2008, Vol. 20, No. 1, 82–107 *Mathematics Education Research Journal*

Do Mathematics Textbooks Cultivate Shallow Teaching? Applying the TIMSS Video Study Criteria to Australian Eighth-grade Mathematics Textbooks

Jill Vincent and Kaye Stacey
University of Melbourne

Australian eighth-grade mathematics lessons were shown by the 1999 TIMSS Video Study to use a high proportion of problems of low procedural complexity, with considerable repetition, and an absence of deductive reasoning. Using definitions from the Video Study, this study re-investigated this 'shallow teaching syndrome' by examining the problems on three topics in nine eighth-grade textbooks from four Australian states for procedural complexity, type of solving processes, degree of repetition, proportion of 'application' problems and proportion of problems requiring deductive reasoning. Overall, there was broad similarity between the characteristics of problems in the textbooks and in the Australian Video Study lessons. There were, however, considerable differences between textbooks and between topics within textbooks. In some books, including the best-selling textbooks in several states, the balance is too far towards repetitive problems of low procedural complexity.

The 1999 Third International Mathematics and Science Study (TIMSS) Video Study (Hiebert et al., 2003) described teaching practices in eighth-grade mathematics and science in the United States and in six countries where students performed well relative to the United States on the TIMSS 1995 assessments: Australia, the Czech Republic, Hong Kong SAR, Japan, the Netherlands, and Switzerland. Mathematics lessons were analysed from many viewpoints, such as the relative proportions of teacher and student speaking, how the lesson was organised in terms of whole class work and individual student work, and the types of problems that students solved. Almost 15000 mathematics problems in 638 eighth-grade mathematics lessons were analysed, with 82% of the problems focusing on number, geometry and algebra. Note that the Video Study uses the term 'problem' for any self-contained mathematical task undertaken by students. Many examples are given below.

The Video Study highlighted many characteristics of Australian lessons, some positive and some negative. On the positive side, for example, the Australian lessons had a higher proportion of real-life contexts than all countries except the Netherlands. Amongst the negative features, three quarters of the problems presented in the Australian lessons were repetitions of preceding similar problems in the lesson. This was the highest proportion of the seven countries. The Australian lessons also included the highest proportion of problems of low procedural complexity (Hiebert et al., p. 71) and virtually no

Australian lessons included proof or verification of results by logical reasoning by either the teacher or the students (p. 74).

Although these characteristics were not unique to Australia, and statistically Australia's position in these rankings was not significantly different from some of the other countries, taken together they are cause for concern to Australian educators. This cluster of features of Australian lessons — low procedural complexity of problems, high proportion of repetition, and absence of mathematical reasoning in classroom discourse — together constitute what we have termed the 'shallow teaching syndrome' (Stacey, 2003).

The findings raise the question of whether the syndrome is indeed a reality in Australian classrooms or simply an artifact of the definitions and procedures of the Video Study. There are many influences on what students learn in the mathematics classroom. However, with textbooks or worksheets reportedly used in at least 90 per cent of the Australian mathematics lessons (Hiebert et al., 2003, p. 114), the findings motivated us to consider the types of mathematical problems that Australian eighth-grade students would be exposed to if teachers were to closely follow a textbook (as will often be the case) and students undertook all of the work provided or, equivalently a statistically random sample of it. Our goal in the current study, then, was to compare 'textbook teaching' with the findings of the Video Study, to determine whether the general picture revealed by the 1999 Video Study would arise if all lessons followed textbooks exactly. In particular the following research questions are addressed:

1. Does the analysis of problems presented in textbooks in 2006 broadly align with the findings of the 1999 Video Study for Australia?
2. Can differences between textbooks be identified using the Video Study criteria?

Literature Review

Textbooks and other curriculum materials have been described by Ball and Cohen (1996) as "the stuff of lessons and units, of what teachers and students do" (p. 6). In a study of the use of mathematics textbooks in English, French and German classrooms, Pepin and Haggarty (2001) found that in some textbooks, exercises predominated, with few connections made between the concepts practised. In others, student exploration, questioning and autonomy were encouraged, and the posing of problems motivated the acquisition of new knowledge. Brändström (2005), who analysed three different Swedish seventh-grade mathematics textbooks, found that even when textbooks were written specifically to cater for students of different ability levels, only a small proportion of questions challenged students beyond the use of procedures.

Mayer, Sims, and Tajika (1995) compared the lesson on addition and subtraction of signed whole numbers in three seventh-grade Japanese mathematics textbooks with the corresponding lesson in four United States mathematics textbooks. All three Japanese books, but only one of the four United States books, emphasised multiple representations (words, symbols, and diagrams) in worked examples. The Japanese textbooks made strong links

between the three representations, supporting the findings of the TIMSS Video Study data, where 54% of problems in the Japanese lessons focused on 'making connections'. Making connections is one of the categories of solving processes to be discussed below.

When new mathematical concepts are introduced to students it might be expected that the teacher would augment the textbook explanations, demonstrate worked examples, and engage students in discussion where connections would be made with other concepts. Love and Pimm (1996) suggest that "the teacher normally acts as a mediator between the student and the text" providing an interpretation of the text that is "based not only on her constructions of the intention of the author, but on her accumulated experience of teaching" (p. 398). If this is the case, then the Video Study is expected to show a higher proportion of connections and deductive reasoning than the textbook analysis. This may not, however, be the case. Pehkonen (2004), for example, concludes that "teachers want the mathematics textbooks to concentrate on the basics, since they believe the basics constitute good and proper mathematics teaching" (p. 519).

Referring to the "back to basics" approach in US mathematics curricula of the 1970s, Schoenfeld (2004) reports that "not surprisingly, students showed little ability at problem solving — after all, curricula had not emphasized aspects of mathematics beyond mastery of core mathematical procedures. But performance on the 'basics' had not improved either" (p. 258). Schoenfeld criticises the extremes as untenable: "an exclusive focus on basics leaves students without the understandings that enable them to use mathematics effectively. A focus on 'process' without attention to skills deprives students of the tools they need for fluid, competent performance" (p. 280–281). Ideally, mathematics textbooks would present a balanced view of the importance of both skills and process.

The style of language chosen by a textbook writer may also influence what teachers and students perceive to be important in mathematics. In a study of the 'voice' of student texts associated with the Connected Mathematics Project (a US middle-school problem-centred curriculum), Herbel-Eisenmann (2007) found that although the authors often referred to 'questions', these were "actually imperatives, which were instructions to direct actions" (p. 354) (e.g., 'make', 'draw', 'use', 'look'). She asserts that the language used by textbook writers can influence students' conception of the nature of mathematics.

Stein and Lane (1996) refer to the commonly-held belief that students who lack adequate preparation in elementary mathematics cannot benefit from more challenging instruction. They suggest that this belief is responsible for an "even greater tendency for middle school instruction to focus on procedural skill" (p. 52). Reporting on the QUASAR (Quantitative Understanding: Amplifying Student Achievement and Reasoning) Project, Stein and Lane conclude from their study that the greatest student gains occurred where students were presented with tasks that encouraged non-algorithmic forms of thinking. Gains were relatively small, though, when tasks were "procedurally based and able to be solved with a single, easily accessible strategy, single representations, and little or no mathematical communication" (p. 74).

Stein and Lane (1996) note that there is a tendency for teachers "to proceduralize tasks due to time constraints" and to "perform the most demanding parts of tasks for students" (p. 60). This observation supports the findings of the TIMSS Video Study. When problems classified on the basis of their implied solution processes ('using a procedure', 'stating a concept' or 'making connections') were followed through to their public solution, only 8% of 'making connections' problems in the Australian lessons were actually solved in this way. The remaining problems were reduced to the use of a procedure or the stating of a concept (Hiebert et al., 2003, p. 104). This contrasted sharply with the findings for Czech Republic, Hong Kong and Japan where approximately 50% of 'making connections' problems were observed to be publicly solved by explicitly making connections.

In summary, the literature shows that many teachers rely on textbooks for instructional materials, which they may or may not supplement to make connections and emphasise mathematics beyond basic skills. The literature gives examples of textbooks that provide problems which extend students considerably beyond the routine use of procedures although textbooks generally do not. The literature also underlines the importance of providing problems that go beyond routine use of procedures to promote deep student learning.

Method

Selecting the Textbooks and Problems

The selected sample included the 2006 best-selling eighth-grade textbooks (textbooks A, B, C, and D) in four Australian states. Each was a clear market-leader in its state. The best-selling textbooks were selected because this gave us the best 'one book' picture of the problems that might be presented to Australian students. The same topics were also analysed in an additional sample of five different textbooks used in the authors' home state (textbooks E, F, G, H, and I). This state was chosen because of accessibility of the textbooks. Because the results are limited to only three topics, and these are not necessarily representative, the textbooks are not named in this paper.

Three topics from different curriculum strands that were common to all the textbooks were selected: addition and subtraction of fractions, solving linear equations by the 'do the same to both sides' method (not guess and check or graphical solving), and plane geometry concerning triangles and quadrilaterals. These three topics will be abbreviated hereafter as 'fractions', 'equations' and 'geometry'. All 3687 problems, including worked examples, from the three mathematical topics were analysed. The three selected topics were common to all states at this level and are also representative of the three most prevalent topic areas in the Video Study — number, algebra and geometry. The problems were drawn from the part of the textbook dedicated to that topic. We did not search the rest of the books to find problems that used knowledge from these topics.

Choosing topics that were comparable across textbooks for the different states was complicated by their slightly different curriculum emphases due to

local content variation, and by whether eighth grade was the first (textbooks A and B) or the second year of secondary schooling (textbooks C to I).

Numbers of problems. The number of problems in the textbooks varied considerably between textbooks and between topics. Textbooks A and B, from states where eighth grade was the first year of secondary schooling, included an extensive treatment of fractions (129 and 119 problems respectively). In states where eighth grade was the second year of secondary schooling, textbooks E, H, and I provided substantial revision, but textbooks C, D, and F each included only a small problem set (12 problems in textbook D) and G had no fractions section (see Appendix 1). The results reported below are therefore based on only two topics for textbook G. For equations, the number of problems ranged from 94 (textbook A) to 360 (textbook D) (see Appendix 2). For geometry, textbook C included only 21 problems, in contrast to 299 problems in textbook I (see Appendix 3), which seems more than any student would do. In nearly all the cases, the number of problems was large enough for the compared proportions to be not unduly affected by the presence or absence of a few unusual problems or unclear classifications.

Classifying the Problems

Each of our selected textbook problems was classified according to five of the Video Study criteria: procedural complexity, type of solving processes, degree of repetition, proportion of 'application' problems and proportion of problems requiring deductive reasoning. Although the Video Study also classified aspects of lesson delivery, the selected variables were applicable to problem statements, and so could be used on textbook problems. The five criteria and associated methods of classification are now described.

Procedural complexity. Hiebert et al. (2003, p. 70) note that complexity of the mathematics presented in lessons is an important feature, but that it is difficult to define and code reliably. The cognitive complexity of a problem depends, for example, on the experience and capability of the student, and on whether the problem requires the student to make connections across different mathematical concepts. In order to define the complexity of problems independently of the student, the Video Study focused on procedural complexity. Problems were classified as being of low, moderate, or high procedural complexity according to the number of steps in a common solution method and whether or not the problem involved the solution of sub-problems. A sub-problem is a task embedded in a larger problem that could itself be coded as a problem (p. 71). Table 1 shows the Video Study definitions for the three levels of procedural complexity and they are used here.

Table 2 shows our classification according to the Video Study definitions for procedural complexity of three problems from the Coding Manual (LessonLab, 2003a). It should be kept in mind that different solution strategies for particular problems may be possible and these may involve different numbers of steps. We do not know if the steps we identified coincided precisely with those identified by the Video Study, as their steps were not made explicit in the examples. However, our classification of low, moderate, or high procedural complexity for

the three problems in Table 2 coincided with those of the Coding Manual. In the main coding, such variations in the solution strategies only rarely changed the classification of procedural complexity.

Table 1
Defining Levels of Procedural Complexity of Problems (Hiebert et al., 2003, p. 71)

Procedural complexity	Definition
Low	Solving the problem, using conventional procedures, requires four or fewer decisions by the students (decisions could be considered small steps).
Moderate	Solving the problem, using conventional procedures, requires more than four decisions by the students and contains 0 or 1 sub-problem.
High	Solving the problem, using conventional procedures, requires more than four decisions by the students and contains 2 or more sub-problems.

Type of solving processes. In recognition of the fact that procedural complexity was insufficient to describe the overall cognitive complexity of problems, the Video Study also classified each problem according to the type of mathematical solving process implicit in the problem statement as using a procedure, stating a concept or making connections. The definitions and an example for each category (from Hiebert et al., 2003) are shown in Table 3. In applying this classification to textbook problems, we needed to interpret each problem in terms of its implied solution strategy. The Video Study found that the solving process implicit in a problem statement was not necessarily the solving process that was actually used in a public presentation of the problem, with a tendency (noted above) for 'making connections' problems to be solved publicly by 'using a procedure' or 'stating a concept'. The 'making connections' category is the most demanding, and, according to Stein and Lane (1996), most linked to student gains.

Repetition. A problem was classified as a repetition if it "was the same, or mostly the same, as a preceding problem in the lesson. It required essentially the same operations to solve although the numerical or algebraic expression might be different" (Hiebert et al., p. 76). For example, the equations $2x + 3 = 8$ and $3x + 5 = 16$ are essentially the same, so solving the second equation is regarded as repetition of the first. Some textbook problems that appeared to be repetitive, on closer inspection revealed differences that would be significant for eighth-grade students. For example, at this level students are still likely to be gaining confidence with addition and subtraction of negative integers so we did not classify $-2x - 7 = -13$ as a repetition of $2x + 3 = 8$.

Table 2
Examples of Applying the Video Study Definitions for Procedural Complexity

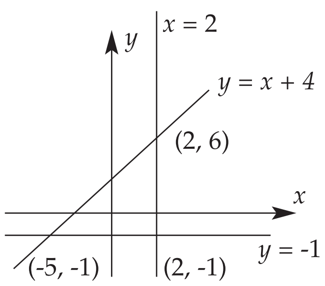
<table>
<tr><td colspan="2">Example 1: Low procedural complexity
Solve the equation: $2x + 7 = 2$</td></tr>
<tr><td>$2x + 7 – 7 = 2 - 7$
$2x = – 5$
$\frac{2x}{2} = \frac{-5}{2}, x = \frac{-5}{2}$</td><td>Step 1: Subtracting 7 from both sides.
Step 2: Recognising that 2 – 7 is -5. For students in eighth grade, working with negative integers is a relatively new skill, so this was regarded as a separate step.
Step 3: Dividing both sides by 2.</td></tr>
<tr><td colspan="2">Example 2: Moderate procedural complexity
Solve the set of equations for x and y: $2y = 3x$; $2x + y = 5$</td></tr>
<tr><td>$2y = 3x$ (1)
$2x + y = 5$ (2)
$2x + 2y = 10$
$4x + 3x = 10$
$7x = 10$
$\frac{7x}{7} = \frac{10}{7}, x = \frac{10}{7}$
$2y = 3x$
$2y = \frac{3x10}{7} = \frac{30}{7}$
$y = \frac{15}{7}$</td><td>Step 1: Inspect equations to decide on an appropriate strategy.
Step 2: Multiply equation (2) by 2.
Step 3: Substitute equation 1 ($2y = 3x$).
Step 4: Add $4x$ and $3x$.
Step 5: Dividing both sides by 7.
Step 6: Substitute in equation (1).
Step 7: Divide both sides by 2.</td></tr>
<tr><td colspan="2">Example 3: High procedural complexity
Graph the following linear inequalities and find the area of intersection: $y < x + 4$; $x > 2$; $y > -1$</td></tr>
<tr><td>x-intercept: -4, y-intercept:4)

</td><td>Steps 1, 2 and 3: Find intercepts for $y < x + 4$ and use in graphing.
Steps 4, 5: Sketch graphs $x = 2$ and $y = –1$.
Steps 6, 7: Find intercepts for $y = x + 4$ and use in graphing.
Sub-problem:
Steps 8, 9: Find coordinates of intersections.
Step 10: Decide on required region.
Sub-problem:
Steps 11, 12: Find base and height of right-angled triangle.
Step 13: Calculate area of triangle.</td></tr>
</table>

Table 3
Defining the Types of Solving Processes Implied by Problem Statements (from Hiebert et al., 2003, p. 98)

Solving process	Description
Using procedures	Problem statement suggested the problem was typically solved by applying a procedure or set of procedures. These include arithmetic with whole numbers, fractions, decimals, manipulating algebraic symbols to simplify expressions and solve equations, finding areas and perimeters of simple plane figures, etc.
	Example: Solve for x in the equation $2x + 5 = 6 - x$.
Stating concepts	Problem statements that called for a mathematical convention or an example of a mathematical concept. Examples: Plot the point (3, 2) on a coordinate plane. Draw an isosceles triangle.
Making connections	Problem statement that implied the problem would focus on constructing relationships among mathematical ideas, facts or procedures. Often, the problem statement suggested that students would engage in special forms of mathematical reasoning such as conjecturing, generalising, and verifying. Example: Graph the equations $y = 2x + 3$, $2y = x - 2$ and $y = -4x$, and examine the role played by the numbers in determining the position and slope of the associated lines.

Exercise or application. The Video Study classified problems as either exercises, for example, practising a procedure on a set of similar problems, or applications, where students apply procedures they have learned in one context to solve problems about a different context. Contrary to common Australian use of the term, applications do not necessarily have real-life references. Hiebert et al. (2003) note that applications require students to make decisions about how and when to use procedures they may have already learned and practised and are therefore "by definition, more conceptually demanding than routine exercises for the same topic" (p. 90). An example of an application problem based on the practised procedure of solving equations is: "The sum of three consecutive integers is 240. Find the integers." (Hiebert et al., 2003, p. 90). The problem in Figure 1 is also classified as an application problem in the Video Study. The classification as an application problem is another way of trying to measure cognitive demand and making links between mathematical ideas, rather than a way of examining the real world relevance of the teaching.

Find the measure of angle x.

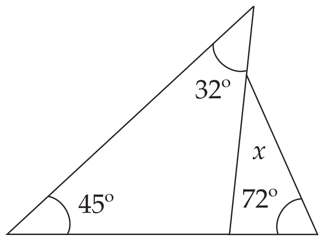


Figure 1. Video Study example of application problem (Hiebert et al., 2003, p. 90).

Proof, verification or derivation (PVD). The Video Study sought to measure the frequency of deductive reasoning, by identifying problems involving proof, verification or derivation (PVD problems). In the Video Study, proof was defined as "the process of establishing the validity of a statement, especially by definition from other statements in accordance with principles of reasoning"; verification was defined as "the act or process of ascertaining the truth or correctness of a rule"; and derivation was defined as "a sequence of statements showing that a result is the necessary consequence of previously accepted statements" (LessonLab, 2003a, p. 66). The Coding Manual indicates that to qualify as a proof, verification or derivation (PVD), the target result must apply to a class of problems (for example, proof of the Pythagorean theorem) rather than a single problem, must be non-numeric, and must be arrived at through deductive reasoning. One example that qualifies is: "prove that the sides of these two congruent triangles are equal" (p. 66). It is also noted that PVDs might be found more frequently in geometry lessons than in algebra or arithmetic lessons, and that not all problem statements that include the word 'proof' are PVDs by the Video Study definition. A problem that asked students to prove, for example, that -6 and 0.5 are solutions to the equation $2x^2 + 11x - 6 = 0$ would not be a PVD, whereas showing that $\left(-b \pm \sqrt{b^2 - 4ac}\right)/2a$ is an expression that generates the solutions to equations of the form $ax^2 + bx + c = 0$ is a PVD (LessonLab, 2003a, p. 67).

Reliability of Coding

Testing the reliability of the coding was the next stage of the study. As noted by Hiebert et al. (2003), "the importance of clearly defining and applying codes to the data, and then making sure that the coders are categorising the data as consistently and accurately as possible, is paramount" (p. 8). In the current study, the first author, in discussion with the second author, classified all problems. Subsequently, two research assistants were trained to apply the coding definitions. Approximately 30% of the 3687 problems were double-coded. Training commenced with an initial stage of familiarising the research assistants with the TIMSS Video criteria. Further training and discussion then took place in

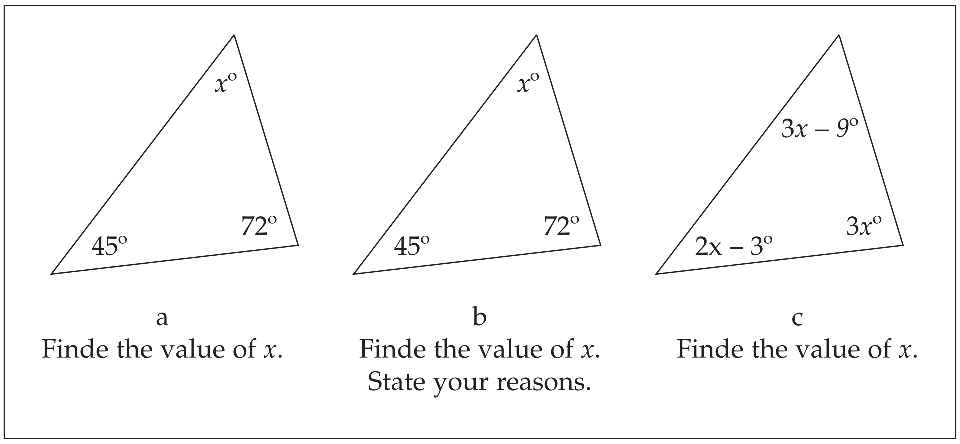


Figure 2. Three problems about finding unknown angles in a triangle

the initial stages wherever patterns of differences emerged in coding. There were initial differences between coders (the first named author and research assistants), for example, in classifying certain triangle and quadrilateral problems that required students to find the size of an unknown angle in a triangle or quadrilateral. Three variants of these problems are illustrated by the examples in Figure 2. All textbooks included a set of problems resembling those in either Figure 2a or Figure 2b. Several textbooks also included problems of the type shown in Figure 2c.

One coder initially classified all three such problems as 'using a procedure', whilst the other classified the first (Figure 2a) as 'using a procedure', the second (Figure 2b) as 'stating a concept' and the third (Figure 2c) as 'making connections'. After discussion, it was agreed that the stated reason in Figure 2b (that the angles of a triangle add to 180o) was implicit in Figure 2a, so both problem variants were therefore classified as 'using a procedure'. By contrast, in Figure 2c the students needed to write an equation using the angle sum and solve it to find the value of x. In this case it was agreed that 'making connections' was a more appropriate classification.

Differences in coding also occurred at first in classifying the procedural complexity of word problems that required students to formulate and solve an equation. When such problems were presented by the textbook in several parts—writing an equation, solving the equation, then interpreting the solution in terms of the problem statement — it was agreed that the problems should in fact be classified as three problems of (usually) low, rather than a single problem of moderate or high, procedural complexity. Occasionally it was also found that an incorrect coding had been inadvertently assigned.

As a measure of reliability of the coding, the percentage agreement between the coders was calculated by dividing the number of agreements by the number of agreements plus disagreements (see Hiebert et al., p. 8). Table 4 shows, for

example, the percentage agreement between coders after training for coding of the solving process, procedural complexity and repetition for the triangle/quadrilateral problems. This level of agreement between coders was at least as high as the reliability of 85% accepted in the TIMSS Video Study (Hiebert et al., 2003, p. 166) and was considered to be acceptable. The other topics had similarly high levels of agreement.

Table 4
Percentage agreement between two coders for classifying triangle/quadrilateral problems by textbooks A to I

	A	B	C	D	E	F	G	H	I
Procedural complexity	90	95	95	89	97	92	88	89	92
Method of solving	96	88	100	92	87	91	91	90	95
Repetition	85	89	100	88	93	97	92	94	95

Results

The results of the coding of the problems on each criterion are discussed in turn. Appendices 1, 2, and 3 give the percentage of problems in each category separately for the three topics (fractions, equations, geometry). The graphs in the sections below summarise the findings by looking at the proportions in the three topics combined.

Procedural complexity

The majority of problems in all textbooks were of low procedural complexity. For fractions, the percentage of problems of low procedural complexity ranged from 56% to 83% (see Appendix 1), which was very similar to the range from 58% to 83% for equations (see Appendix 2). Most textbooks had a higher proportion of problems of low procedural complexity for geometry than for the other two topics, ranging from 73% to 96% (see Appendix 3).

Problems of high procedural complexity were rare (0% - 4%) (see Appendices 1, 2 and 3) in all textbooks. This was lower than for the Video Study Australian lessons, where 8% of problems were classified as high procedural complexity.

To enable a better comparison with the Video Study data, the moderate and high procedural complexity categories were combined to make a two-level classification of low procedural complexity against moderate/high procedural complexity. Figure 3 shows the percentages of moderate/high complexity problems in each textbook. It is important to note that two of the textbooks with the lowest percentage of problems with moderate/high procedural complexity (B and D) were the best selling textbooks in their respective states.

Figure 3 also enables comparison with two measures from the Video Study. The Australian lessons in the Video study had 24% of problems of moderate/high procedural complexity on average, matched only by two

textbooks (E and I). Figure 3 also shows the considerably higher median percentage of moderate/high procedural complexity problems for the seven countries in the Video Study (34%). The median percentage has been used in this and later comparison graphs because the high outlier country of Japan, where mathematics is taught very differently, skews other measures of central tendency.

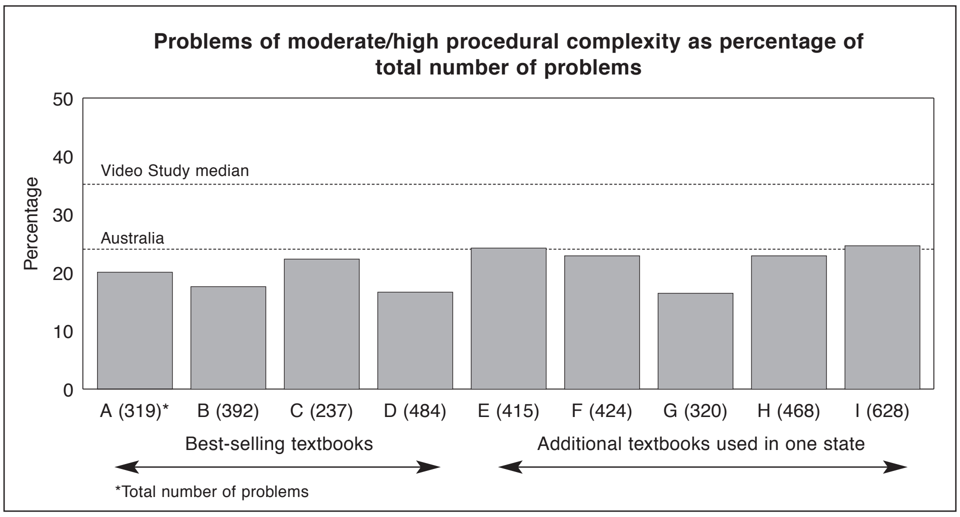


Figure 3. Percentage of problems of moderate/high procedural complexity.

Type of Solving Process

For fractions and for equations, the majority of problems in all books were classified as 'using procedures'. The proportion of 'using procedures' problems for fractions ranged, for example, from 71% (H) to 100% (D) (see Appendix 1) and from 55% to 88% for equations (see Appendix 2). Geometry had lower proportions of 'using procedures' problems with all but one textbook in the range of 33% to 58% (see Appendix 3); instead many more geometry problems were classified as 'stating concepts'.

Examination of the types of problems that were classified as 'using procedures' and 'stating a concept' led us to conclude that these two classifications were generally of similar (low) cognitive demand in the eighth grade context. The problems in Figures 2a and 2b above illustrate the similarity. Other examples are problems requiring students to use a procedure such as solving a two-step linear equation, or to state a concept such as naming the type of triangle in a diagram. The comparison most relevant to the shallow teaching syndrome is therefore the proportion of 'making connections' problems.

Over the three topics, in the majority of books and including three of the four best-selling textbooks (A, B, and D), less than 15% of problems required students to make connections (see Figure 4). Only textbooks C, H and I exceeded the

Video Study median of 17% 'making connections' problems. With research demonstrating the importance of students making connections to students' learning of mathematics (see, e.g., Stein & Lane, 1996), the differences between books and the generally low levels are cause for concern.

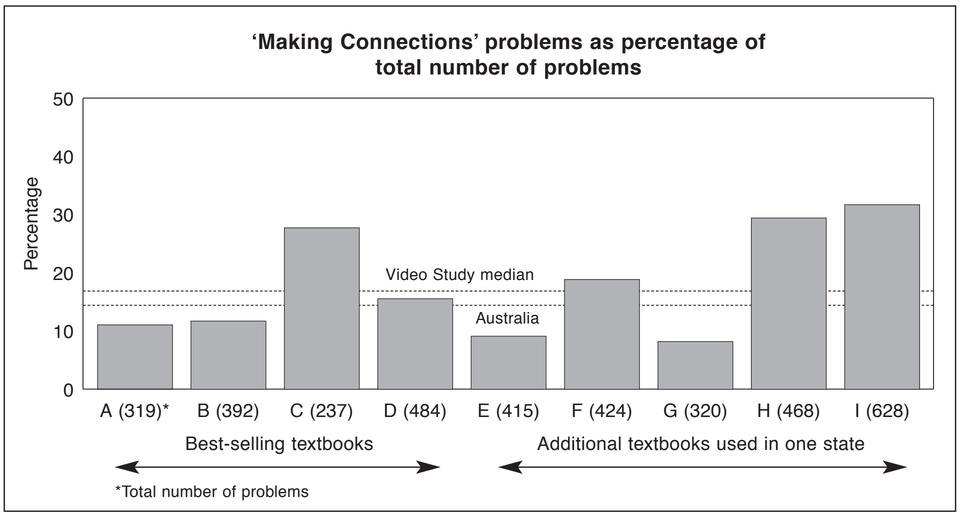


Figure 4. 'Making connections' problems as a percentage of the total number of problems.

Geometry generally had a higher proportion of 'making connections' problems, with only three of the books (B, C, and E) not reaching the Video Study median of 17% over all topics. In geometry, books D, H, and I included questions where students were required to make connections between various geometric concepts as, for example, in making a chain of reasoning using several known properties to find an unknown angle. By contrast, textbooks B, C, E, and G included few problems that needed more than application of one principle (e.g., students were given two interior angles of a triangle and asked to find the third angle or to find the exterior angle).

Repetition

The percentage of problems that were repetition was generally in the range of 25% to 60%, varying between textbooks and topics. In textbook C, for example, 50% of equations problems (Appendix 2) and 71% of geometry problems (Appendix 3) were repetition. However, in this textbook, 'fractions' was marked as a revision topic, and in the brief set of problems provided, only 6% were repetition (Appendix 1). As shown in Figure 5, the best selling textbooks tended to have medium to high levels of repetition. Repetition was frequently associated with problem sets that required students to use a procedure or to state a simple

concept such as "name the type of triangle". By contrast, F, H and I had relatively low levels of repetition (around one problem in four) averaged over all three topics. As Figure 5 shows, the overall level of repetition in all books was lower than the 76% repetition identified in the Australian Video Study lessons and the international median.

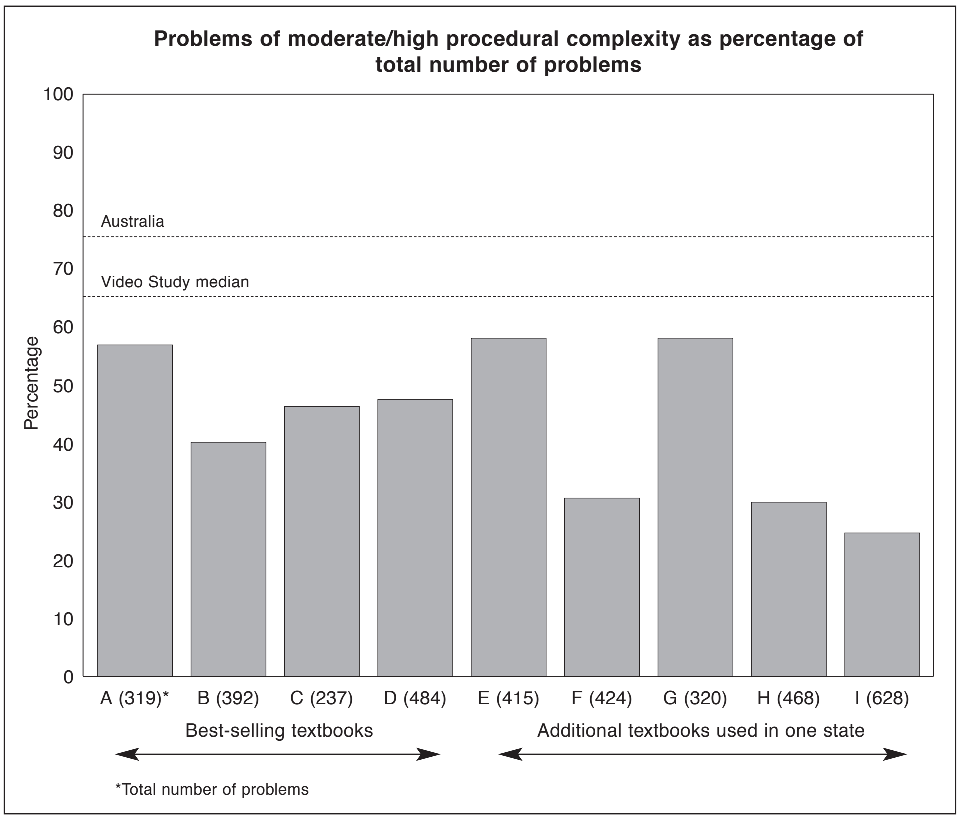


Figure 5. Percentage of total number of problems for three topics that were repetitions.

Related to the issue of repetition (but not measured by it) was the positioning of worked examples relative to problem sets. Most books tended to place all worked examples before the related problem set, so students working on problems had to select which worked example to follow. In textbook A, however, each worked example was followed immediately by repetitions of that worked example, before the next worked example, and so on. Whilst this design provided strong scaffolding, it reduced the extent to which students needed to choose an appropriate solution strategy. This is another aspect of repetition for textbook writers to consider.

Exercise or Application

Classification of textbook problems as either exercises (practising of procedures) or applications (apply procedure from one context in another, not necessarily a real life context) indicates varying emphasis according to the topic and the textbook. Appendices 1, 2 and 3 show that the proportion of application problems ranged from 0% (textbook A for equations and textbooks C, D and F for fractions) to 61% (textbook I for geometry). As shown in Figure 6, for all three topics combined, the percentages of applications were substantially below, and usually substantially below, the Video Study median (and the Australian mean) of 45% application problems (Hiebert et al., 2003, p. 91).

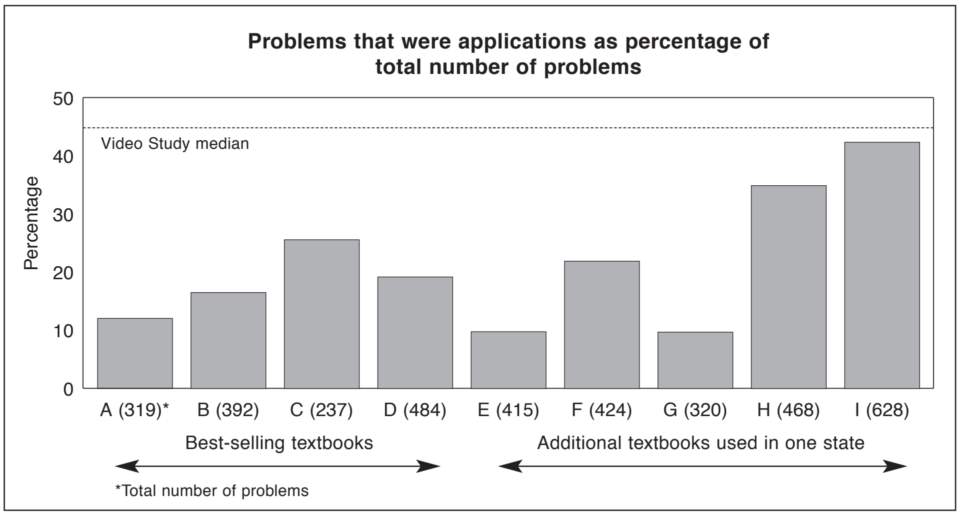


Figure 6. Percentage of application problems for three topics.

Proof, Verification or Derivation (PVD)

Problems that could be classified as PVDs were found only in the geometry sections, and then in only six of the nine textbooks (see Appendix 3). The graph in Figure 7 therefore shows the percentage of the total number of geometry problems that were PVDs, not the overall percentage as with the other graphs. The six textbooks with PVDs included demonstration proofs, for example, that the angles of a triangle added to 180o or that an exterior angle of a triangle was equal in size to the sum of the two interior opposite angles. Textbook C had one demonstration proof, but this was the only PVD problem: there were no problems for students to do. In textbooks A, B and D, some problems that required students to write informal explanations satisfied the Video Study definition for being classified as a PVD. Examples include "Can a triangle have more than one obtuse angle. Explain" (textbook A) and "Why are the opposite

angles of a parallelogram equal" (textbook D). Only books H and I emphasised deductive proof and explicitly asked students to construct proofs as a logical sequence of reasoning in a similar way to that shown in the worked examples.

Comparing Textbooks with the Video Study Findings

Overall, there was a broad similarity in the proportions of problems in each category in the textbooks and the Australian Video Study lessons, although considerable variation existed between textbooks and topics. None of the textbooks had a higher percentage of problems of high or moderate procedural complexity that those documented in the Video study (Figure 3).

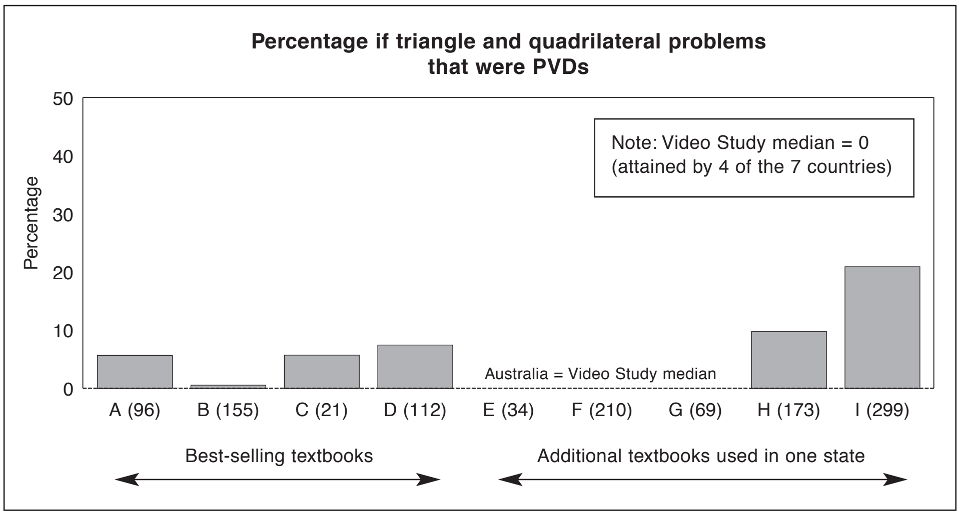


Figure 7. Percentage of geometry problems that were PVDs.

It seems reasonable to conclude that Australian eighth grade students are not exposed to many complex problems. The percentage of 'making connections' problems in the Video study seemed to be aligned with the percentages in the textbooks, with some textbooks above and some below: this was also near the international Video study median. Repetition in every textbook was lower than the Australian average for the video study. This is surprising, since we expected that textbooks would be higher because they provide a bank of homework problems for skills practice. This seems an important result to explain; we do not have a convincing argument that it is an artifact of methodology. It seems that students are seeing less variation in the problems in a lesson than amongst the problems in a textbook. Every textbook showed a lower percentage of application problems than the Australian video study mean. Since this was another way of looking at links across topics (rather than real-world relevance,

on which Australia was in the middle band in the Video study), the very low percentages in some textbooks show that some students are exposed to only very basic problems. This textbook analysis confirmed that PVD problems are very rare in Australian textbooks at eighth grade. It would be interesting to examine textbooks for older students to see where PVD first has a significant presence.

There are several important reasons why the findings of the Video Study and this analysis of textbook problems would be expected to differ and this study does not enable us to examine these in detail. As discussed elsewhere in this paper, the problems selected for a class may not be representative of the problems which students do. Furthermore, the Video Study did not consider each lesson in the context of what preceded and followed. As noted by Clarke, Keitel, & Shimizu (2006, p. 2), "analyses of classroom practice that do not take into account the situation of the lesson within the enfolding topic, ignore one of the major influences on the teacher's purposeful selection of instructional strategies". A lesson that focused on making connections between different concepts, for example, may have been preceded by lessons in which students practised using procedures. By contrast, the present analysis of textbook problems represents complete units of work on their respective topics. On the other hand, in any comparison of textbook problems with the problems observed in the Video Study lessons, it should be kept in mind that the textbook is also a resource for homework tasks, which may often focus on the completion of routine practice exercises, leaving class time for introducing new topics, and solving problems which present greater difficulty to students. It is common, too, for students not to do all the textbook problems, which could skew the percentages of problems in students' actual experience towards either the low or high end. A further barrier to direct comparison with the Video study is the time difference, since we compared the 1999 Video study with current textbooks. The success of this method of measuring textbook problems suggests that a comparison of 1999 and current textbooks may be warranted.

Application of Video Study Criteria

The Video Study definition of low procedural complexity (four or fewer steps and no sub-problems) seemed insufficiently sensitive for problems in the either grade Australian textbooks. Consider, for example, the equations $2x + 3 = 11$ and $\frac{3(x-4)}{5} + 7 = 12$. Both may be solved in four or fewer steps and would therefore be classified as low procedural complexity. The second equation, however, places greater cognitive demand on students in deciding the order of performing operations and involves more complex arithmetic. It could be argued, then, that the high proportion of problems of low-procedural complexity is to some extent an artifact of the Video Study definitions, which seem too strict in the current Australian context.

When considering the type of solving process, it is important to note that the solving processes implicit in textbook problems are not necessarily matched by the solving processes that actually occur. Although a textbook problem may be

written with the intention of prompting students to make connections between different mathematical concepts, these connections may not be made. Also, as suggested by Stein and Lane (1996), teachers may 'proceduralise' problems in order to assist students to reach a solution. By contrast, it is also likely that a teacher, in demonstrating the solution of a textbook problem, may engage students in making connections, even in cases where the problem implies the use of a procedure. Hiebert et al. (2003, p. 97) give the following example of how solving the equation $2x + 3 = 11$ may be straightforward 'using procedures' or might be exploited by skilful questioning to become 'making connections'.

> The problem statement suggests that a procedure will be used to find x, perhaps subtracting 3 from both sides of the equation and then dividing both sides by 2, yielding $x = 4$. If the problem actually is solved in this way (by the teacher or the students) without further examination, the mathematical processes suggested by the statement and those used while solving the problem are the same. The processes could be called "using procedures." But imagine that the teacher asks some additional questions as the problem is being solved: "If the equation was written $11 = 2x + 3$, would the solution be the same?" or "Is it OK to divide both sides of the equation by any number?" And imagine that the teacher follows the questions with a discussion on, for example, the concept of transforming equations in ways that preserve equivalence.

Hence the kinds of processes that students actually engage in may be quite different from the intended or implied processes (see also Hiebert et al., 2003, p. 103). Nevertheless, classifying textbook problems according to the implied solving process indicates the extent to which students are exposed to each of the solving processes if the teacher were to follow the textbook closely. It is of concern, then, that five of the nine textbooks, including three that were the best-selling books in their respective states, fell below the Video Study international median for the category of making connections.

Textbook Problems

The study has shown substantial differences between textbooks in the balance of problem types. Having two classifications, one for procedural complexity and one for type of solving processes, highlights the fact that lower procedural complexity does not necessarily imply lower quality of problems in terms of challenging students to make connections or to reason, and vice versa. Whilst most of the textbooks had a predominance of problems of low procedural complexity, of greater concern is the low level of making connections problems in several books, including three of the four best-selling textbooks in their respective states. These findings may support Pehkonen's (2004) claim that teachers want textbooks to focus on the basics. The lack of emphasis on making connections is also in agreement with the findings of Stein and Lane (1996) and Pepin and Haggarty (2001). As Schoenfeld (2004) asserts, there needs to be a balance between acquiring mathematical skills and experiencing the processes which enable students to use mathematics.

In the case of geometry, a typical Australian eighth grade textbook problem with low procedural complexity gives students two angles of a triangle and asks them to find the third angle. Only two steps are required: finding the sum of the two given angles and subtracting the sum from 180o, a routine procedure for eighth-grade students. By contrast, the application problem shown in Figure 8a, although still classified as low procedural complexity, requires students to recognise that they need to modify the diagram in some way in order to find the unknown angle, X. This problem formed part of one of the Japanese Video Study lessons (LessonLab, 2003b, example 5.4: application). In the previous lesson the teacher had summarised three ways of approaching angle problems: parallel lines, triangles or quadrilaterals.

During the solving of the problem in the Video lesson, the teacher observed students use each of these different approaches and three students were invited to demonstrate their methods. The problem can be solved in three steps, as shown using the parallel line method (see Figure 8b): adding the third parallel line, recognising the equal alternate angles, and adding to find the required angle. However, this requires "making connections" insight: to recognise that drawing the extra line will allow them to solve the problem by applying their knowledge of angles associated with parallel lines. (Alternative solutions require similar insights.) As for the above equation example, this problem also illustrates that the cognitive complexity of a problem can be at odds with its low procedural complexi\ty classification.

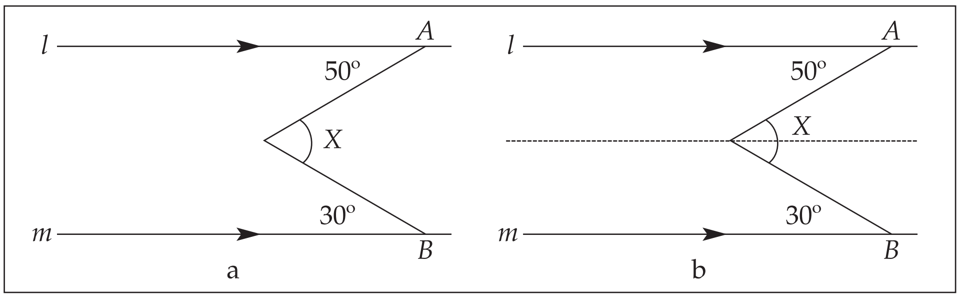


Figure 8. Geometry application problem from a Japanese lesson showing mismatch between procedural complexity and cognitive complexity.

Whilst the focus in most textbooks in the sample was on procedures and stating concepts, some books did include problems designed specifically to help students (and perhaps inexperienced teachers) to make connections. Consider, for example, the problem from textbook H shown in Figure 9. The question has a metacognitive aim, with the intention of focusing students' attention on when it is preferable to expand the brackets first. Problems of this type may be considered in terms of Herbel-Eisenmann's (2007) discussion of the 'voice' of a mathematics textbook'. Use of words such as 'prefer' suggest to students that alternative solution pathways are possible and that they may exert some choice over their particular solving strategy.

(a) Solve the equation $3(2x + 3) = 27$

(i) by first expanding the brackets (ii) by first dividing both sides by 3

(b) Which method do you prefer? Explain.

(c) Which method would you use if you were solving the equation $3(3x - 5) = 11$? Explain.

Figure 9. Equation-solving problem from textbook H.

Several textbooks included equations problems that qualified as moderate procedural complexity on the basis of requiring more than four steps, for example, $6(p + 3) = 5(2 - p) + 7p$ -12 (textbook D). However, apart from deciding upon the order of steps, the student simply repeats a greater number of times the same types of operations: expanding brackets, dealing with positive and negative signs, collecting like terms, and so on. It is important that students should be able to solve equations involving multiple steps, so these moderately complex equations have a place. In analysing the textbooks, it appeared that some writers were choosing between including these problems with moderate/high procedural complexity or problems that were 'making connections' or 'applications'.

Problems of high procedural complexity were rare, and in fact most textbooks had no problems that matched the definition for this classification. Potentially high procedural complexity problems were generally broken down into a number of clearly stated sub-problems, each of which then became a separate problem of low (or occasionally moderate) procedural complexity. This common problem structure is largely the reason for the virtual absence of problems of high procedural complexity in the sample of textbooks. Whilst this type of scaffolding initially may be necessary in order to develop students' confidence in tackling problems with high procedural complexity, it removes the necessity for students to make connections themselves between information given in the problem statement and to plan a path through sub-problems in order to reach the target result. We could argue, however, that although this scaffolding reduces cognitive complexity, it may actually train students to make connections.

Our examination of the range of problems in the textbooks suggests that a two-way classification on both procedural complexity and type of solving process provides a useful picture of the underlying balance of problems in textbooks. This method of analysis is illustrated in Table 5, which shows the distribution of equation-solving problems for textbook D. The analysis highlights the bias towards problems of low procedural complexity that require only the use of a procedure or the stating of a concept. For this topic, this textbook provides little to assist students in making connections across mathematical topics; nor does it cater well for high-achieving students, for whom the balance should be towards a greater proportion of moderate and high procedural complexity problems that require them to plan solution pathways.

Table 5
Number of 'solving equations' problems for textbook D classified by both procedural complexity and type of solving process (N = 360)

	Procedural complexity	
Type of solving process	Low	Moderate/high
Using a procedure/stating a concept	277	39
Making connections	20	24

Limitations of the current study

When interpreting the textbook analyses, it should be kept in mind that only three topics were analysed, and that no attempt was made to search the textbook for problems relating to these three topics in other sections of the books. Analysis of the entire textbooks may have presented a different picture. Whether eighth grade is the first or second year of secondary school would also be expected to have some effect on the balance of problem types in textbooks from different states, but this is not evident in the data presented.

It became evident during the classification that using the percentage of problems in each category as the basic measure may be less desirable than looking at the percentage of time spent. In both lessons and in private work, problems of higher procedural complexity and applications and those requiring students to make connections might be expected to take more time to complete or, put another way, several low procedural complexity problems could be solved in the same time as one problem of high procedural complexity. Time spent on problems of each type would perhaps be a more appropriate measure. Even using this measure, though, the Video Study notes that an average of 65% of the private work time of students in the Australian lessons was spent 'repeating procedures' (see Hiebert et al., p. 104; we interpret this to mean repeating the same types of 'using procedures' problems that had been encountered earlier in the lesson). Although this was not the highest for the seven countries (84% for the Czech Republic and 81% for Hong Kong), the equivalent measure for the Japanese lessons was only 28%.

The three topics sampled for this study showed significant differences, both within textbooks and between textbooks. Different decisions about how a topic should be treated, and possibly different chapters being written by different authors, led to differences in the balances of problem types. However, there are also reasonably consistent differences between the three topics across textbooks that reveal differences between the intellectual processes they require or with which they are traditionally associated. Geometry, for example, was the only topic where PVD problems were observed.

Conclusions

For the three topics investigated, textbooks were found to vary in their relative

emphases on procedural skills and applications, in the level to which they supported the making of connections, and, in the case of geometry, in their inclusion of proof. Whilst our study shows a broad similarity between textbook problems and the Australian Video Study lesson problems, we should be cautious in our interpretation of the findings. Different types of problems play different pedagogical roles. It is important that textbooks provide students with sufficient problems so that procedures may be practised and become a secure part of a student's mathematical toolbox. A certain level of repetition is therefore desirable. Likewise there should be sufficient problems for students to learn to apply those practised skills, to encourage them to make connections between different aspects of mathematics and to recognise underlying mathematical concepts. There should also be problems that stimulate reflection and reasoning, and require students to plan a solution pathway by identifying sub-problems. Some high procedural complexity problems are also desirable so that students gain experience in sustaining a chain of reasoning without error.

It was evident during the classification process that the classifications do not show which are 'good' problems, and that there are problems in all categories that provoke and do not provoke mathematical thought. A problem such as 'plot the point (3, 2)', for example, is classified as 'stating concepts', but it may stimulate less learning than a simple 'using procedures' problem. It is not that 'using procedures' problems and problems of low procedural complexity are 'bad' of themselves, but that their dominance curtails the experiences that students have of mathematical thinking.

What is important, then, is that students are presented with a balanced curriculum experience. The balance will need to be different for high and low achieving students, but all students need exposure to the full range of problem types. Our analysis suggests that in some books, including the best-selling textbooks in several states, the balance is too far towards repetitive problems of low procedural complexity that require little more than using procedures.

Different teachers obviously have different expectations of a textbook, depending on how they choose to supplement the textbook with their own resources and pedagogical experience. Whilst there are books that support teachers in providing students with a strong basis for conceptual understanding, making connections and reasoning, some textbooks, including best-selling textbooks, can be seen to be contributing to shallow teaching. This is of particular concern in classrooms where teachers rely heavily on 'textbook teaching'. It is also the case that these books provide little support for new teachers or teachers lacking strong mathematical backgrounds. There are important messages here for curriculum writers and textbook writers, and perhaps a strong argument for textbooks to have an accompanying teachers' guide that focuses on the pedagogical intentions of the textbook material.

Acknowledgements

The authors would like to acknowledge the advice of Gail Hood and David Rasmussen who were involved in the TIMSS Video Study.

References

Ball, D., & Cohen, D. (1996). Reform by the book: What is—or might be—the role of curriculum materials in teacher learning and instructional reform? *Educational Researcher, 25*(9), 6-8, 14.

Brändström, A. (2005). Differentiated tasks in mathematics textbooks: An analysis of the levels of difficulty. Licentiate Thesis, Luleå University of Technology, Sweden.

Clarke, D., Keitel, C., & Shimizu, Y. (Eds.) (2006). *Mathematics classrooms in twelve countries: The insider's perspective*. Rotterdam: Sense Publications.

Herbel-Eisenmann, B. (2007). From intended curriculum to written curriculum: Examining the "voice" of a mathematics textbook. *Journal for Research in Mathematics Education, 38*(4), 344-369.

Hiebert, J., Gallimore, R., Garnier, H., Givvin, K. B., Hollingsworth, H., Jacobs, J., Chui, A. M-Y., Wearne, D., Smith, M., Kersting, N., Manaster, A., Tseng, E., Etterbeck, W., Manaster, C., Gonzales, P., & Stigler, J. (2003). *Teaching mathematics in seven countries: Results from the TIMSS 1999 Video Study*. Washington, DC: National Centre for Education Statistics, U.S. Department of Education.

LessonLab (2003a). *TIMSS-R video math coding manual*: www.llri.org/html/TIMSS/download/TIMSS%201999%20Video%20Coding%20Manual.pdf

LessonLab (2003b). TIMSS 1999 Video Study: Mathematics. Video clip examples.

Love, E., & Pimm, D. (1996). 'This is so': A text on texts. In A. Bishop, K. Clements, C. Keitel, J. Kilpatrick, & C. Laborde (Eds.), *International handbook of mathematics education*, pp. 371-409. Dordrecht, The Netherlands: Kluwer.

Mayer, R. E., Sims, V., & Tajika, H. (1995). A comparison of how textbooks teach mathematical problem solving in Japan and the United States. *American Educational Research Journal, 32*(2), 443-460.

Pehkonen, L. (2004). The magic circle of the textbook — an option or an obstacle for teacher change. In M. J. Høines & A. B. Fuglestad (Eds.), *Proceedings of the 28th conference of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp.107-136). Bergen: PME.

Pepin, B., & Haggarty, L. (2001). Mathematics textbooks and their use in English, French and German classrooms: A way to understand teaching and learning cultures. *Zentralblatt für Didaktik der Mathematik, 33*(5), 158-175.

Schoenfeld, A. (2004). The math wars. *Educational Policy, 18*(1), 253-286.

Stacey, K. (2003). The need to increase attention to mathematical reasoning. In H. Hollingsworth, J. Lokan, & B. McCrae, *Teaching mathematics in Australia: Results from the TIMSS 1999 video study* (pp. 119-122). Melbourne: ACER.

Stein, M. K., & Lane, S. (1996). Instructional tasks and the development of student capacity to think and reason: An analysis of the relationship between teaching and learning in a reform mathematics project. *Educational Research and Evaluation, 2*(1), 50-80.

Authors

Jill Vincent, Science and Mathematics Education, Melbourne Graduate School of Education, University of Melbourne, VIC 3010, Australia. Email: jlvinc@unimelb.edu.au

Kaye Stacey, Science and Mathematics Education, Melbourne Graduate School of Education, University of Melbourne, VIC 3010, Australia. Email: k.stacey@unimelb.edu.au

Appendix 1

'Addition and Subtraction of Fractions' Problems (Including Worked Examples) According to Five TIMSS Video Study Criteria for Sample of Australian Eighth-grade Mathematics Textbooks

Textbook	Number of problems	Procedural complexity (percentage of problems)			Solving process (percentage of problems)			Repetition %	Exercise or Application (percentage of problems)		PVD %
		Low	Moderate	High	Using procedures	Stating concepts	Making connections		Exercise	Application	
A	129	78	21	1	88	7	5	60	88	12	0
B	119	76	23	1	92	3	4	58	89	11	0
C	16	56	44	0	75	0	25	6	100	0	0
D	12	83	17	0	100	0	0	50	100	0	0
E	79	76	23	1	87	8	5	56	95	5	0
F	22	68	32	0	86	14	0	27	100	0	0
G[a]	0										
H	79	75	22	4	71	6	23	34	73	27	0
I	74	80	20	0	92	3	5	32	95	5	0

Note. [a]No fractions section

Appendix 2

'Solving Linear Equations' Problems (Including Worked Examples) According to Five TIMSS Video Study Criteria for Sample of Australian Eighth-grade Mathematics Textbooks

Textbook	Number of problems	Procedural complexity (percentage of problems)			Solving process (percentage of problems)			Repetition %	Exercise or Application (percentage of problems)		PVD %
		Low	Moderate	High	Using procedures	Stating concepts	Making connections		Exercise	Application	
A	94	73	27	0	76	21	3	60	100	0	0
B	118	77	20	3	55	22	23	32	77	23	0
C	200	79	19	3	58	12	31	50	71	29	0
D	360	83	17	0	88	0	12	57	87	13	0
E	302	74	26	0	83	6	11	61	90	10	0
F	192	58	41	1	64	18	19	28	81	19	0
G	251	82	17	1	8	6	7	59	93	7	0
H	216	75	24	1	60	8	32	30	77	23	0
I	255	75	22	4	56	9	35	33	67	33	0

Appendix 3

'Triangles and Quadrilaterals' Problems (Including Worked Examples) According to Five TIMSS Video Study Criteria for Sample of Australian Eighth-grade Mathematics Textbooks

Textbook	Number of problems	Procedural complexity (percentage of problems)			Solving process (percentage of problems)			Repetition %	Exercise or Application (percentage of problems)		PVD %
		Low	Moderate	High	Using procedures	Stating concepts	Making connections		Exercise	Application	
A	96	87	13	0	57	15	28	48	74	26	5
B	155	94	6	0	41	48	11	37	85	15	1
C	21	95	5	0	90	5	5	71	90	10	5
D	112	92	8	0	58	17	25	26	62	38	8
E	34	88	12	0	41	53	6	41	82	18	0
F	210	96	4	0	41	36	22	33	74	26	0
G	69	88	12	0	45	38	17	52	83	17	0
H	173	82	17	1	33	32	35	28	48	52	9
I	299	73	26	1	39	21	40	16	39	61	21