How Heavy is my Rock? An Exploration of Students' Understanding of the Measurement of Weight

Michael Drake
Victoria University of Wellington
<Michael.drake@vuw.ac.nz>

New Zealand and Australian curricula require students to learn about weight/mass for at least six years. However, little research identifies what should be taught. This study reports cognitive interviews with 17 Year 9 students who were asked *how heavy is my rock?* Only one student demonstrated some understanding of how to use analogue kitchen scales, most had multiple errors. Results suggest that teachers would benefit from better guidance about teaching the skill set needed for such a task.

Curriculum documents in Australia and New Zealand introduce weight/mass measurement in the first year of schooling, with learning about the attribute continuing until at least Year 6 (Australian Curriculum, Assessment and Reporting Authority, ACARA, 2012; NZmaths, 2013). However, there has been little research that focuses on students' developing understanding of weight/mass, and what there is tends to focus on the early years (Cheeseman, McDonough, & Clarke, 2011). This lack of evidence is problematic for mathematics educators for two reasons. Firstly, in relation to weight/mass, the learning progressions that underpin the curricula are not based on evidence, but rather educated guesses. Secondly, beyond the first years of school there is little guidance for teachers seeking to develop students' understanding of weight/mass. What should be focused upon?

When measuring weight, a common task is to weigh food on a set of analogue kitchen scales. Here the term weight (rather than mass) is appropriate because spring balance scales react to the pull of gravity on whatever is placed in the pan (Van de Walle, Karp, & Bay-Williams, 2013). In The Australian Curriculum (TAC) this task is part of the content described for Year 4 (ACARA, 2012). In The New Zealand Curriculum (NZC) the task fits best with Years 5 and 6 (Curriculum Level 3) (NZmaths, 2012). Are these levels suitable? Are there particular issues to which the supporting documents for these curricula should alert teachers? To answer these questions, this paper explores students' understanding of weight measurement by examining their responses when asked to weigh a rock on a set of kitchen scales. However, in the absence of literature related to measuring weight on scales, the literature related to measuring length with a scale, namely the broken ruler problem, and a previous study by the author on factors influencing students' ability to use linear scales (Drake, 2011) have been used to inform the exploration.

Literature review

Studies of students' developing understanding of weight or mass are uncommon (Cheeseman et al., 2011). The National Council of Teachers of Mathematics research compendia do not address the subject (Grouws, 1992; Lester, 2007) and books for teachers about effective practice have little, if anything, about using analogue weighing scales (e.g., see Bobis, Mulligan, & Lowrie, 2013; Van de Walle et al., 2013). The study by Cheeseman et al. (2011) reports the developing understanding of mass up to the end of Grade 2. By the end of Grade 1, 1% of the 479 students were able to use a set of analogue kitchen weighing scales, while 6% of the 256 students were able to complete the same tasks by the end of

In V. Steinle, L. Ball & C. Bardini (Eds.), Mathematics education: Yesterday, today and tomorrow (Proceedings of the 36th annual conference of the Mathematics Education Research Group of Australasia). Melbourne, VIC: MERGA.

Grade 2. However, their paper reports, but does not elaborate, on this finding. Cheeseman, McDonough, and Ferguson (2012) report the results of a teaching experiment with Year 1 and 2 students. They showed that a short intervention (5 one-hour periods) could lead to large gains in understanding the measurement of mass, with 53% of Year 1 and 84% of Year 2 students being able to use standard units by the end of the intervention. However, these percentages seem to involve using balance, digital, and analogue scales; further details were not provided. In other studies, MacDonald (2010, 2011) discusses young children's concepts of mass, but does not address the use of weighing scales.

In place of research into students' understanding of weight and weighing scales, research into students' understanding of length was used to inform the study. The broken ruler problem has been widely used as a measure of length understanding (e.g., Irwin & Ell, 2002; Lehrer, Jenkins, & Osana 1998). Different forms of the problem can involve measuring an object with a broken ruler, or identifying the length of an object drawn next to an offset ruler where the start of the object does not align with zero. Results for this problem suggest that learning to use a measuring scale is not easy. For example, results from the National Assessment of Educational Progress (NAEP) consistently show that since the 1980s, roughly 60% of American 13-year-olds are successful with the broken ruler problem (Kloosterman, Rutledge, & Kenney, 2009). One commonly reported error is that some students simply read the number on the ruler that aligns with the end of the object; another is that students count from 1 to measure the object, starting with 1 at the beginning of the object (e.g., Lehrer, 2003). Such errors are interpreted to mean that the students do not understand the fundamental nature of units and unit iteration (e.g., Irwin & Ell, 2002; Lehrer et al., 1998). By using a problem similar to the broken ruler problem for this research, results across both problems may be compared.

In a previous paper, Drake (2011) focused on the sources of error that 184 Year 7 and 8 students had when working with linear scales on number lines and in measurement and graphing. Using a linear scale was not straightforward for the students; some made multiple errors in a single problem. In general, students' understanding of linear scale was related to their understanding of number and measurement. In particular, student understanding of the conventions used on a linear scale, their understanding of the role played by the marks and the spaces, and their knowledge of partitioning strategies could all affect their response.

A commonly used conceptual model for planning the teaching of measurement attributes is described by Zevenbergen, Dole, and Wright (2004) and Cheeseman et al. (2011). The first step involves focusing on identifying the attribute, the second, comparing and ordering a set of items (seriation), and the third informal measurement using uniform but non-standard units – to give the idea that different units can be used in different contexts, and to establish the need for common units. At the fourth step, formal measurement with standard units is developed while the fifth considers application to problem-solving contexts – including the development of formulae. In Bobis et al. (2013) a similar learning framework of six steps is introduced: identify the attribute; informal measurement; develop standard units and the structure of the iterated unit; measure using formal units; identify and using relationships between units; and application of measurement ideas to life-like contexts. In neither progression, nor in the examples of activities that exemplify the progressions, is it clear where learning to use a set of analogue kitchen scales fits. For example, in Zevenbergen at al. (2004), step 4 seems logical and a single line tells readers to "(p)rovide lots of experiences with weighing everyday items" (p. 270). However, the emphasis at this step is on developing a sense of the basic unit, appropriate language and symbols to describe it, the size of familiar referents, and estimation in relation to the unit. By comparison, Cheeseman et al. (2011) use a simple analogue scale to measure students' ability to apply knowledge, skills and concepts in step 5, whereas use of analogue scales seems to be part of step 4 in Cheeseman et al. (2012).

Method

Seventeen Year 9 students (aged 13 to 14 years) from a coeducational secondary school in the Wellington region of New Zealand participated in the research. The school is of medium size and is situated in a high socio-economic area. The students were chosen by the Head of Mathematics (HOD) from across Year 9 to represent students of a range of abilities. The HOD used mathematics Progressive Achievement Test (PAT) results to identify their ability. PATs are developed by the New Zealand Council for Educational Research (NZCER) and are nationally normed (NZCER, 2012). All Year 9 students sat this test after entry to the school so the results provided both recent and valid measures of mathematical knowledge. The research took place early in term 2, after students had already completed their measurement unit for the year.

The task reported is one of 30 items from various mathematical domains that were used in an interview to identify student understanding of linear scale. The appropriateness of the task for the students was measured against NZC (NZmaths, 2013). This identified that students should be using linear scales and whole numbers of metric units for weight (mass) at CL3 (the standard programme for Years 5 & 6). By comparison, the National Standards documentation indicates that by the end of Year 4, students can measure the weight of objects by reading linear scales with standard units to the nearest whole number (NZmaths, 2013). Students from Year 9 should thus have the skills to successfully complete the task.

For the task, when the equipment was revealed, each student saw a rock resting in the pan of a set of kitchen weighing scales. The scales had gauges both in pounds and kilograms as is common for scales produced in China and sold internationally. The kilogram gauge (the one students were expected to use) had every kilogram and 100g marked and labelled, with each 100g interval being further partitioned into multiples of 25g. As the rock weighed 390g, when it was in the pan the needle clearly did not align with one of the marks on the gauge (Figure 1a). The scales had not been properly zeroed and had been set at -75g; that is, the needle rested at -75g when the pan was empty (Figure 1b). Students were told that they could do whatever they needed to do with the scales to work out "how heavy is my rock?" To ensure that the same situation was presented to each student, the scales were carefully reset between interviews.

Cognitive interviews (Presser et al., 2004) were used to elicit responses to a semi-structured questionnaire designed by the author. Once an item had been answered, students were asked "how did you work that out?" The goal was to prompt verbal think-alouds, that is a retrospective report of "the thought processes involved in interpreting a question and arriving at an answer" (Presser et al., 2004, p. 4). Responses were captured on audiotape, with the interviewer revoicing unclear explanations to ensure the accuracy of the record. As a source of triangulation, extensive field notes of the students' responses and non-verbal signals were taken. Unstructured follow-up questions were used when, in the opinion of the interviewer, the explanation given did not clearly indicate how the student worked out their answer, or when more information might be enlightening. Students were not asked follow-up questions when, in the opinion of the interviewer, doing so might stress an already struggling student. After answering, some students changed their answer as a consequence of being asked to explain how they worked it out. In this situation the recorded answer was updated. At times, explaining their answer showed that a student was not really sure. In these cases the student's response was left as originally stated.

Responses were transcribed by the researcher, with each transcript being compared to the field notes. Data were then analysed using content analysis, that is, "a careful, detailed, systematic examination and interpretation of a particular body of material in an effort to identify patterns, themes, biases, and meanings" (Berg, 2007, pp. 303-304). Each response was analysed four ways.

1. Familiarity with the equipment (identified through students' oral comments and field notes describing issues with using the equipment).
2. Accuracy of response (a response of 380-395 grams was coded as correct, given the weight of the rock was visibly between the marks on the gauge).
3. The strategy explained by the student when asked to identify how they worked out their answer.
4. Sources of error when responding. (Coding based on Drake, 2011.

Figure 1a. Non zeroed scales with the rock.

Figure 1b. Set of scales (non-zeroed) without the rock.

Results

Reactions to this item indicated that all students recognised the instrument. No student asked about the dual gauges on the dial, and all worked with the gauge marked in kilograms. One student commented that she did not know what the little lines on the gauge meant as she usually used electric scales.

Table 1 shows the weights students gave for the rock. It shows that only one of the 17 students was close to providing an answer coded as correct; the other 16 students (94%) failed to ascertain if the scales were properly zeroed. Only two students lifted the rock from the pan and replaced it; one of these (Student 12) answered 400g. However, even for this student, a check of the instrument after the interview indicated that the needle was visibly between 375 and 400g, so a response of less than 400g would have been more accurate.

Interviewer (I):	How did you get 400 grams?
Student 12 (S12):	I took that off [lifts rock]. There are 4 things, little marks, in each space. Well, there's 3 of them but they equal 4 and each big line is 100. Divide 100 by 4 and you get 25, and if you take out the rock it is less than zero by 3 of the little marks so that's 75. Then put the rock back in [replaces rock] and get whatever it is and add on 75.

Table 1
How heavy is my rock?

Weight	30.1	300	300.2	310	315	320	325	328	400	Not sure
Frequency	1	3	1	1	1	3	4	1	1	1

If the issue of zeroing the scales is not considered, five of the 17 students (29%) could be coded as producing a correct answer (a number between 305 and 320g). However, when asked to explain his reasoning Student 7 (with an answer of 320g) signalled that he had miscounted the partitions "5 gaps, 100 grams, 1 gap past 300". Three of the other four indicated their response was a form of guess. For example, Student 9 who gave the answer 315 commented:

> …I knew it had gone past 300 so, sort of between one of the ummm marks on the way to 400. So I was sort of trying to figure out how it was even but I couldn't quite find out so I sort of guessed it was 15.

Overall, nine of the 17 students (53%) indicated that they were unsure how the 100g intervals had been partitioned. For example, all three students who originally answered 300 later indicated this form of problem, as Student 13 illustrates:

S(13):	300 and about, *ish*. I dunno.
I:	300ish?
S(13):	Yeah. Roughly 300.
I:	So when you are saying 300*ish*, what was it that made you think it was the *ish* part?
S(13):	Because there are 3 gaps in between so I don't know what those gaps are. I mean 3 lines.

Student explanations revealed that various strategies were used to obtain an answer. Four students (24%) used a form of skip counting. For example, Student 5 (who answered 320) gave a skip count but quickly realised her original count was incorrect, then was unable to find the correct count.

S(5):	I went 0,20, 40, 60, 80 Oh. That doesn't work. Oh I don't know. …
I:	So you are not quite sure?
S(5):	No.
	Oh! Does it go up in 30s? Oh no, cos that doesn't equal 400.

Six students (35%) were able to work out how the 100g intervals had been partitioned, but each seemed to use a different strategy to do so. As identified earlier, Student 12 divided 100 by 4, but by way of contrast Student 16 seemed to know that if there were three marks in the middle, each mark represented 25g: "*since there are only three in the middle that must be 25 for each line.*" Student 2 used half and half again: "*Well, in the middle it's just 350… Then I just halved 50 and that's 25*", while Student 14 used fractions "*the marks show there's four spaces so it must be quarters of the hundred so 325.*" Note that two of these strategies demonstrate an understanding of multiplicative thinking beyond applying the multiplication tables up to 10 × 10.

In some instances the strategy a student used was unclear. This is initially true of Student 4, who was able to identify how each 100g interval was partitioned.

S(4): You read those marks, here, and they are 25 grams.
I: And how did you work that out?
S(4): There are 4 spaces.
I: Yep. Keep going.
S(4): Each space is 25.
I: [After pause] So, some people go 25, 50, 75, 100 to check or did you go
S(4): [Interrupts] Yep.

In this case, follow-up prompts were important because the numbers 100, 4, and 25 can be linked in a variety of ways, from the skip count in 25s, to multiplication ($4 \times 25 = 100$), and recall of knowledge from having done similar problems previously. However, it was still not always possible to identify clearly the strategy a student used to answer the question. Overall, two students did not elaborate further, while several others were not able to explain their reasoning. For example, Student 6 (who answered 328) appeared to guess when he explained "*It's about 300, but I'm not sure how much more*."

Discussion

New Zealand students of a similar age to those answering the NAEP broken ruler problem performed poorly on a task of similar design. Only one of the 17 students provided a response that indicated an understanding of how to use analogue kitchen weighing scales. Sixteen students simply read the number that aligned with the needle and did not check the scales were properly zeroed. Lehrer (2003) reports a similar response by students working with broken rulers – reading the number that aligns with the end of an object. Such errors with the broken ruler problem have often been attributed to a lack of understanding of the nature of the unit and unit iteration (e.g., Irwin & Ell, 2002; Lehrer et al., 1998). However, it is more difficult to suggest that a lack of these understandings impacts upon weighing a rock on a set of kitchen scales. Perhaps not understanding the role of zero as the starting point of a measure is a better explanation given the error occurs in both tasks.

That 13 or 14 years-old students should find the kitchen scales problem so difficult is of concern but is predictable given that many students internationally have difficulty with the broken ruler problem. Three possible causes other than the small sample size can be identified: The first is the increasing use of digital weighing scales, yet the way the students responded to the scales indicated that such instruments were not unfamiliar. Second, measurement should be a practical topic, yet a perusal of text books often reveals book-based tasks, and reading a drawing of a weighing scale removes the practical consideration of checking that the instrument is properly zeroed. Third, the students have rarely been required to measure more accurately than to a labelled mark. Observations by the author of assessment items involving both measurement scales and graphs support this possibility.

That only six students could work out how each 100g interval was partitioned is also of concern. The responses of students who were able to do this indicate that the task requires a high level of number understanding, such as the ability to skip count in 25s or some form of multiplicative thinking based on knowledge beyond the recall of basic multiplication or division facts. One student, Student 7, was not able to identify correctly the number of pieces into which the interval had been partitioned, while other responses indicate a lack of familiarity with strategies (such as skip counting in 25s) that could be used to identify the value of each partition. These findings support the previous study by Drake (2011) which identified that learning to use linear scales is complex and that there are multiple potential sources of error. A new finding is that some students tend to read to the closest mark rather than trying for greater accuracy. The low success rate and multiple possible sources of error suggest that teachers would benefit from greater guidance about issues to address with students learning to use analogue weighing scales than is currently available from NZC and its supporting documentation (NZmaths, 2013), or can be found in books on effective practice (e.g., Bobis et al., 2013; Van de Walle et al., 2013; Zevenbergen et al., 2004).

In NZC, students are not expected to have memorised their basic multiplication facts until they have mastered CL3 (Year 6) (NZmaths, 2013), suggesting that moving beyond these facts could be expected of students in Year 7. The strategies used by the 'successful' students in this research tend to indicate that the task is beyond the number understanding that can be expected of New Zealand students in Year 6. A similar issue appears to exist with TAC, in which students are expected to have mastered their basic multiplication and division facts by the end of Year 4, the same year in which they are expected to be able to measure with a range of measuring instruments to the nearest graduation (ACARA, 2012).

Conclusion

In terms of learning about weight/mass, not learning to use a set of analogue scales is akin to learning to measure length without learning to use a ruler. Current practice assumes that learning to use such a set of scales "just happens" through exposure to weighing tasks. This is evidenced by the lack of emphasis on using analogue scales both in curricula and learning progressions, as well as the lack of guidance for teachers in curriculum support documents and books on effective practice. Results from this research and Drake (2011) suggest a more deliberate approach is needed, one that helps teachers unpack the skill set that seems to be necessary. In New Zealand, the NZC and National Standards information for measurement (NZmaths, 2013) would benefit from clarification as the requirements for measuring weight/mass do not seem to align with the development of number understanding. TAC (ACARA, 2012) appears to have a similar issue.

While results from a sample of 17 students must be viewed with caution, they do raise issues for further consideration. For example, students make the error of reading directly from the gauge when answering both the broken ruler and the kitchen scales problems so it may be useful to conduct and report research across attributes that use measurement scales to identify the reason. Current studies seem to focus within individual attributes.

The findings of this research, when considered alongside those of Drake (2011), also suggest that common models for developing measurement concepts do not recognise the complexity of the kitchen scales task. For example, the progressions in Zevenbergen et al. (2004) and Bobis et al. (2013) both have stages involving measurement with standard units which should fit the task. Yet while both the broken ruler problem and the kitchen scales problem involve measuring with standard units, students who were able to work out the value of a partition on the kitchen scales used thinking that was significantly in advance of the skills needed for the broken ruler problem. This lack of recognition of the demands of the kitchen scales problem may be due to the models being developed with geometric measurement (length, area, volume) in mind. Perhaps it is time that their higher stages are revisited to ensure they adequately address the measurement of other attributes.

References

ACARA. (2012). *The Australian curriculum*. Retrieved January 7, 2013, from http://www.australiancurriculum.edu.au

Berg, B. (2007). *Qualitative research methods for the social sciences* (6th ed.). Boston: Pearson Educational.

Bobis, J., Mulligan, J., & Lowrie, T. (2013). *Mathematics for children* (4th ed.). Frenchs Forest, NSW: Pearson.

Cheeseman, J., McDonough, A., & Clarke, D. (2011). Investigating children's understanding of the measurement of mass. In J. Clark, B. Kissane, J. Mousley, T. Spencer, & S. Thornton (Eds.), *Mathematics: Traditions and [new] practices (Proceedings of the 34th annual conference of the Mathematics Education Research Group of Australasia and the Australian Association of Mathematics Teachers)* (pp. 174-182). Adelaide: AAMT & MERGA.

Cheeseman, J., McDonough, A., & Ferguson, S. (2012). The effects of creating rich learning environments for children to measure mass. In J. Dindyal, L. Cheng, & S. Ng (Eds.), *Mathematics education: Expanding horizons: Proceedings of the 35th annual conference of the Mathematics Education Research Group of Australasia* (pp. 482-490). Adelaide: MERGA.

Drake, M. (2011). Factors impacting on students' ability to use linear scales. *The International Journal of Learning, 18*(1), 588-604.

Grouws, D. (Ed.). (1992). *Handbook of research on mathematics teaching and learning.* New York: Macmillan.

Irwin, K., & Ell, F. (2002). Visualising and the move from informal to formal linear measurement. In B. Barton, K. Irwin, M. P fannkuch, & M. Thomas (Eds.), *Mathematics education in the South Pacific*: *Proceedings of the 25th annual conference of the Mathematics Education Research Group of Australasia, Auckland* (pp. 358-365). Sydney: MERGA.

Kloosterman, P., Rutledge, Z., & Kenney, P. (2009). Exploring results of the NAEP: 1980s to the present: Results of the long-term trend assessment (LTT) for middle grades show positive advancement. *Mathematics Teaching in the Middle School, 14*(6), 357-365.

Lehrer, R. (2003). Developing understanding of measurement. In J. Kilpatrick, W. Martin, & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 179-192). Reston, VA: NCTM.

Lehrer, R., Jenkins, M., & Osana, H. (1998). Longitudinal study of children's reasoning about space and geometry. In R. Lehrer & D. Chazan (Eds.), *Designing learning environments for developing understanding of geometry and space.* Mahwah, NJ: Lawrence Erlbaum.

Lester, F. Jr. (Ed.). (2007). *Second handbook of research on mathematics teaching and learning: A project of the National Council of Teachers of Mathematics*. Charlotte, NC: Information Age.

MacDonald, A. (2010). Heavy thinking: Young children's theorising about mass. *Australian Primary Mathematics Classroom, 15*(4), 4-8.

MacDonald, A. (2011). Young children's representations of their developing measurement understanding. In J. Dindyal, L. Cheng, & S. Ng (Eds.), *Mathematics education: Expanding horizons: Proceedings of the 35th annual conference of the Mathematics Education Research Group of Australasia* (pp. 482-490). Adelaide: MERGA.

NZCER. (2012). *Progressive Achievement Tests (PATs).* Retrieved January 7, 2013 from http://www.nzcer.org.nz/tests/pats

NZmaths (2013). *NZC and Standards.* Retrieved January 7, 2013 from http://www.nzmaths.co.nz/nzc-and-standards

Presser, S., Rothgeb, J., Couper, M., Lessler, J., Martin, E., Martin, J., & Singer, E. (Eds.). (2004). *Methods for testing and evaluating survey questionnaires.* Hoboken, NJ: John Wiley.

Van de Walle, J., Karp, K., & Bay-Williams, J. (2013). *Elementary and middle school mathematics: Teaching developmentally* (8th ed.). New Jersey: Pearson.

Zevenbergen, R., Dole, S., & Wright, R. (2004). *Teaching mathematics in primary schools*. Crows Nest, NSW: Allen & Unwin.

.