GRADE-RELATED TRENDS IN THE PREVALENCE AND PERSISTENCE OF DECIMAL MISCONCEPTIONS

Vicki Steinle & Kaye Stacey1
University of Melbourne

Over a period of about 3 years, 3204 students in Grades 4 to 10 completed 9862 tests to identify and track their interpretation of decimal notation. Analysis of the longitudinal data demonstrates that different misconceptions persist among students to different degrees and in different patterns across the grades. Estimating the prevalence of misconceptions is complex due to the nature of longitudinal data. Best estimates are provided of grade prevalences and the proportions of students affected during schooling.

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 requires the co-ordination of many ideas and draws on previous learning and fundamental metaphors of number and direction both to advantage and disadvantage (Stacey, Helme & Steinle, 2001). 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 that has both cross-sectional and longitudinal components, so that the prevalence of different misconceptions can be determined and the paths that students take between these misconceptions can be traced over some years. This quantitative work is set in a context where we have also examined more closely particular misconceptions and have provided some explanations in terms of how the students may be thinking about decimal notation, drawing on interview and written data from school students and teacher education students (Steinle & Stacey, 2001; Stacey, Helme & Steinle, 2001) and examined the effectiveness of targeted teaching, although these aspects are not discussed in this paper.

This paper reports two aspects of the longitudinal data. Firstly, we report the prevalence of certain groups of misconceptions by grade level and show that there are interesting variations amongst the patterns in how these misconceptions persist amongst students of different ages. We also discuss how an estimate of the lifetime prevalence of these misconceptions might be obtained from the longitudinal data. In previous papers, we have made preliminary reports on the prevalence of expertise and misconceptions, as well as the paths that students commonly follow on the way to expertise (Steinle & Stacey, 1998, Stacey & Steinle, 1999a and 1999b). The present paper extends these analyses by using a larger dataset collected over a longer timeframe and by using more sophisticated analyses. We intend that the results will be useful to researchers working on students' understanding of decimals and to those interested more generally in the aetiology of students' ideas. We also intend that the discussion of the technical difficulties of analysis of longitudinal data will be of broad interest.

---

[1] This study was funded by the Australian Research Council, under a grant to Kaye Stacey and Liz Sonenberg. We wish to thank the teachers and students who provided data for the study.

DATA COLLECTION

Sample and Procedures

The Decimal Comparison Test (described below) is used to classify students' thinking about decimal notation. The test, which takes less than 10 minutes, was administered by volunteer class teachers to intact classes of students in 13 volunteer schools in 6 suburban regions of a large Australian city. Researchers analysed the tests, allocating a code that indicates the students' thinking, and returned all results to the classroom teacher with explanations. By including both secondary schools and their "feeder" primary schools, not only could students be tracked from grade to grade within a school, but also across the primary-secondary transition (Grade 6 to Grade 7). The results of individual students were traced throughout the study and the longitudinal results were analysed with the computer program "STATA".

In total, there were 3204 students in Grades 4 to 10, who completed 9862 tests between 1995 and 1999. Schools were requested to test students at six monthly intervals but not all teachers volunteered and others delayed testing for various reasons. These procedures have important effects on the data which need to be managed: there are a large number of students who are tested only once (or whose later tests could not be confidently identified), many students have broken sequences of tests due to their own absence from school on the day of the testing or their class not having been tested in a given semester. Overall, however, a very rich data set was collected: the maximum number of tests completed by a student was 7 (49 students), the average number of tests per student was 3.1, and the average time between tests was 8 months, offering an unprecedented opportunity to see the development of students' ideas on one topic over time. The regions represented all socio-economic groups, but the sample is only representative and voluntary, not randomly selected.

Decimal Misconceptions and their Diagnosis

It is well established that the task of choosing the larger of a pair of decimal numbers (or ordering a larger set) is very useful for revealing how a student interprets decimal notation. Some students consistently choose the longer of the pair (e.g. they will say 4.63 is a larger number than 4.8) and others will choose the shorter (e.g. they will say 5.62 is a larger number than 5.736). We label these behaviours as Longer-is-larger (L) and Shorter-is-larger (S). There are many different patterns of thinking which lead to these behaviours, which have been explored by Fuglestad (1998), Resnick et al (1989), Sackur-Grisvard and Leonard (1985), Stacey and Steinle (1998), Swan (1983) and others. Consequently, although L and S are sometimes referred to in the literature as misconceptions, they are actually both behaviours arising from clusters of misconceptions, which space precludes us from discussing here.

The behaviours L and S were identified on the basis of the pattern of responses to the Decimal Comparison Test, which has been based on similar tests reported in the literature. This test consists of one sheet of paper with 30 pairs of decimal numbers (referred to as 30 items) with the instruction: *For each pair of decimal numbers, circle the larger*. The thinking pattern of the student is diagnosed by a detailed analysis of the pattern of responses to the 30 items, and the test is allocated one of 11 "fine codes". One of these (A1) represents expertise on this task with the student being correct on all item types, nine represent misconceptions which are reasonably well understood and one is for test responses with unclassified patterns. The patterns of thinking corresponding to these "fine codes" are grouped according to whether they exhibit L or S or other behaviour as explained below. Because of space limitations, the analysis presented in this paper is at the level of the L and S groups of misconceptions (the "coarse codes"), not the finer classification.

Table 1 provides the ten core items from the Decimal Comparison Test that serve to identify the L and S students. For convenience, the larger decimal is always given first in Table 1, but not on the test. The code A (Apparent-expert) indicates that a student has answered correctly on these ten core items, and the code U (Unclassified) is used on papers that do not fit the A, L or S response patterns. To allow for "careless errors", the A, L or S code is allocated even if there is one deviation from the expected pattern. So, for example, a student with 0 or 1 correct on Type 1 and 4 or 5 correct on Type 2 would be classified as L. The ten core items have been carefully chosen as "normal" items to avoid the complications, many of which arise from visible and invisible zeros (Steinle and Stacey, 2001), which distinguish between the fine classifications. Note again that A, L, S and U are behaviours, rather than particular ways of thinking. A student could be classified A, for example, by being truly an expert or by simply selecting the number with the larger digit in the tenths column as the larger, in which case they may not be able to choose the larger from 7.942 and 7.94.

Behaviours	Type 1 items					Type 2 items				
	4.8 4.63	0.5 0.36	0.8 0.75	0.37 0.216	3.92 3.4813	5.736 5.62	0.75 0.5	0.426 0.3	2.8325 2.516	7.942 7.63
Longer-is-larger (L)	X	X	X	X	X	√	√	√	√	√
Shorter-is-larger (S)	√	√	√	√	√	X	X	X	X	X
Apparent-expert (A)	√	√	√	√	√	√	√	√	√	√
Unclassified (U)	none of the above									

Table 1: Expected responses by students with particular behaviours

Figures 1 and 2 illustrate how students' responses fit the A, L, S groups. Figure 1 shows the 36 possible scores on Type 1 and Types 2 items if students answered randomly with probability being correct on any item of 0.5. The majority of students would, if guessing, have scores of 2 or 3 on both Types 1 and 2. (Note that the (5,5) corner corresponding to A is hidden in this graph, because the expected numbers are very small.) A choice other than one half for the probability of being correct or unequal probabilities for Type 1 and Type 2 will shift the "mountain" of probabilities from the centre, but not change the overall single peaked mountain shape. The contrasting data in Figure 2 comes from the 3531 tests completed in 1997 by students from Grades 5 to 10. There are three clear peaks. A considerable number of students answer all 10 items correctly hence the column at the back corner (5,5) in Figure 2 is considerably taller than all other columns (and has been truncated to 800 from 1664). The very low columns in the centre of the graph demonstrate that students, on the whole, are not choosing randomly on this test. Secondly, the very tall columns at the side corners, corresponding to L and S behaviours, confirm the validity of the L and S constructs. In fact, if only the 1867 papers with one or more errors in these 10 items are considered, then 59% are coded as L or S. Hence, Figure 2 shows that the sets of Type 1 and Type 2 items are internally homogeneous but different to each other and hence that the classifications to L, S and A are indeed meaningful.

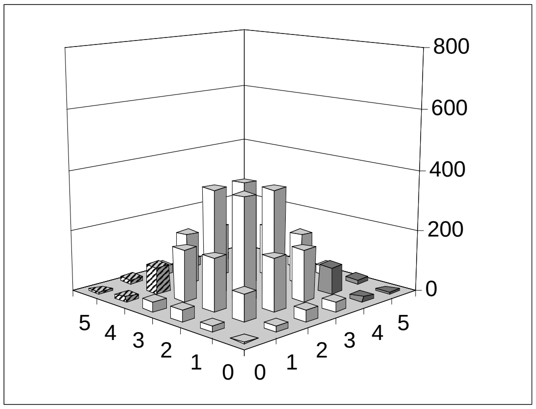

Figure 1: Distribution of A, L, S, U assuming answers selected at random.

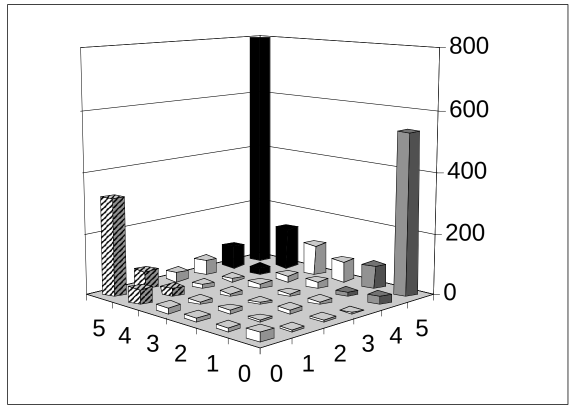

Figure 2. Actual distribution based on data from 1997 (Truncated at (5,5) corner).

(Key: A black, L grey, S striped, U unshaded)

PREVALENCE ACROSS THE GRADES

An important use of the longitudinal data is to report on how many students are likely to be affected by each of the decimal misconceptions. This simple notion requires careful thought. One useful measure is the percentage of students who demonstrate the behaviour at any given grade and another is the percentage of students who are likely to demonstrate the behaviour at some time in their schooling. In defining useful measures, we draw on epidemiology (Hope et al, 1998, p 791) where the *point prevalence* of a disease is defined as the number of cases at a point in time divided by the population at risk and the *period prevalence* is the number of cases at any time during the study period, divided by the population at risk. This concept moves closer to the notion of lifetime risk for a disease.

The measure which corresponds to point prevalence in our context is “grade prevalence”, which is intended to indicate the percentage of students in a given grade who are likely to exhibit the behaviour. The naïve measure for this would be the percentage of tests from a given grade receiving a certain code. However, we have used only the 3204 first tests that students have done, so that individuals are not included twice (especially in the data from one grade) and to avoid the repeat-test effect. Data shows that students who undertake the test repeatedly improve more than others, most probably because teachers who take the effort to test their students are more likely to give decimal notation an emphasis in their teaching. Additionally, these students probably have above average attendance at school. The grade prevalence results are affected by the constitution of our sample, as well as the strictness of the rules for classifying behaviour (as mentioned above, we permit one error per type, for example). The data set is of sufficient size that over 250 first tests have been available at each grade.

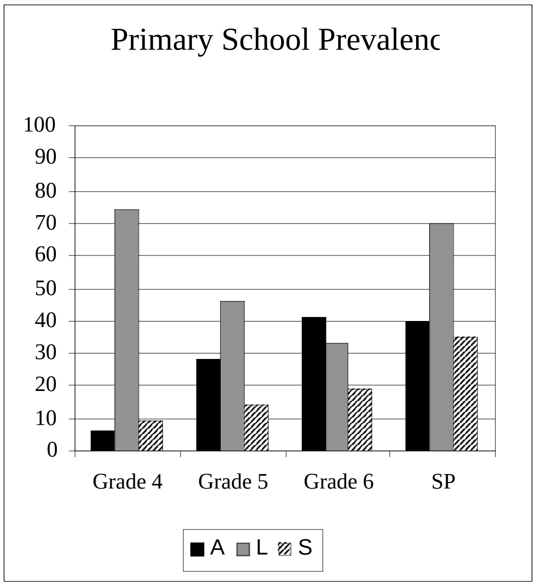

Figure 3. Primary school grade and schooling prevalences for A, L and S.

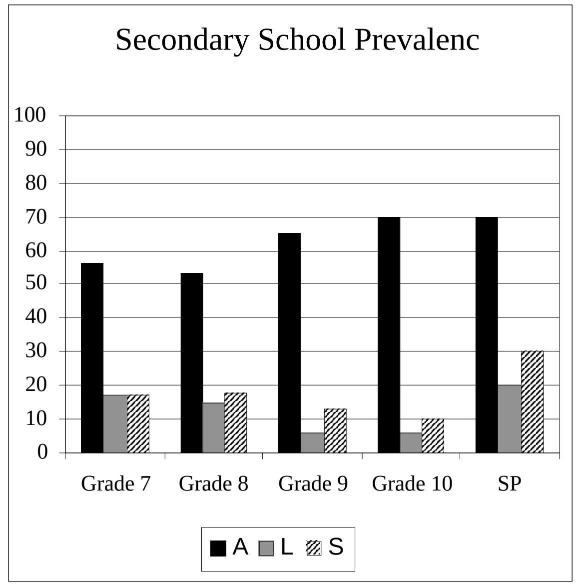

Figure 4. Secondary school grade and schooling prevalences for A, L and S.

The grade prevalences for Grades 4 to 10 are shown in Figures 3 and 4 (along with the schooling prevalence (SP), which is discussed below). The grade prevalence of L drops steadily (often nearly halving across successive grades), indicating that L is principally a misconception of younger students. In contrast, the grade prevalence of S stays reasonably constant, with 10 to 20% of students in the middle grade levels exhibiting this behaviour. In the next section, we will show that there are also very different patterns of students' movements into and out of these misconception groups. The percentage of students coded as A increases, as would be expected. However, it falls significantly short of 100%, rising only to 70% at Grade 10, indicating that problems with understanding decimals continue beyond the compulsory years of schooling for a significant group of students. Data not shown in Figure 3 shows that at all grade levels, about 12% of the tests coded A are not A1. This indicates that about 12% of students who can deal with the Type 1 and 2 items, cannot deal with those with zeros after the point, the same digits in the tenths column or other complicating features. They may have no real understanding of decimal notation, despite being coded as A.

PERSISTENCE OF MISCONCEPTIONS ACROSS THE GRADES

The calculation of grade prevalence above uses only cross-sectional data. In this section, we use the longitudinal data to describe how persistent each misconception code is. Figure 5 shows the probabilities of a student retesting in the same code on the next test that they complete. This data is based on an average of 332 students per data point with a maximum of 1202 students (Grade 7 A to A) and only two data points based on less than 100 students (Grades 9/10 L and S).

Figure 5 shows that students testing as A have a high probability (around 90%) of retesting as A on the next test that they do, regardless of their initial grade level. Although not shown, the retesting probabilities for A1 are similar. The probabilities for retesting as an L reflect the decreasing numbers of students in this code, and the fact that the L code is populated mainly by young students (Figures 3 & 4). The data for S demonstrates that students about Grade 8 (based on 195 students) are especially likely to be retained in this group (for explanation see Stacey, Helme, & Steinle, 2001). The different trends at Grades 9 and 10 may be due to the smaller sample or it may reflect the fact that students at this level who are still in the L or S codes have special learning difficulties.

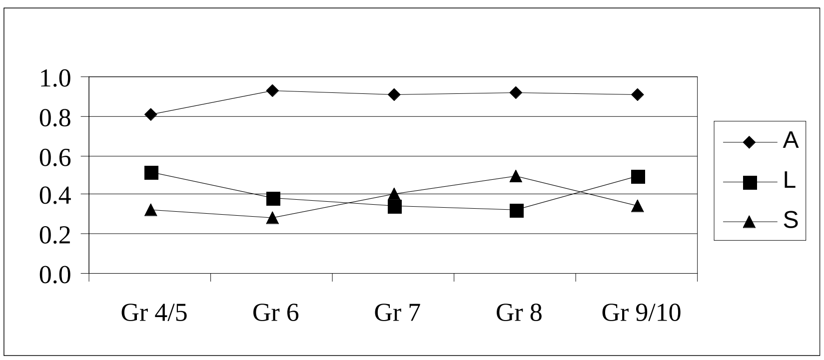

Figure 5. Probability of retesting in the same code, by grade of initial test.

Whereas the behaviours of groups L and S seem similar in Figure 5, Table 2 reveals the differences. The probabilities in the second row reflect the data in Figure 5, combined across grade levels. However, the first row gives the probability that a student testing in a given code had tested in that code on their immediately prior test. (Note that the data sets are necessarily not quite the same: row 1 is based on second to final tests, row 2 is based on first to penultimate tests etc). Whereas only one quarter of the L students came from another code on the preceding test, on average, half of the S students were previously in another code (mainly L or U). To oversimplify, many students begin in the L group at Grade 4/5 and then they move out at some stage, whereas code S recruits across the middle years of schooling, with students moving in and out with reasonable probability. This data is affected by the length of the study, which may have concluded before the student returned to the code and hence it provides a lower limit on the real data. This phenomenon is central to discussion of the schooling prevalence in the next section.

	A	A1	L	S	U

Probability of an immediately previous test of same code	0.77	0.74	0.73	0.48	0.71
Probability of an immediately following test of same code	0.91	0.89	0.44	0.38	0.71

Table 2. Probabilities of a student having previous and following tests in same code

SCHOOLING PREVALENCE

In addition to the grade prevalences, it is also of interest to know how many students are affected by the misconception during their school lives (or at least between Grades 4 and 10). This quantity, which we call *schooling prevalence*, is analogous to the lifetime risk for a disease. A first proposal is to calculate the percentage of the 3204 students who were allocated a given code on any test that they completed during the course of the study. This would be reasonable if all 3204 students had been completely tracked from Grades 4 to 10. However, it is inappropriate for our sample, because students enter and leave the study at all stages from Grade 4 to Grade 10. Students who were in Grade 8 when they entered the study, for example, may never test as L even though they were L beforehand. Calculating the schooling prevalence therefore requires care. A first decision is that the data better supports calculation of primary schooling prevalence and secondary schooling prevalence separately. The 3 years of the longitudinal data nearly matches the complete timespan of interest (Grades 4 – 6 and Grades 7 – 10). Happily, this information is useful, matching the separate professional concerns of primary and secondary teachers.

Table 3 shows the schooling prevalences based on the sample of primary (secondary) students who have completed at least 4 tests in the primary (secondary) school. The calculated prevalence is the number of students who have at least one test in the code, as a percentage of the restricted sample. No student is in both samples. Table 3 data for both L and S are compatible with the grade prevalences (Figures 3 and 4) and the persistence patterns. Comparison with the grade prevalences in Figures 3 and 4 shows that the prevalence for A is very much higher. This is because of the repeated test effect discussed above. Given this and the pattern of the persistence data (Figure 5), our best estimate for the schooling prevalences of A (and A1) is obtained from the maximum of the grade prevalences, say 40% for primary and 70% for secondary. Round numbers are used to emphasize that these are estimates. These best estimates have been graphed as SP in Figures 3 and 4.

	A	A1	L	S	U
Calculated primary schooling prevalence (n=333) *(Best estimate bracketed)*	68% *(40%)*	60% *(30%)*	71% *(70%)*	35% *(35%)*	44%
Calculated secondary schooling prevalence (n=763) *(Best estimate bracketed)*	89% *(70%)*	83% *(60%)*	20% *(20%)*	27% *(30%)*	29%

Table 3: Schooling prevalence calculated from restricted sample (students with 4+ tests in primary (secondary) school) and overall best estimates (*brackets*).

CONCLUSION

This paper has reported best estimates for the percentages of students in each of the Grades 4 to 10 likely to exhibit certain types of behaviour (L, S and A) related to misconceptions about decimal numeration. The different behaviours exhibit different prevalence, persistence and patterns across grades. Using the longitudinal data and knowledge of the prevalence and persistence, estimates have been made of the percentage of students who will exhibit the given behaviours at some time during their primary schooling (effectively after Grade 4 for this topic) and secondary schooling. These results are summarized in Figures 3 and 4. The results show the importance of addressing this decimal numeration explicitly. More generally, the paper has demonstrated some of the challenges and potential of dealing with longitudinal data and has provided a case study of the development of students' ideas on a topic across the grade levels.

References:

Fuglestad, A. B. (1998). Computer support for diagnostic teaching: A case of decimal numbers. *Nordic Studies in Mathematics Education (Nordisk Matematikk Didaktikk), 6*(3-4), 25-50.

Hope, R.A., Longmore, J. M., McManus, S. K., & Wood-Allum, C. A. (1998). *Oxford handbook of clinical medicine*. 4th Edition. Oxford: Oxford University Press.

Resnick, L. B., Nesher, P., Leonard, F., Magone, M., Omanson, S., & Peled, I. (1989). Conceptual bases of arithmetic errors: The case of decimal fractions. *Journal for Research in Mathematics Education, 20*(1), 8-27.

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. (1999a). Understanding decimals: The path to expertise. In J. Truran & K. Truran (Eds.), *Making the difference. Proceedings of the 22nd Annual Conference of the Mathematics Education Research Group of Australasia* (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.

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* (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* (Vol. 2, pp. 434-441). Sydney: MERGA.

Swan, M. (1983). *Teaching decimal place value: A comparative study of conflict and positive only approaches*. Nottingham: Shell Centre for Mathematical Education.