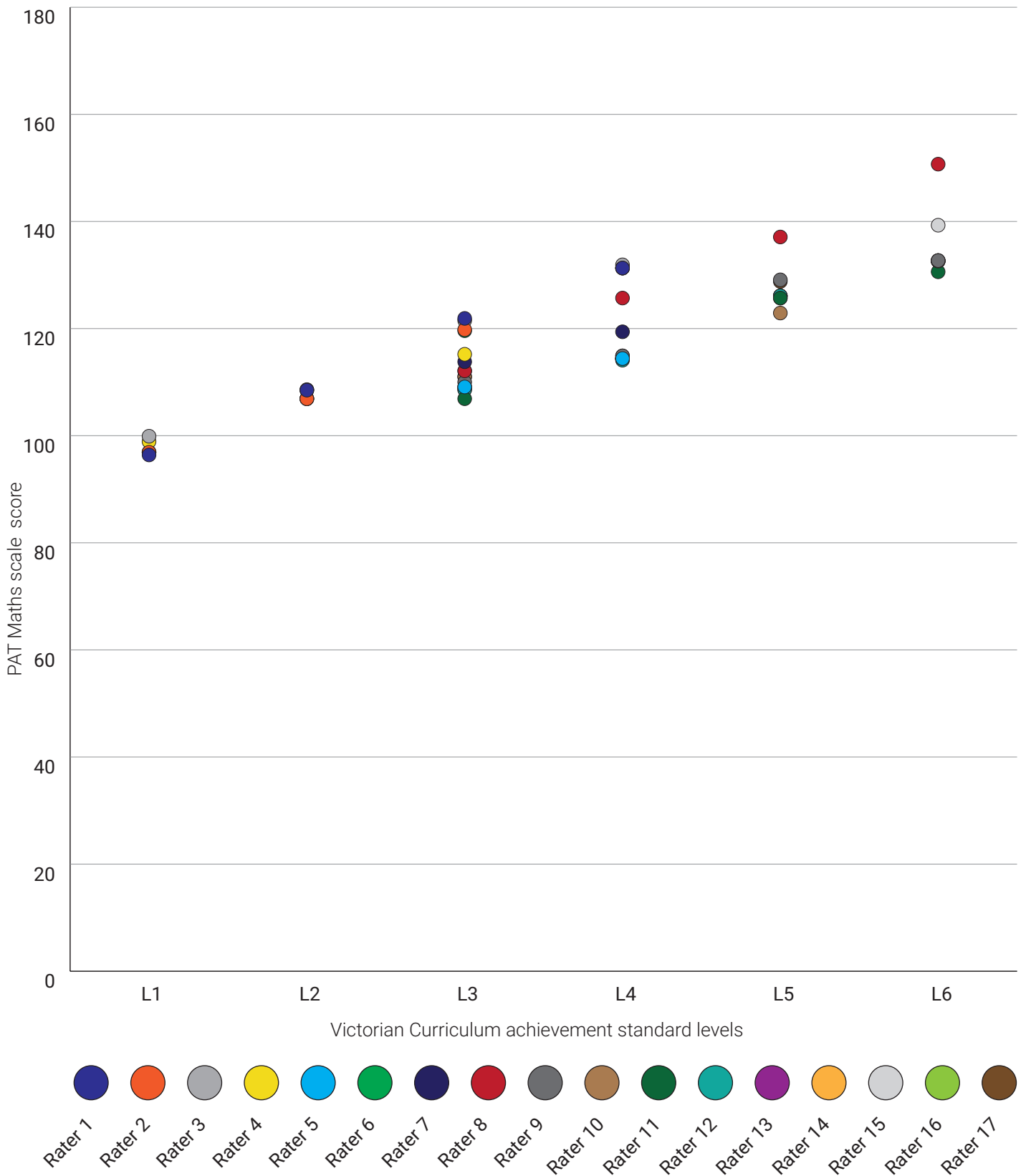
As expected, the opportunity to discuss placement of the bookmarks with their colleagues resulted in raters coming closer together in their placement. The differences between the two groups remained for the shared bookmarks at Level 3 and Level 4. At Level 3, the lower level group lowered their average bookmark placement, and the upper level group raised their average bookmark placement slightly. At Level 4, the lower level group once again lowered their average bookmark placement. The upper level group also lowered their placement, so a gap remained between the groups.

Using the mean bookmark location for each achievement standard level, the impact results shown in Figure 3 (page 9) were observed. Using the mean bookmark placement and incorporating a margin of two standard errors around that point, we can estimate a scale score range in which each achievement standard level sits. For example, we can estimate that a student who has achieved a PAT Maths scale score of at least 97.5 (± 1.0) is likely to be reaching Achievement Standard Level 1. Similarly, a student achieving a scale score of at least 134.9 (± 3.8) is likely to be reaching Achievement Standard Level 6.

Figure 3 *Mathematics achievement level bookmarks and PAT Maths Australian norms*

Aligning Victorian Curriculum achievement standards with PAT Maths and Reading scales
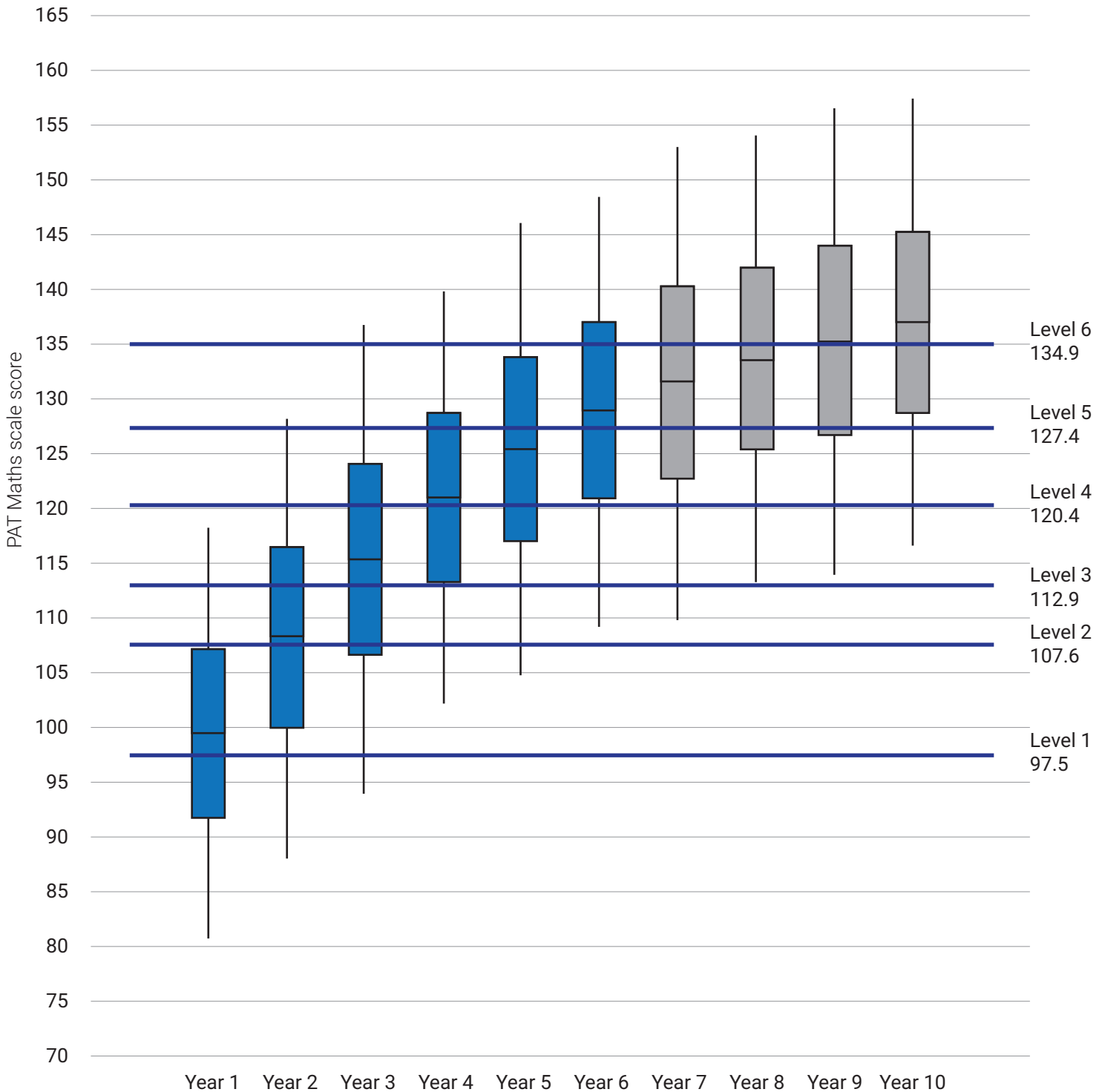
Discussion

For Mathematics, there was a reasonable consensus reached in Round 2, with standard errors reduced significantly compared to Round 1. For achievement standard Levels 1 to 5, the cut points hovered around the median scores for the corresponding cohorts in the PAT Maths Australian norms:

- The Level 1 cut score was set just below the median score for the year 1 PAT norm.
- The Level 2 cut score was set at the median score for the year 2 PAT norm.
- The Level 3 cut score was set just below the median score for year 3 PAT norm.
- The Level 4 cut score was set at the median score for the year 4 PAT norm.
- The Level 5 cut score was set just above the median score for the year 5 PAT norm.

This suggests that, for years 1 to 5, the achievement standard levels correspond quite well to the achievement levels of Australian students who complete the PAT tests.

The cut score for achievement standard Level 6 broke from this pattern. The cut score was set well above the median score for the year 6 norm cohort, and just below the upper quartile for year 6. The curriculum content and corresponding achievement standard for Victorian students in year 6 may be an ambitious target for a large proportion of year 6 students. At this year level, a lot of new and difficult concepts are introduced, or become significantly more complex in terms of the conceptual connections and interrelationships that need to be made. For example, the notion of integers is first introduced and number lines combining negative numbers and fractions are encountered. Although in year 5, students were dealing only with unit fractions, adding and subtracting fractions with the same denominator and dealing with very simple decimals, by year 6 they are expected to do far more complex calculations with fractions and decimals and also connect them to percentages and percentage calculations, such as sale discounts. The complexity of the Mathematics required, and the conceptual connections that need to be made, is far more advanced than in year 5 and represents a larger conceptual leap compared to the increase in curriculum expectation between one year level and another for the lower achievement standard levels. So, although the year 6 cut score seems like a significant outlier, a qualitative analysis of the curriculum expectations compared to previous year levels does help to explain this result.

Conclusion

For Mathematics, the results of this standard setting exercise suggest a fairly good alignment between Victorian Curriculum v1.0 achievement standards and the PAT norms for the corresponding year levels, for all levels apart from year 6, where the achievement standard for Victorian Curriculum v1.0 seemed a high relative to the Australian year 6 PAT norm. With the release of Victorian Curriculum v2.0, it is worth considering how the cut scores may shift if this exercise were to be repeated using the new standards. A qualitative look at the changes in content suggests that for years 1, 3 and 5, the cut score would likely be at a higher point on the PAT scale, while for years 2 and 6, it would likely be at a similar level. The effect of changes to the Victorian Curriculum – Mathematics is discussed in more depth in the Appendix.

Reading

Results

The Reading team reached a reasonable consensus in Round 2, and the resultant achievement standard level cut scores are fairly well distributed along the PAT scale. But there are some issues with the alignment between the achievement standard levels, which describe expected end of year development, and the median PAT Reading scale scores for each year level, according to the Australian norms. These findings raise some interesting questions about how well curriculum standards and descriptions can be aligned with an empirical scale when it comes to Reading.

Discussion

The main difficulty in this standard-setting exercise concerned the broadness of the achievement standards and curriculum descriptions and the lack of guidance they provide regarding how Reading skills progress. For example, achievement standard Level 2 for Reading and Viewing states that students "identify literal and implied meaning, main ideas and supporting detail." (Vic. Curriculum 2022) In PAT Reading, test items provide a distinction between simple and more complex interpretations. The PAT Reading achievement band description (90–99) for the strand 'Interpret Implied' indicates that students at this level "use their everyday knowledge to infer simple explanations for familiar events and predict likely outcomes." The progression of this skill is evident from the description of PAT Reading achievement band 100–109, at which students "infer main ideas when clues are scattered and there is some competing information."

The Victorian Curriculum achievement standard Level 3 for Reading and Viewing does not clearly build on the Level 2 description. This standard refers again to students' ability to "identify literal and implied meaning", but in this case by "connecting ideas in different parts of a text." (Vic. Curriculum 2022). In the PAT Reading achievement band descriptions, distinctions are made between simple connections, such as those "across adjacent sentences" (90–99) and more challenging connections "between events that are separated by one or two sentences" (100–109). These differences in purpose and approach between the curriculum and PAT assessments meant that raters were required to make significant inferences about the standards and descriptions when positioning them along the PAT scale.

Conclusion

Given the issues raised in the discussion, it is not appropriate to release the alignment of the Victorian Curriculum achievement standards to the PAT Reading scale at this stage. Questions remain about how we can align Reading assessments with the curriculum in a way that provides meaningful information to teachers about where students are in their learning and how they can be supported to progress. These questions are worthy of further scrutiny, as they concern how teachers can plan their lessons and collect evidence of student achievement in a way that best supports student development while still meeting the requirements of the curriculum.

References

Adams, R. J., Vista, A., Scoular, C., Awwal, N., Griffin, P., & Care, E. (2015). Automatic coding procedures for collaborative problem solving. In P. Griffin, & E. Care. Assessment and teaching of 21st century skills: Methods and approach. (pp. 115–132). Springer.

Cizek, G. J. (2006). Standard setting. *Handbook of test development,* 225–258.

Hambleton, R. K. (1998). *Handbook for the development of performance standards: Meeting the requirements of Title I.* Council of Chief State School Officers.

Appendix

Victorian Curriculum changes for Mathematics and their likely effects

After this standard-setting exercise was completed, the Victorian Curriculum – Mathematics v2.0 was released. A qualitative comparison of the changes to the achievement standards and implications for the findings for Mathematics is discussed here.

For achievement standard Level 1, more complex skills and reasoning processes have been added in v2.0. Namely, in the Number strand, counting elements in sets and representing cardinal numbers has increased from 20 to 120. There is now an explicit reference to partitioning – a strategy for counting and representing value – is also present, as well as 'skip counting'. Problem solving, formally via addition and subtraction (for numbers to 20), and informally (sharing and grouping) via multiplication and division is also included. Explicit reference to "Mathematical modelling to solve practical problems" is also new. In the Measurement strand, an emphasis on "communicating reasoning" and using "informal units" has now been added.

The effect of these changes is that the achievement standard Level 1 has been raised in comparison to v1.0 of the Victorian Curriculum, and so, if this exercise were to be repeated, the PAT Maths scale cut score for Level 1 would likely shift so that it is above the median scale score for the year 1 PAT norm.

For the achievement standard Level 2, some content elements that appear in v1.0 have been dropped and others have been added. Changes include the following:

- For the Number strand, an emphasis on how 'partitioning' can assist an understanding of place value, as well as in 'calculations'. An explicit reference to "Mathematical modelling to solve practical problems" occurs again, in both additive and multiplicative applications (as opposed to just additive). The following statement of proficiency is also new: "They recall and demonstrate proficiency with addition and subtraction facts within 20 and multiplication facts for twos."
- For the Measurement strand, the need to distinguish weeks, months and seasons from a calendar has been dropped; only days between events is mentioned.
- For the Space strand, the reference to "one-step transformations" of shapes has been replaced by "quarter, half, three-quarter and full measures of turn in everyday situations" (which may be linked with the analogue clock). Following directions and pathways also now appears. The reference to 3D objects has been dropped.

The cumulative effect of the additions and removals of content in the achievement standard Level 2 is harder to gauge than for the achievement standard Level 1. It is likely that the cut score for Level 2 would remain steady or only slightly increase if this exercise were to be repeated using the Victorian Curriculum 2.0.

For Level 3, much more specific detail has been added to the standard. Changes to the achievement standard are summarised below:

- Explicit reference to "Mathematical modelling to solve practical problems" occurs again, as in previous levels.
- Recalling multiplication facts has been reduced from "single-digit" to specifically 2s, 3s, 4s, 5s and 10s.
- For money transactions, "counting out change" has been replaced by determining the reasonableness of financial calculations and estimates.
- Specifically halves, thirds, quarters, fifths, eighths (and their multiples) have been dropped for the more generic "unit fractions and their multiples".
- Using the properties of odd and even numbers has been added in v2.0, as opposed to merely identifying them in v1.0.
- Exploring number patterns involving addition, subtraction and multiplication has been replaced by the more generic "exploring simple patterns". There is a new emphasis on creating patterns via an "algorithm".
- There is an extra requirement that angles, understood as a measure of turn, are compared to right angles.
- Making models of 3D objects is replaced by "compare and classify" objects (dropped reference to '3D').

- Statistical investigations are now to include numerical data (as well as categorical).
- The reference to the concept of symmetry has been removed.
- Explicit reference to finding unknown values in number sentences has been added.

The cumulative effect of this changes is also likely to shift the cut score up compared to v1.0.

The achievement standard Level 4 for v2.0 has the following changes compared to v1.0:
- Recalling 10 x 10 number facts extended to calculations where all four operations are applied for greatest efficiency.
- References to 'purchasing problems' and 'finding unknowns in number sentences' now cast as "use Mathematical modelling to solve financial and other practical problems", with a requirement to "interpret the results of the situation".
- The emphasis on 'unpacking' number sequences and patterns has been removed. In its place is the creating of number patterns via algorithms. The explicit reference to understanding decimal place value and (implicitly) multiplying and dividing by powers of 10 is new. Also new is the phrase "numerical equations", but only with respect to addition and subtraction.
- Reference to symmetry is now elaborated as 'line and rotational'. There is a new requirement that (at least some of) the shapes and objects studied are (approximately) 'real world'.
- For statistics, there is now a need to informally discuss data distribution shapes and variation. This awareness of statistical variation flows into probability with chance experiments.
- Apart from statistics, the reference to explore concepts "with digital technology" has been dropped.

The cumulative effect of these changes is likely to shift the cut score up to above the median scale score for the year 4 PAT norm cohort, due to the greater emphasis on reasoning and strategic competence and the inclusion of more complex concepts such as the notion of variation and probability as well as more complex symmetry concepts.

The changes in the achievement standard Level 5 in v2.0 are broadly outlined below:
- The reference to using the four operations to solve "simple" problems is now elaborated explicitly in terms of the multiplication and division demands, with further mention of the other operators in the sentence on modelling.
- Applications to budgets and solving unknowns in number sentences is again couched in terms of using Mathematical modelling, also with an emphasis on interpreting the results.
- Linking common percentages to decimals and fractions is new. An explicit reference to "numerical equations" is also new, and has been made, this time with respect to multiplication/division. Continuing number patterns involving adding/subtracting fractions has been removed, but "designing and using" algorithms to identify multiples and factors has entered.
- Estimating and measuring angles has been removed.
- The type of data collected is espoused and extended to nominal-ordinal categorical distinctions, as well as discrete numerical. In terms of interpreting and comparing data displays, the mode and shape of the distribution is now explicitly referenced, as well as a new, specific reference to line graphs. "Constructing various displays" of data sets, or something similar, is missing from the wording in v2.0 but was present in v1.0.
- (Theoretical) probability as value between 0 and 1, based on equally likely outcomes, is now just estimated likelihood (empirically quantified). This may mean that situations not involving equally likely outcomes could also be explored.

There is a significant conceptual shift in the achievement standard Level 5 for v2.0 compared to v1.0, including some content (the link between fractions, decimals and percentages) shifting from Level 6 to Level 5. As a result, if a similar standard-setting exercise were conducted with v2.0, a lower percentage of the year 5 PAT norm cohort would be expected to achieve the standard compared to the v1.0 standard. The cut score would likely significantly increase compared to v1.0.

For the achievement standard Level 6, changes include the following:

- The distinctly different treatment of applying the four operators to whole numbers and decimals is subsumed under a single statement regarding just decimals. Moreover, this statement includes representing measurements in the metric system.

- The emphasis on understanding fractions, decimals and percentages as equivalent forms has been replaced by creating any one of those number forms in applications. Moreover, problem solving with fractions and percentages, particularly discounts, has been replaced by the standard wording "Mathematical modelling to solve financial and other practical problems" using percentages and rational numbers. Here, it is also with an emphasis on justifying choices and estimating approximate solutions.

- Number sentences with order of operations, including brackets, is dropped, but is picked up by finding unknowns via "numerical equations" involving combinations of operations.

- The reference to multiplying and dividing by powers of 10, in the context of decimals, is gone (it has moved to v2.0 Level 5). Unlike Level 5, there was no reference to number patterns in Level 6 v1.0. Nevertheless, there is now the familiar reference to "designing and use algorithms to generate sets of numbers".

- Solving problems involving time and area has been elaborated, but reference to solving problems involving length has been reduced to just unit conversion. Making comparisons between volume and capacity has been reduced to just unit conversions for capacity. Explicit reference to volume has been removed.

- Investigating transformations in the plane now explicitly refers to tessellating patterns using combinations of transformations.

- The emphasis on "constructing" prisms is now on identifying parallel cross-sections of prisms.

- Reference to pyramids has been dropped.

- In statistical investigations, data types now explicitly include continuous numerical data, as well as the other data types, and reference is made to comparing distributions.

- Representing simple probabilities as a ratio has now been dropped.

At Level 6, some difficult concepts have been removed and others have been added, so the cumulative effects of the changes to the achievement standard Level 6 are not likely to significantly shift the cut score at this level compared to that determined for v1.0, which was already quite high (above the median scale score for the year 6 PAT norm cohort). If there is any effect, it may only be a slight increase.