Steinle, V. & Stacey, K. (2003) Exploring the right, probing questions to uncover decimal misconceptions. In Leicha Bragg, Coral Campbell, Georgina Herbert, Judith Mousley (Eds.), *Merino. Mathematics Education Research: Innovation, Networking, Opportunity* (Proceedings of the 26th Annual Conference of the Mathematics Education Research Group of Australasia, Vol. 2, pp. 634-641), MERGA , Deakin University, Geelong

Exploring the right, probing questions to uncover decimal misconceptions

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

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

This paper analyses in detail 3531 tests involving 24 decimal comparison items, to establish the homogeneity of items and to identify items that throw additional light on students' thinking about decimals. The analysis confirms the dominance of ways of thinking previously identified, but it also provides evidence for new variations and combinations. Two features (involving only hundredths and being greater than one) explain variation in facility within the item sets. The features suggest particular teaching to target student needs.

> It is only by asking the right, probing questions that we discover deep misconceptions, and only by knowing which misconceptions are likely do we know which questions are worth asking (Swan, 1983, p 65)

The above quote from Swan reveals the duality between the questions that we ask and the student conceptions and misconceptions that we can find. In order to develop a test or interview to uncover students' thinking about a mathematical topic, the ways in which students think need to be known, but the questions need to be known before the thinking can be probed. Test or interview development is therefore a step-by-step procedure of constant refinement building on previous results. In Stacey and Steinle (1998), we analysed the results of 1853 students who had completed a test (which we call DCT1) of 25 decimal comparison items, based on the test of Resnick, Nesher, Leonard, Magone, Omanson, & Peled (1989). Pairs of decimal numbers are presented and the student instructed to circle the larger number in each pair. The test results were initially used to identify students exhibiting expertise or one of three misconceptions, and then by carefully analysing the results, further misconceptions were identified. In addition, some items that had previously been thought to present precisely the same task to students (i.e. they belong to the same 'item type') were found to behave differently. (We define an 'item type' as a set of comparison items that should be answered identically, either all correct or incorrect, by any given student who is consistently applying their own way of thinking about decimals.) As a consequence, new item types were discovered along with a refined understanding of students' misconceptions. This work resulted in the creation of a new version of the decimal comparison test, called DCT2, which has better diagnostic power. This paper undertakes for DCT2 the type of analysis undertaken for DCT1 in the 1998 paper. The aim is to see if some item types should be further subdivided and thus if further student misconceptions are evident in responses to the new test.

Diagnosing Decimal Misconceptions

Various *ways of thinking* about decimal numeration (see Appendix) have been detected in interviews and written work by other researchers, (see for example, Resnick *et al,* 1989), and the authors, (for example Stacey, Helme & Steinle, 2001). These *ways of thinking* are grouped and coded according to the logical consequences of this thinking on certain canonical items. Students with any of the ways of thinking labelled as L1, L2 or L4 will tend to choose the decimal with more digits after the decimal point as the larger number but these groups behave differently when the item has special features (often zeros). Similarly, students coded as S1, S3 and S5 have a way of thinking that leads them to often choose the number with fewer digits as the larger. Note that some *ways of thinking* cannot be separated by DCT1 or DCT2 because students are not able to indicate if they think the decimals in a pair are equal. In particular, the Zero Test (referred to as DCT0 in Steinle & Stacey, 2001) is used to separate *reciprocal thinking* from *negative thinking* (both S3 in this paper).

The classification scheme in Table 1 indicates how codes (intended to indicate the students' ways of thinking) are allocated to completed tests on the basis of the scores on each item type. For example, a test is coded as A1 (*task expert*) when the scores on Types 1 to 6 are all High. A test coded as A2 follows the pattern (Hi, Hi, Hi, Lo, Hi, Hi). Any other test that starts as (Hi, Hi, ….) but does not completely match the pattern for A1 or A2 is coded as A3, hence the "else" in Table 1. Similarly, the codes L1 to S5 (see Table 1) are assigned on responses to item types 1 to 6. Any other test is coded as U (unclassified). A small subset of these has almost every item was incorrect (code U2) and the remaining U tests are coded as U1.

Table 1
Classification scheme for allocating codes to tests on performance on groups of items

Item Type (number)			Codes											
			A1	A2	A3	L1	L2	L3	L4	S1	S3	S5	U1	U2
Core	1	(5)	Hi	Hi	Hi	Lo	Lo	Lo	Lo	Hi	Hi	Hi	Else and min 6 correct	Else and max 5 correct
	2	(5)	Hi	Hi	Hi	Hi	Hi	Hi	Hi	Lo	Lo	Lo		
Non-Core	3	(4)	Hi	Hi	Else	Lo	Hi	Lo	Else	Hi	Hi	Else		
	4	(4)	Hi	Lo		Hi	Hi	Hi		Lo	Lo			
	5	(3)	Hi	Hi		Hi	Hi	Hi		Hi	Lo			
	6	(3)	Hi	Hi		Hi	Hi	Lo		Hi	Lo			

Hi=High (at most one error in the set of items for that type) and Lo=Low (at most one item correct in set)

Evaluation of the Homogeneity of the Item Types

During 1997, the test DCT2 was used to gather data on students from Grades 5 to 10 and the responses to each item were recorded. Of the 3531 tests in total, exactly 1200 (34%) were correct on every item and coded as A1. These tests contribute no more to this paper, since it is an analysis of error patterns. The remaining 2331 tests had at least one error. Table 2 demonstrates the strong tendency of students to score either Low or High on an item type. In each row, the most frequent score is 'all-correct', and the second most frequent is 'all-wrong'.

The first analysis considers the consistency of responses to each item type by the students within a given code. Table 2 indicates that a score of High or Low on any item type is allowed to vary by 1 from the extremes of all-correct or all-wrong, in order that one "careless error" does not automatically result in the whole test being unclassified. In the following discussion, we will use variations in this careless error rate to search simultaneously for further systematic errors and for items that do not belong to their assigned types. Some of the careless errors will not, in fact, be "careless".

Table 2

Percentage distribution of scores of each item type for 2331 tests

Type of item (number)			Score on item type					
			0	1	2	3	4	5
Core	1	(5)	26	6	6	8	11	43
	2	(5)	19	5	4	5	8	59
Non-Core	3	(4)	18	7	7	8	60	-
	4	(4)	30	6	5	7	53	-
	5	(3)	13	4	6	77	-	-
	6	(3)	14	6	6	74	-	-

Low scores=light shading, High scores= dark shading

Table 3 indicates where this search should be carried out. It gives the proportion of tests (in each code) that follow the exact predictions within an item type. Consider tests assigned the code L1. Table 1 indicates that the score for Type 1 must be Low. Of the 428 L1 tests, 92% matched the prediction of 0 on Type 1 and thus 8% of L1 tests scored 1 out of 5. For Type 2 items, 97% matched the prediction of 5 out of 5. Only 82% of the L1 tests, however, matched the prediction (of 0) on Type 3. This indicates that there may be a Type 3 item that does not belong as it elicits different responses to the other items in the group. The lowest proportions in each column of Table 3 are therefore bolded to highlight points for investigation. (Note that no predictions are made by the coding system for A3, L4 and S5 for Types 3 to 6). One of the first features to note in Table 3 is that the bold entries do not all appear in one row as might be expected. This highlights the fact that students within the different codes react differently to features of various items.

Table 3

Proportion of tests that follow exact predictions within each item type, by code

Item Type		A1	A2	A3	L1	L2	L4	S1	S3	S5
		n=443	n=127	n=132	n=428	n=122	n=99	n=106	n=225	n=121
Core	1	**0.76**	0.90	**0.67**	0.92	**0.75**	**0.75**	0.81	0.88	0.78
	2	0.88	0.79	0.85	0.97	0.95	0.80	**0.65**	0.93	**0.74**
Non-Core	3	0.90	0.95		**0.82**	**0.76**		0.93	0.94	
	4	0.84	**0.72**		0.97	0.92		0.92	0.99	
	5	0.92	0.94		0.99	0.98		0.86	0.82	
	6	0.99	0.94		0.94	0.99		**0.74**	0.87	

Table 4 contains details of facility (percentage correct) by code for the 24 items of Types 1 to 6. The items appear as rows and are ranked by overall facility within each of the types. The overall facilities on these 24 items range from a maximum of 83% (Q4, 1.85/1.84, Type 5) to a minimum of 56% (Q9, 0.37/0.216, Type 1). Data is given in Table 4 to the nearest percentage point for space considerations, but calculations throughout have been done with accurate figures.

Steinle, V. & Stacey, K. (2003) Exploring the right, probing questions to uncover decimal misconceptions. In Leicha Bragg, Coral Campbell, Georgina Herbert, Judith Mousley (Eds.), *Merino. Mathematics Education Research: Innovation, Networking, Opportunity* (Proceedings of the 26th Annual Conference of the Mathematics Education Research Group of Australasia, Vol. 2, pp. 634-641), MERGA , Deakin University, Geelong

Table 4
Facility on 24 test items (used in the classification scheme) by code

Description			Decimal Pair		Sub-type	Overall	A1	A2	A3	L1	L2	L4	S1	S3	S5	U1	U2
			larger	smaller		n=2331	n=443	n=127	n=132	n=428	n=122	n=99	n=106	n=225	n=121	n=501	n=27
Core items	Type1	Q6	4.8	4.63	α	64	97	99	96	4	14	11	97	95	93	67	11
		Q7	0.5	0.36	β	61	97	98	95	2	2	4	95	98	96	57	15
		Q8	0.8	0.75	β	61	98	98	95	1	0	2	95	97	96	57	11
		Q10	3.92	3.4813	γ	60	94	98	88	1	8	8	98	98	96	53	7
		Q9	0.37	0.216	δ	56	91	97	92	0	1	0	95	100	97	42	4
	Type2	Q17	0.75	0.5	β	72	96	94	95	98	99	97	24	4	13	63	0
		Q20	7.942	7.63	γ	72	100	96	96	100	100	92	2	1	7	64	4
		Q19	2.8325	2.516	γ	71	98	98	97	99	98	95	3	0	1	63	0
		Q16	5.736	5.62	γ	71	95	93	98	99	98	98	4	1	4	65	4
		Q18	0.426	0.3	δ	70	99	97	98	100	100	98	3	0	2	59	0
Non-Core items	Type3	Q12	4.7	4.08	α	73	96	98	72	11	95	55	98	96	90	82	19
		Q15	8.514	8.052573	γ	71	100	98	64	3	98	55	99	100	88	76	7
		Q14	2.621	2.0687986	γ	70	97	99	64	2	92	43	97	100	88	76	11
		Q13	3.72	3.073	γ	69	97	99	64	2	92	44	99	99	92	73	7
	Type4	Q22	17.353	17.35	γ	63	98	9	73	99	100	76	4	1	17	57	7
		Q24	3.2618	3.26	γ	62	96	6	68	100	99	75	2	0	13	55	4
		Q23	8.24563	8.245	γ	61	96	4	64	100	98	67	1	0	17	53	11
		Q21	4.4502	4.45	γ	60	93	8	58	99	94	75	1	0	16	54	0
	Type5	Q4	1.85	1.84	α	83	99	100	82	100	99	96	97	4	76	84	0
		Q5	3.76	3.71	α	83	98	98	83	100	100	95	99	4	70	84	0
		Q3	0.4	0.3	β	81	95	96	87	100	98	95	90	10	60	79	0
	Type6	Q26	0.42	0.35	β	81	99	97	94	98	99	92	93	8	43	78	7
		Q27	2.954	2.186	γ	79	100	98	94	96	100	86	89	2	31	80	0
		Q28	0.872	0.813	δ	79	100	98	91	100	100	87	92	2	29	77	0

Initial inspection of the overall facilities in Table 4 shows that facilities of items within each type are all extremely close, indicating considerable success for the definitions. Type 1, however, has an anomalously large overall facility range (from 56% to 64%) and this is mirrored in the facilities for several of the codes considered individually (A1, A3, L2, L4, U1). This prompts closer examination of the items. The sub-types (α, β, γ, δ) listed next to each item in Table 4 have been determined by considering the interaction of two features as in Table 5. Our hypothesis, which is supported by the data in Table 4, is that these four sub-types explain most of the variation in overall facility within any particular item type. Consider the five Type 1 items. The only α item (Q6, 4.8/4.63) has the highest facility 64%, the two β items have facility of 61%, then 60% for the γ item and then 56% for the δ item. Inspection of the other five item types in Table 4 reveals the *same order* (α, then β then γ then δ) for those that are present. It is therefore to be expected that there are variations in student thinking according to these features, which cut across other misconceptions. We propose that the large facility range for Type 1 is because it is the only item type containing all four sub-types. The better facility for comparisons involving decimals with a maximum of two digits probably reflects the effect of specific teaching (*even in the presence of other major misconceptions*) and the better facility for items with non-zero integer part relates to special difficulties with zero (see Stacey *et al*, 2001).

Table 5
Definition of four sub-types from the interaction of two features

Feature 2: Integer part	Feature 1: Number of digits after the decimal points	
	Only one or two digits	More than two digits
Non-zero integer part	α	γ
Zero integer part	β	δ

The remainder of Table 4 (that is the item facilities by code) will now be discussed, with the intention of explaining the bolded entries in Table 3. Light shading has been applied to cells that indicate a variation of at most 5% from the predicted values of 100% and 0%. Dark shading has been applied to cells that are not part of the classification scheme and so have no predictions (e.g. U1 and A3 on Types 3 to 6). The following discussion will then concentrate on the remaining unshaded cells. The U2 column is shaded because small numbers (n=27) do not support detailed analysis.

According to Table 3, Type 1 items caused the most inconsistent responses for the A1 tests (only 76% scored 5 out of 5). Table 4 reveals that, for A1 tests, the two Type 1 items with the lowest facilities are Q10 (3.92/3.4813) and Q9 (0.37/0.216), which are sub-types γ and δ (i.e. involve decimals with more than two digits after the point). Within the A1 group, which is nominally the group of *task experts*, the data from Type 1 items indicates that some, perhaps up to 6% of this sample (and hence about 4% of the original 3531 sample) can only work with decimals to two places. How then, are they correct on the other item types? There is no corresponding drop in facilities for the A1 codes on the Type 2 items (three γ and one δ). Our claim is that these students can correctly order decimals with tenths and hundredths only (not necessarily with understanding) and when this breaks down, the “repair” (in the sense of Brown (1982), i.e. the default strategy) is to choose the longer decimal. This results in correct choices being made on *all* Type 2 items (whatever the sub-type) and correct choices on Type 4 items, and hence tests likely to be coded as A1. This claim is further substantiated by the lower facility by A1 students (79%) on a supplementary item Q29 (0.04/0.038), as students choosing the longer decimal as larger will choose incorrectly. The patterns of facilities for the code A3 also show generally less success with sub-types γ and δ, indicating that there may be a similar group who mix expert decisions on decimals to two places with longer-is-larger thinking within code A3.

According to Table 3, Type 4 items caused the most inconsistent responses for the A2 tests (only 72% scored 0 out of 4). Table 4 reveals that instead of the predicted 0%, the four facilities range from 4% to 9%. Our claim is that, rather than a feature of the items causing inconsistency, this phenomenon is due to a break down in the strategies used by A2 students and a consequence of the coding rules. An A2 student with *money thinking*, or some similar analogy, deals well with comparisons that can be determined within the first two places and hence is correct on all other item types. The prediction is, however, that they are guessing on Type 4 items. If they are coded as A2, it is because they have a score of 0 or 1 on the Type 4 items, nearly always choosing the shorter (incorrect) decimal. If they consistently chose the longer decimal, they would score 3 or 4 out of 4 and be coded as A1. Otherwise they have a score of 2 and the resulting code A3. Hence, the inconsistency of A2 students' choices on Type 4 is due to their strategy rather than the particular variations between items. This claim is supported by the lower facility by A2 students (35%) on a supplementary item Q1 (0.457/0.4), as students choosing the shorter decimal will choose incorrectly. The *money thinking* group is therefore spread across A1, A2 and A3 (making different repairs when their strategy breaks down).

The L1, L2 and L4 students generally choose longer decimals as larger. The most inconsistent response for the L1 tests was on Q12 (4.7/4.08) with the relatively high facility of 11% and this is likely to be due to its sub-type, α. Students in L2 (see two variations in Appendix) were inconsistent on Type 1 items. Table 4 reveals that this is due to two items: Q6 (4.8/4.63, sub-type α) has a relatively high facility of 14% and Q10 (3.92/3.4813, sub-type γ) with 8%. In fact, L4 also has higher facilities for these two items. This leads us to conclude feature 2 of Table 5 is important for L2 and L4 students, who may use different strategies for decimals with the integer part zero and non-zero.

For both S1 and S5, Table 3 indicates that Type 2 items cause the most inconsistency, and Table 4 indicates that this is due mainly to Q17 (0.75/0.5) with higher than expected facilities of 24% (S1) and 13% (S5). While this may be due to the sub-type of the item (β) we have an alternative hypothesis. Code S1 contains the *denominator focussed* students (see Appendix), a generally able group (see Nesher & Peled's fraction rule (1986)), who interpret decimals as fractions but then have difficulty co-ordinating numerator and denominator. They may well recognise 0.5 as a half and 0.75 as three quarters and thus compare correctly. For this reason, Q17 should not be included in the test as a Type 2 item. Amongst the S3 codes, there is a tendency for the β items to have increased facility (e.g. up 6% in Types 5 and 6). This may be related to some combination of *place value number line thinking* and *reciprocal or negative thinking.* This completes the exploration of the bolded entries in Table 3, except for the performance of the S1 code on Type 6 items for which we have no explanation.

Conclusion

On the whole, the data confirms that the item types used in DCT2 are strongly homogeneous, and hence largely serve to identify groups of students with the same conceptions of decimal numbers. Furthermore, the relationship between the misconceptions or *ways of thinking* and the codes allocated to the completed tests is supported. This paper, however, was intended to look beyond the groups identified by previous analyses. Evidence is presented of the presence of two features that interact to give four sub-types of items *within* the item types already defined. The previously defined item types remain dominant, but the additional features identify groupings within these.

For students who have the misconception that longer decimals are larger (most young students), items that involve decimals greater than one are more likely to be answered correctly than those whose value is less than one. Primary teachers should capitalise on this feature and first focus on decimals greater than one (ensuring that students have a good idea of their meaning), and then ensure that the students extend their understanding to decimals less than one. Evidence has been provided to show that among A1 students (the *task experts* on this test) some have little generalised understanding of decimals. Two strategies that break down in certain situations leave their mark in the data. First, choosing according to the digit(s) in the first column(s) after the point very often produces correct answers, but when these digits are the same, students often revert to choosing the longer decimal as the larger. Second, comparing digits from left to right until the larger digit is found is a correct strategy, but students often "repair" it incorrectly and choose on length to overcome an impasse.

Evidence in the data from DCT2 shows some of the variations and combinations of misconceptions that students employ to compare decimals. The L and S groups of misconceptions that are outlined in the Appendix, and whose existence have been supported by interview data from many sources, are supplemented by various pieces of partial correct knowledge to make a large variety of response patterns that can be detected in the data. Many students do not have a coherent view of the quantitative meaning of decimal numbers, but seem to accrete isolated facts, often onto an erroneous base. Students need plenty of activities which put decimal numbers, fractions, positive and negative numbers together (and on a number line); excellent basic instruction in place value which extends beyond hundredths; many opportunities to generalise to very large and to very small numbers; and they should not routinely round to two decimal places.

References

Brown, J. S. & VanLehn, K. (1982). Towards a Generative Theory of "Bugs". In T.P. Carpenter, J.M. Moser & T.A. Romberg (Eds), *Addition and Subtraction: A Cognitive Perspective* (pp117-135). Hillsdale, New Jersey, Lawrence Erlbaum.

Nesher, P. & Peled, I. (1986). Shifts In Reasoning: The Case of Extending Number Concepts. *Educational Studies In Mathematics, 17,* 67-79.

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.

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

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. (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.

Steinle, V., & Stacey, K. (2002) Further evidence of conceptual difficulties with decimal notation. In B. Barton, K.C. Irwin, M. Pfannkuch, & M. Thomas (Eds.), Mathematics Education in the South Pacific. *Proceedings of the 25th Annual Conference of the Mathematics Education Research Group of Australasia, Auckland* (Vol. 2, pp. 633-640). Sydney: MERGA.

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

Appendix: Known ways of thinking associated with codes

Task expert (A1): Correctly completes the task of comparing decimals. Various correct and incomplete strategies might be used singly or in combination throughout the test. Students may "fully understand" or rely on rote rules.

Money thinking (any A): Treats first 2 decimal places like the (whole) number of cents (or cm) so unsure when these are equal. Sees decimals as discrete. Difficulties with Type 4 (eg 4.45/4.4502) as both numbers are like $4.45, and then may truncate or round or guess. (Will be coded as A2 if consistently chooses incorrectly on these items, else A1 or A3.)

First digits only thinking and ***Failed left to right thinking (any A):*** *First digits only* makes comparison only with the first digits (one or two places) after the decimal point but strategy fails when these are equal. *Failed left to right thinking* refers to an incomplete version of a correct procedure. When comparing 3.26 with 3.2618 digits from left to right, the "1" needs to be compared with the "invisible zeros" at the end of the 3.26 to successfully complete the algorithm. Like money thinking, these students are generally correct but need to guess when their procedures fail.

Whole number thinking (L1): Treats decimal portion as another whole number, so 4.8 < 4.75 as 8 < 75. Two variations: *Numerator focussed thinking* chooses 0.53>0.006 as 53>6, while *String length thinking* chooses 0.53 < 0.006 as 006 has 3 digits & 53 has two.

Column overflow thinking (L2): Correctly chooses 4.03 < 4.2 as 3 hundredths < 2 tenths, but incorrectly chooses 4.8 < 4.75 as 8 tenths < 75 **tenths.** The presence of a zero indicates the need to use a new "name". Generally correct on equal length decimals.

Zero-makes-small thinking (L2): Uses *whole number thinking (L1)* with an additional (isolated) fact that a zero after the decimal point 'makes the number smaller'. Correctly chooses 4.03 < 4.2 as the zero in 4.03 makes it small, but incorrectly chooses 4.8 < 4.75.

Reverse thinking (L3): Believes right-most columns have largest place value, so compares from the right-most column first, either due to *mishearing* column names (hundred**ths** as hundreds etc) OR an overgeneralisation of symmetry (larger value columns on outside). So, 4.8 < 4.75 as 5 hundred 7 tens > 8 tens, and 0.42 < 0.35 as 2 < 5.

Denominator focussed thinking (S1): Reads a one digit decimal as a number of tenths, a two digit decimal as a number of hundredths etc and then incorrectly generalises the fact that 1 tenth is greater than 1 hundredth to 'any number in the tenths is greater than any number in the hundredths'.

Place value number line thinking (S1): Works from false analogy between place value columns and number lines. Moving from far left to far right, numbers are indicated in this sequence, numbers in the hundreds (3 digits) then tens (2 digits) then single digit numbers (including 0 which is a 'whole number') then single digit decimals (tenths), two digit decimals (hundredths), three digit decimals (thousandths) etc. Thinks 0.6 less than zero, because zero is in the ones column and 0.6 is in the tenths.

Reciprocal thinking or Negative thinking (S3): Treats decimal portion as another whole number but then as something analogous to the denominator of a fraction (reciprocal) OR as a number 'on the other side of zero' or less than zero (not necessarily negative!). So, 4.82 < 4.3 as 1/82 < 1/3 or as -82 < -3. 'The larger it looks the smaller it is'. Generally makes *in*correct judgements on equal length decimals.

Misread/misrule (U2): Students who get nearly all questions wrong. Either a *task expert (A1)* who misreads the instructions, circling the *smaller* number throughout the test, OR a student following a correct comparison rule (like A1) but then believing that there is a reversal in size (by loose analogy with fractions and negative numbers). Support for *misrule* being widespread is that two thirds of these students select 1.3 > 0.86, whilst being incorrect on almost every item with the same integer part.