Further Evidence of Conceptual Difficulties with Decimal Notation

Vicki Steinle
University of Melbourne
<v.steinle@unimelb.edu.au>

Kaye Stacey
University of Melbourne
<k.stacey@unimelb.edu.au>

Over 3000 students in Grades 4 to 10 completed about 10000 tests to identify and track their conceptions about decimal notation. This paper reports the incidence of two particular error-patterns on these tests and how this incidence changes over time, from cross-sectional and longitudinal viewpoints. At some time, approximately 3% of students answered almost all test items incorrectly and over 10% were unable to compare decimal numbers with the same tenths and hundredths digits. Explanations for these error-patterns are given.

It is well recognised that many students have difficulty understanding decimal notation. The reasons for this lie both in the nature of the mathematical and psychological aspects of the task and in the teaching they receive. Understanding decimal notation is a complex challenge, which draws on previous learning and fundamental metaphors of number and direction, both to advantage and disadvantage. As a consequence, there are a wide variety of erroneous ways in which students interpret decimal numbers, often referred to as decimal *misconceptions*. This paper reports results from a study of 3204 students in Grades 4 to 10, who completed 9856 tests between 1995 and 1999. Each item on the Decimal Comparison Test required the student to identify the larger number from a pair of decimal numbers (e.g. 0.8 and 0.75). This comparison task has been widely used because patterns of errors among carefully chosen sets of comparison items can reveal a great deal of information about how students interpret decimal notation. Every test is allocated a code according to the pattern of errors on the 30 items. The study has both a cross-sectional and longitudinal component, so that the prevalence of different test codes can be determined and the *paths* that students take, between the codes, over some years can be traced. In previous papers, (eg Stacey & Steinle, 1998), we have reported on the development of the 30-item Decimal Comparison Test; how inferences about students' thinking can be made reliably from the codes allocated to the tests; the incidence of expertise and other codes, as well as the paths through the codes which students commonly follow on the way to expertise (Steinle & Stacey, 1998; Stacey & Steinle, 1999a and 1999b). We have also examined more closely particular test codes and have provided some explanations in terms of how the students may be thinking about decimal notation, drawing on data from school students and teacher education students in several countries (Stacey, Helme, Steinle, Baturo, Irwin & Bana, 2001; Steinle & Stacey, 2001; Stacey, Helme & Steinle, 2001) and examined the effectiveness of targeted teaching (Helme & Stacey, 2000). This paper focuses on two test codes that are particularly intriguing; following up in this large sample, observations reported earlier in interviews and smaller scale studies. We report the incidence of these codes by grade level from both a cross-sectional and longitudinal point of view (explained below) and investigate the behaviour of such students on later tests to obtain additional insight into how they might be thinking.

B. Barton, K. C. Irwin, M. Pfannkuch, & M. O. J. Thomas (Eds.) *Mathematics Education in the South Pacific* (Proceedings of the 25th annual conference of the Mathematics Education Research Group of Australasia, Auckland, pp. 633-640). Sydney: MERGA. ©2002 MERGA Inc.

Methodology and Definitions

Volunteer secondary schools and some of their feeder primary schools in six geographical areas of Melbourne (labelled A to F) participated in testing in volunteer classes at six monthly intervals (although not every school tested students on every occasion). Table 1 shows the number of tests by grade level and school group. The sample is not random, but the six regions included very low, medium and high socio-economic areas. For this reason, the results of the school groups are reported separately, to provide an estimate the range of variation that might be in the whole population. Small differences between the basic numbers of the large sample reported in this paper and elsewhere are due to ongoing checking of the database. Larger differences from earlier published work are due to further data collection. Full details of the methodology can be found in Steinle et al (1998) and Stacey et al (1999b). Students took about 10 minutes to complete the test. Tests were coded by the researchers and results promptly returned to the teacher.

Table 1
Number of Tests by Grade and School Group

Grade	School Group						Total number of tests
	A	B	C	D	E	F	
4	39	29	178	90	0	0	336
5	158	66	355	263	121	0	963
6	232	140	454	415	206	0	1447
7	329	356	615	478	314	205	2297
8	345	327	459	293	222	456	2102
9	290	336	196	109	180	534	1645
10	278	249	0	0	50	489	1066
Total	1671	1503	2257	1648	1093	1684	9856

Definitions and Hypotheses

Twelve codes based on error-patterns have been identified from the 30-item test and these have been associated with different *ways of thinking* or *misconceptions* about decimal notation. Some of these ways of thinking are well established in other research (e.g., Hiebert & Wearne, 1986; Irwin, 1996; Putt, 1995; Sackur-Grisvard and Leonard, 1985) while others have been recently identified (e.g., Stacey et al, 1998). Space does not permit a full description of all of the misconceptions, but Figure 1 gives a brief description. For a particular test, the error-patterns on ten *core* test items results in a coarse code of A, L, S or U. These core test items are especially created to avoid particular difficulties for students- e.g. the digit zero is not used in the decimal columns. The codes are then further subdivided (e.g. into L1, L2 and L4) according to performance on a range of specialised items. The code A1 refers to expert behaviour on the comparison task, although this does not imply that the student has full understanding of decimals beyond this "simple" comparison task (e.g. they may not be able to place numbers on a number line).

This paper explores codes A2 and U2. Students testing as A2 (being one of the A codes) seem to be expert from all the core items, such as comparison of 0.8 with 0.75 (which all L students get wrong) and of 0.7 with 0.85 (which all S students get wrong). However, they are unable to select the larger number from items such as 8.41 and 8.41253. In Steinle et al (2001), we reported results of 315 students given a different version of the test that allowed them to select the larger number from each pair or to state that the numbers are equal. About 4% of students stated that numbers like 8.41 and 8.41253 are equal, indicating that they believe the decimal number system is discrete. Like dollars and cents, they believe the amounts after the hundredths (say) are not real but perhaps some sort of rounding error. Other students do not believe numbers such as 8.41 and 8.41253 are equal, but they make errors specifically on this item type. We hypothesise that they have no strategy for deciding which is larger, and so need to guess because they do not know (how) to add zeros and therefore their left-to-right digit-by-digit comparison strategy (which works in all other comparisons) fails when confronted with comparing the 2 in 8.41253 with a blank space in 8.41. In this paper we investigate the incidence of A2 in this larger and younger sample (noting that students who guess correctly on such items will be coded as A1 and hence not detected) and the later paths that these A2 students follow.

U2 is an intriguing code. These students select the smaller decimal to be the larger (nearly) completely consistently, so that they must be using a rule that is of comparable complexity to the correct rule. There are several possible explanations: they may be expert students who misread the instructions (in which case, we might expect A1 codes on their earlier and subsequent tests); they may be mischievous expert students having fun puzzling the researchers; alternatively, they may be correctly comparing digits according to place value, but then mysteriously reverse their otherwise correct answers. Amongst 103 Australian student teachers, Steinle et al (2001) found that 18% think that decimals which start as 0.xxx are less than zero (which they often insist is *not* the same as being negative), so it is possible that they also believe that just as -3 is greater than -4, so 0.3 is greater than 0.4 (although 1.3 is greater than less than 1.4). All of these explanations lead to the prediction that students who test as U2 are likely to have adjacent tests as A1, rather than other codes. This paper investigates this hypothesis.

Code	Description
A1	Task expert – successful on all comparisons (small allowance for carelessness)
A2, A3	Successful on core comparisons: errors in "unusual" comparisons
L1, L2, L4	Different ways of thinking which result in the general behaviour of choosing the longer decimal as the larger (e.g. may interpret decimal part as a whole number)
S1, S2, S3, S4	Different ways of thinking which result in the general behaviour of choosing the shorter decimal as the larger (e.g. may interpret decimal part as fraction)
U2	A maximum of 5 correct answers on entire test (i.e. at least 25 errors)
U1	Unclassified – not fitting into any other code

Figure 1. The classification system for decimal comparison test.

Results

Cross-sectional Incidence of U2 and A2

The incidence of any code can be considered from two points of view, which we call cross-sectional incidence and longitudinal incidence. The cross-sectional incidence measures the percentage of a given population who exhibit the behaviour at a particular point in time, whereas the longitudinal incidence measures the percentage of the population who exhibit the behaviour *at some time*. To illustrate with an analogy, if we consider the population of people at a shopping centre on one day, only a small percentage, say 25%, will be teenagers, but all the people in the shopping centre will be (or have been) teenagers at some time of their lives. So the cross-sectional incidence may be 25%, but the longitudinal incidence is nearly 100%. Both these measures are important. For decimal misconceptions, it is useful to know how many of a given school population (say the students in Grade 5) may be thinking in a certain way, but it is also useful to know what percentage of students think that way at some stage during their school career. Note that, typically, the longitudinal incidence will be greater than the cross-sectional incidence.

Table 2 presents the cross-sectional incidence for codes U2 and A2 by grade of the student at the time of testing. At this stage, no attempt is made to follow individual students; thus many students contribute several times to the calculations in this table as they get older. All of the cells of the table are percentages based on large numbers of tests. Overall 0.9% of the tests are in category U2, including an average of 0.2% up to Grade 7 and about 1.5% thereafter. Thus there is a general increasing percentage of U2 students with age, and Table 2 demonstrates that this is generally evident in the results of each school group. Paradoxically, the increase with age is not surprising, because a student needs to follow a complex incorrect thinking pattern very consistently to make so many errors. The probability of making at least 25 errors from 30 multiple-choice items by random guessing of all items is less than 0.5%, so the primary school results are consistent with this, but the secondary school incidence indicates a real effect.

Table 2 also shows the cross-sectional incidence of code A2. This has a higher incidence than U2, with an overall average of 4.2% of tests. Once again, incidence generally increases with age. In every school group, the primary school incidence is less than the secondary school incidence.

Table 2
Cross-Sectional Incidence (%) of U2 and A2 Tests by Grade and School Group

Percentage incidence of U2 by school group						Grade	Percentage incidence of A2 by school group					
A	B	C	D	E	F		A	B	C	D	E	F
0.0	0.0	0.0	0.0	-	-	4	0	0	0.6	2.2	-	-
0.6	0.0	0.0	0.0	0.0	-	5	0.6	3.0	2.8	3.8	0.0	-
0.0	0.0	0.4	0.0	0.0	-	6	3.9	4.3	2.9	3.4	3.4	-
0.0	1.1	0.3	0.0	0.3	0.5	7	6.1	4.5	3.4	5.2	4.1	4.4
2.6	1.5	1.1	1.7	1.8	1.1	8	3.2	4.3	4.6	3.8	4.5	6.8
0.7	1.5	2.6	0.9	0.5	1.7	9	4.1	3.6	4.6	2.8	2.8	8.2
2.5	1.6	-	-	0.0	2.0	10	4.7	3.6	-	-	2.0	6.7
1.1	1.2	0.6	0.4	0.5	1.5	Total	3.9	3.9	3.3	3.9	3.3	6.9

-indicates no testing in this cell

Longitudinal Incidence of U2 and A2

The longitudinal incidence is an analysis by student rather than by test. As noted above, although there are 9856 tests in this study, there are only 3204 students so that each student completed an average of 3.1 tests each. In particular, the number of tests per student varied from only one (772 students or 24% of the sample) to 7 tests for the 49 students present every time. Table 3 shows the percentage of students in each school group who at some stage tested as U2 or A2. This table demonstrates that U2 is relatively rare behaviour, exhibited at some stage by only about one student in forty (2.3%), but about one student in ten (10.9%) exhibits A2 behaviour at some stage. School group F, has the highest longitudinal incidence of both categories. School groups A and B are relatively high in U2 and relatively low in A2, whilst school groups D and E are the reverse.

Note that the longitudinal incidence, as with every measure from a sample, is affected by the structure of the sample. When one relies upon volunteer teachers and volunteer students in volunteer schools, the sampling will always be less than ideal. The longitudinal estimates here are probably an underestimate of the "real" incidence, because 24% of students have been tested only once. By testing them on only one occasion, the chance of them displaying the U2 or A2 behaviour is less than if they had been tested more times. Had there been more testing of the same students, then the number of U2 or A2 tests would have increased (or at least stayed constant), whilst the denominator (the number of students) would have been the same. This is just one instance of the difficulties of dealing with the influence of structure of the sample on results. In this case, however, it does not explain the higher longitudinal incidences of School Group F, which had a below average number of tests per student (2.8 compared to 3.1 for the whole sample). Both cross-sectional and longitudinal incidences are high, indicating greater prevalence of misconceptions.

Table 3
Longitudinal Incidences of U2 and A2 by School Group

School Group	Number of students	Average number of tests per student	Percentage of students with one test	Longitudinal incidence (percentage of students)	
				U2	A2
A	709	2.4	33	2.3	8.5
B	679	2.2	38	2.4	8.4
C	549	4.1	10	2.4	11.1
D	397	4.2	10	1.3	12.8
E	258	4.2	9	1.9	13.2
F	612	2.8	28	3.3	14.2
Overall	3204	3.1	24	2.3	10.9

What Happens to U2 and A2 Students on Their Next Tests?

Students who have a test which was coded as either U2 or A2 have been tracked to their next test. This subsequent test may be allocated the same code as before or a different code and this can be used to give some idea of the way in which students' conceptions of decimals develop. The numbers of students with exactly 1, 2, 3 or 4 occurrences of U2 are 64, 9, 2, 0, respectively, and of A2 are 300, 34, 14, 2, respectively. No student had more than 4 occurrences of either of these categories. The numbers of students drop rapidly down the list, but it is important to remember that once again this is due in part to the structure of the sample, as fewer students have undertaken the larger numbers of tests. It does not necessarily mean students have stopped this behaviour. This section examines the behaviour at the next test of students testing as U2 or A2. Both Tables 5 and 7 give the conditional probability of the code of the next test; sometimes adjacent grades have been combined, so that the number of students is reasonably large in each cell.

The Paths of U2 Students

There were 51 tests occasions where a student with a U2 test completed another test later. Note that this does not refer to 51 *students* as the 11 students with several occurrences of U2 contribute several times to this count. The probabilities of an A1 or U2 test following a U2 test, split by the grade of the U2 test, are given in Table 5. The probability of a younger U2 student next testing as A1 is 50%. This probability decreases with age while the probability of retesting as U2 increases. We expected "truly expert" mis-reading or mischievous U2 students to retest as A1 or U2. In fact, the U2 students are approximately equally likely to retest as A1, U2 or another code. The high probability of retesting in another code indicates that there is a significant group of U2 students who are not mischievous experts or mis-reading but instead have a misconception. Most likely they are reversing the order of the decimals because they think they are "less than zero". The high incidence of this amongst the teacher education students (Steinle et al, 2001) provides evidence to support the latter explanation.

Table 5
Conditional Probabilities of The Test Following U2

Grade of U2 test	Pr(A1/U2)	Pr(U2/U2)	Pr(other/U2)
Grades 5-8 (n=32)	0.50	0.16	0.34
Grades 9-10 (n=19)	0.32	0.37	0.31
Overall (n=51)	0.43	0.24	0.33

Looking at the test histories of students who often tested as U2 provides other detail. Table 6 gives the test histories of the eleven students in the whole sample who tested as U2 on more than one occasion. They come from all school groups, but are in the older grades. These students are very persistent in their U2 behaviour. On only 2 of the 14 possible occasions is a U2 test followed by any other code (on both occasions A1, expert). While Student 600705121 oscillates between U2 and A1 (suggesting misreading), the other students are either consistently mischievous or consistently using a wrong way of thinking. It is also very interesting that the last test of every one of these students (except Student 600803016) is U2. This suggests that they may well be stuck in the way of thinking that produces this U2 code.

Table 6
Test Histories of the Eleven Students with More Than One Occurrence of U2.

School Group	ID Number	Grade 5		Grade 6		Grade 7		Grade 8		Grade 9		Grade 10	
A	100705035						A1	U2	U2				
A	100905043										U2	U2	U2
B	210503018		A2				U2	U2					
B	200703081						A2				U2	U2	
C	310501006		L2		S2	S4	U2			U2			
D	400704093					A1	U1	U1	U2	U2			
E	500703045						U2	U2					
F	600703055						L4		A3	U2	U2		
F	600705121						U2	A1	U2				
F	600803007								A1		U2	U2	U2
F	600803016								A1		U2	U2	A1

The Paths of A2 Students

The conditional probabilities in Table 7 show that, following an A2 test, a student is most likely to become an expert (A1), although this probability decreases with age. This is a general phenomenon in this domain – the chance of becoming an expert decreases as non-expert students become older (see also Stacey et al, 1999b). The chance that they will retest with the same code (A2) increases with age, however, so that by Grades 9 and 10 it is as high as the chance of becoming an expert. The other columns of Table 7 demonstrate that there is some movement to other codes, and this is similar to what has been found for the sample as a whole: older students are unlikely to move to an L code, but the chance of becoming an S remains at about 10% (Stacey et al, 1999a and 1999b).

Table 7
Conditional Probabilities for the Test Following A2

Grade of A2 test	Test following the A2 test					
	A1	A2	A3	U1	Any L	Any S
Grades 4/5 (n=24)	0.67	0.04	0.00	0.13	0.13	0.04
Grade 6 (n=33)	0.67	0.09	0.03	0.06	0.06	0.09
Grade 7 (n=88)	0.55	0.16	0.02	0.08	0.03	0.16
Grade 8 (n=79)	0.54	0.24	0.06	0.09	0.01	0.05
Grades 9/10 (n=56)	0.32	0.29	0.16	0.14	0.00	0.09
Overall (n=280)	0.53	0.19	0.06	0.10	0.03	0.10

Tracing individual students with several A2 tests is also interesting, but there are too many to show all the individual paths; recall above that 50 students have two or more A2 tests. Thus, Table 8 shows the test histories of the students with at least 2 A2 tests, from the school group and cohort with the largest percentage of A2 tests. Only one student has become an expert at their final test, and 6 have remained A2. There are 2 students with 4 A2 tests. The A3 tests may indicate a student guessing inconsistently with the failing left-to-right digit-by-digit comparison strategy and the A1 tests may even arise from correct guessing. In this extreme example, it is likely that teaching which always emphasizes rounding answers to a fixed number of decimal places is responsible.

Table 8
Details of All Students from Cohort 1989 School Group F with At Least Two A2 Tests

ID Number	Total number of tests	Number of A2 tests	Grade 7	Grade 8	Grade 9	Grade 10
600703014	4	4	A2	A2	A2 A2	
600703027	5	2	S1	A2	U1 A2	U1
600703040	5	2	L4	A1	A1 A2	A2
600805003	4	3		L4	A2 A2	A2
600805016	3	3		A2	A2 A2	
600805020	3	3		A2	A2 A2	
600805047	4	3		A2	A2 A2	A1
600805061	4	4		A2	A2 A2	A2
600805072	4	2		S1	A2 A2	A3
600805085	3	3		A2	A2 A2	A3

Discussion and Conclusion

Examining the two case studies of A2 and U2 codes has demonstrated the value of the longitudinal data. The U2 group is small, comparable with chance in the primary school, but increasing through the secondary school. Whilst some of the U2 tests may come from mischievous experts, tracking their next tests does not indicate that most of them have mastered the decimal comparison task. It is likely that they are still struggling to integrate various ideas about negative numbers, decimals, fractions and place value. The A2 group is much larger, with 10% of students testing this way at some time. The existence of this group points to the need for teachers to treat the rounding process carefully and to continue to attend to basic place value issues (such as when zeros can be inserted into a numeral without changing its value) in the junior secondary years. Future studies will report other aspects of the longitudinal data.

References

Helme, S., & Stacey, K. (2000) Can minimal support for teachers make a difference to students' understanding of decimals? *Mathematics Teacher Education and Development. 2*, 105 – 120.

Hiebert, J., & Wearne, D. (1986). Procedures over concepts: The acquisition of decimal number knowledge. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 199-223). Hillsdale, NJ: Lawrence Erlbaum.

Irwin, K. (1996). Making Sense of Decimals. In J. Mulligan and M. Mitchelmore (Eds.). *Children's Number Learning: A Research Monograph of MERGA/AAMT.* (pp 243 – 257) Adelaide: AAMT.

Putt, I. J. (1995). Pre-service teachers' ordering of decimal nNumbers: When more is smaller and less is larger! *Focus on Learning Problems in Mathematics, 17* (3), 1-15.

Sackur-Grisvard, C., & Leonard, F. (1985). Intermediate cognitive organization in the process of learning a mathematical concept: The order of positive decimal numbers. *Cognition and Instruction, 2*, 157-174.

Stacey, K., & Steinle, V. (1998). Refining the classification of students' interpretations of decimal notation. *Hiroshima Journal of Mathematics Education, 6*, 1-21.

Stacey, K., & Steinle, V. (1999a). Understanding decimals: The path to expertise. In J. & K. Truran (Eds.), *Making the difference.* (Proceedings of the 22nd annual conference of the Mathematics Education Research Group of Australasia, Adelaide, pp. 446-453). Sydney: MERGA.

Stacey, K., & Steinle, V. (1999b). A longitudinal study of children's thinking about decimals: A preliminary analysis. In O. Zaslavsky (Ed.) *Proceedings of the 23rd Conference of the International Group for the Psychology of Mathematics Education. Vol. 4.* pp 233-240. Haifa, PME.

Stacey, K., Helme, S. & Steinle, V. (2001) Confusions between decimals, fractions and negative numbers: A consequence of the mirror as a conceptual metaphor in three different ways. In M. van den Heuvel-Panhuizen (Ed) *Proceedings of the 25th Conference of the International Group for the Psychology of Mathematics Education.* Vol 4. (pp 217 – 224). Utrecht: PME.

Stacey, K., Helme, S., Steinle, V., Baturo, A., Irwin, K., & Bana, J. (2001) Preservice teachers' knowledge of difficulties in decimal numeration. *Journal of Mathematics Teacher Education.* 4 (3), 205-225.

Steinle, V., & Stacey, K. (1998). The incidence of misconceptions of decimal notation amongst students in grades 5 to 10. In C. Kanes, M. Goos, & E. Warren (Eds.), *Teaching mathematics in new times* (Proceedings of the 21st annual conference of the Mathematics Education Research Group of Australasia, Brisbane, pp. 548-555). Sydney: MERGA.

Steinle, V., & Stacey, K. (2001). Visible and invisible zeros: sources of confusion in decimal notation . In J. Bobis, B. Perry, M. Mitchelmore (Eds.) *Numeracy and Beyond.* (Proceedings of the 24th Annual Conference of the Mathematics Education Research Group of Australasia, Sydney, vol 2, pp. 434-441). Sydney: MERGA.