YEAR 8 STUDENTS' PROOF RESPONSES:
SOME PRELIMINARY FINDINGS FROM FOUR GROUPS

Celia Hoyles and Dietmar Küchemann

Institute of Education, University of London

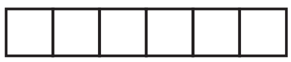
We report on the performance of four high attaining groups of Year 8 students on two written questions, one in algebra, one in geometry, designed to test mathematical reasoning. Preliminary findings suggest that performance is not always consistent between classes, between questions and compared to an overall measure of mathematical attainment.

In this paper we report some preliminary findings from the first year survey conducted as part of a longitudinal study of mathematical reasoning. Information about the study as a whole, and its aims, can be found on the project's website at www.ioe.ac.uk/proof. The study follows on from a survey conducted in 1998 of Year 10 students' conceptions of proof (see Healy and Hoyles, 2000).

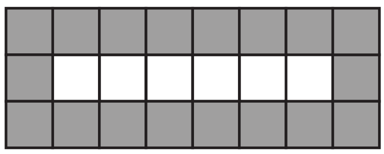
A 50 minute written Proof Survey was administered in June 2000 to high attaining Year 8 students from 63 randomly selected schools within nine geographical areas that spanned England. We report here on two open response questions from the survey, one in algebra (A1) and one in geometry (G1). Frequencies for the sample as a whole are given (nearly 3000 students) and for four groups of students (P1, P2, Q and R). Groups P1 and P2 are parallel top sets from a non selective suburban school, Q is a group of 25 high attaining mathematics students selected from four mixed ability classes in a highly selective grammar school, and R is a top set from an urban comprehensive school.

A1 Lisa has some white square tiles and some grey square tiles.
They are all the same size.

She makes a row of white tiles.



She surrounds the white tiles by a single layer of grey tiles.



How many grey tiles does she need to surround a row of 60 white tiles? . . . .

Show how you obtained your answer.

Figure 1: The first algebra question

Responses to a question about generalising a structure

Question A1 is concerned with generalisation within a familiar setting (tile patterns) and is shown above (Figure 1). (There is extensive work on a generalisation perspective on introducing algebra; see Mason, 1996.)

As well as providing an answer, students were asked to show how it was obtained and responses were coded into 5 broad categories. Correct answers, with clear evidence that a correct structure had been used, were coded 3, 4 or 5, depending on the degree of generality with which the structure was expressed (Table 1)[1].

Code 1	Incorrect answer (180); use of an incorrect number pattern
Code 2	Incorrect answer (eg 120); partial use of correct structure (eg doubles but does not add 6)
Code 3	Correct answer (126); use of correct structure in the specific case of the question with no indication of generality
Code 4	Correct answer (126); use of correct structure indicating its generality
Code 5	Correct answer (126); use of correct structure (expressed in variables)
Code 9	Miscellaneous incorrect answers (including no response)

Table 1: Response codes for question A1

Recognising Structure

Figure 2 shows a typical, if rather minimal, code 3 response. Though there is little in the way of explanation, it is clear that the pattern for 60 white tiles has been seen as two rows of grey tiles (one above the row of white tiles, the other below), with a total of 6 tiles at the ends of the rows.

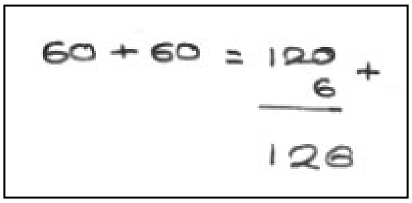


Figure 2: A typical code 3 response to question A1

A characteristic of code 3 responses is that they are couched in terms of a specific number (60) of white tiles. Some students expressed the pattern in more general terms, though it is worth pointing out that this is not necessary to answer the question correctly. Responses like “double and add 6” were coded 4, while responses involving a named variable, like “double the number of white tiles and add 6” and “$2 \times w + 6$” were coded 5.

Spotting an incorrect pattern

The given diagram in question A1 shows 6 white tiles surrounded by 18 grey tiles, and students were asked for the number of grey tiles needed to surround a row of 60 white tiles. A substantial number of students gave the answer 180, on the basis that since 60 is 10 times 6, the number of grey tiles will be 10 times 18, or (less commonly) on the basis that since 18 is 3 times 6, the number of grey tiles will be 3 times 60. Both were given a code of 1. Pattern spotting responses were not unexpected. However, we were surprised first by the overall frequency of code 1 responses (37 %, N = 2796), second by the observation that there could be marked differences in frequency between comparable classes, and third that in some groups or classes, code 1 responses were often given by students who performed well on our Baseline Maths Test[2] and on other questions in the Proof Survey.

The solid black columns in Figure 3 show the frequencies of the codes for question A1 for the total Year 8 sample (N = 2796). As can be seen, over one third of all the students gave code 1 responses, which is more than the proportion who gave code 3 responses and almost as large as the proportion of students who answered the question correctly (ie gave a code 3, 4 or 5 response).

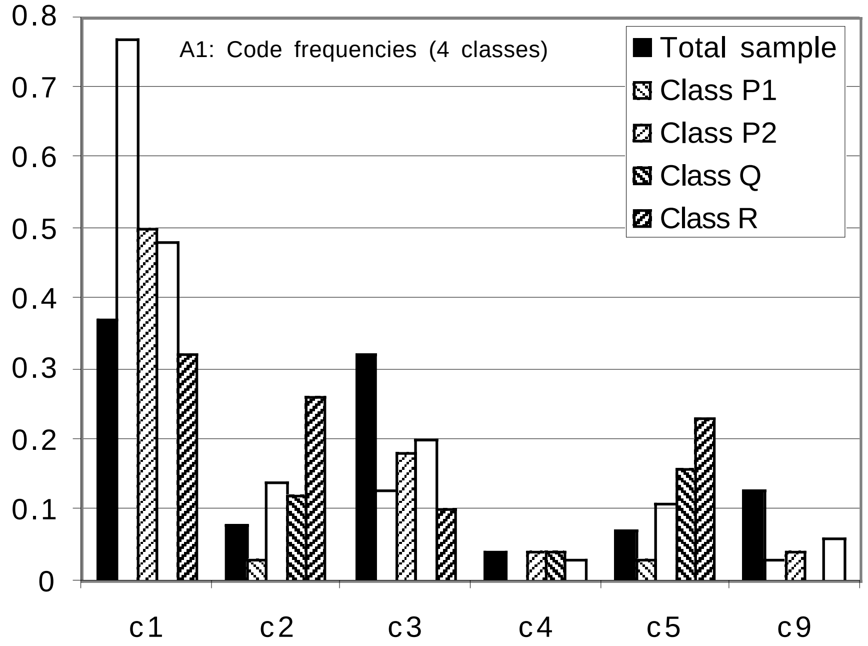


Figure 3: A1 Code frequencies for total sample and for four groups

Figure 3 also compares the responses for the total sample with the responses for classes P1, P2, Q and R. Two features are of particular interest. One is the difference, especially for code 1, between classes P1 and P2. These are parallel classes from the same school, which strongly points to the operation of teacher influences, though at this stage we do not know what these might be. A second interesting feature is the relatively high frequency for code 1 for group Q. The students in this group were selected for their high mathematical attainment within an already selective school. Many of those who gave a code 1 response to question A1 gave high level responses to other questions on the proof survey; many also scored very highly on the Baseline Maths Test, as shown in Table 2. Again, the reasons for the high frequency of code 1 responses for group Q are not yet known, though the textbook used in the school might provide a clue[3].

	Group Q	A1 codes					
		c1	**c2**	**c3**	**c4**	**c5**	**c9**
Baseline Mathematics Test scores	**15**			1			
	16						
	17						
	18	1	1				
	19	5	1	1			
	20	2	1	1			
	21	3		1		3	
	22	1		1	1	1	

Table 2: A1 codes against Baseline Test scores for group Q (N = 25)

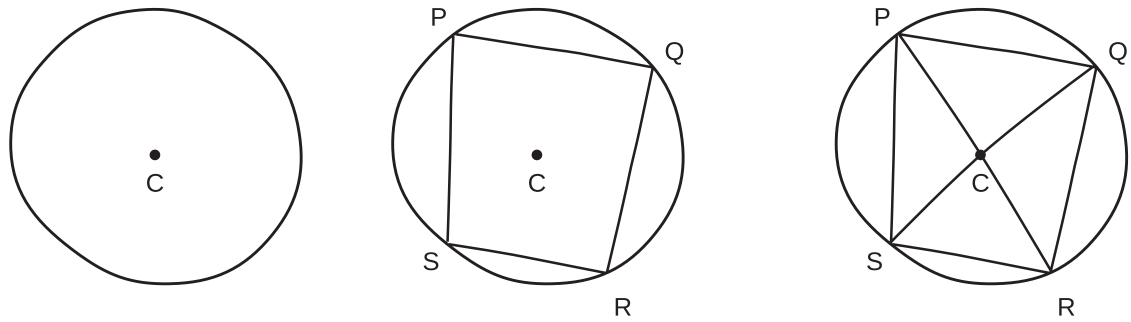
Responses to a question to distinguish perceptual from geometrical reasoning

Question G1 (Figure 4) is concerned with how far students use perception or geometry in proofs (see Lehrer and Chazan, 1998). A geometric diagram is presented which supports a conjecture that turns out to be false. Students are asked whether or not they agree with the conjecture and to explain their decision.

G1 Darren sketches a circle. He calls the centre C.

He then draws a quadrilateral PQRS, whose corners lie on the circle.

He then draws the diagonals of the quadrilateral.



Darren says
"Whatever quadrilateral I draw with corners on a circle,
the diagonals will always cross at the centre of the circle".

Is Darren right? . . . .

Explain your answer.

Figure 4: The first geometry question

Responses to question G1 were coded into 6 broad categories, which are shown below (Table 3)[4].

Code 11	Incorrect answer (Yes)
Code 12	Correct answer (No); no valid explanation
Code 2	Correct answer (No); weak explanation or drawing of weak counter example
Code 3	Correct answer (No); clear description or drawing of counter example
Code 4	Correct answer (No); analytic reasoning (dynamic or static)
Code 9	Miscellaneous incorrect answers (including no response)

Table 3: Response codes for question G1

Responses that agreed with the false conjecture were coded 11. Correct responses but with no valid explanation (and which might well have been the result of guesswork) were coded 12. The code 2 responses were correct but the supporting explanations were weak; they would consist either of a not-incorrect but vague verbal statement such as "If the sides are different the diagonals will not meet at the centre", or of a drawing where the quadrilateral was almost a rectangle and where the supposed centre of the circle and the point of intersection of the diagonals were close together.

Figure 5 shows a typical code 3 response, namely a drawing of a 'decisive' counter-example: here the quadrilateral is plainly not a rectangle and the diagonals clearly do not intersect at the centre of the circle. Such drawings were not unexpected; however we were surprised (and delighted) by drawings such as the one in Figure 6, where the centre of the circle lies outside the quadrilateral, so that the centre and the point of intersection of the diagonals cannot possibly coincide. This was also classed as a code 3 response.

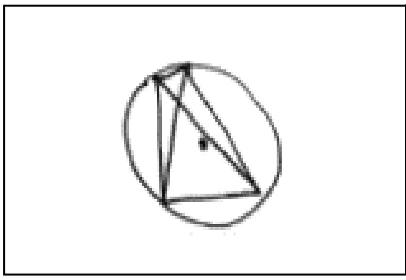


Figure 5: A typical, 'decisive' code 3 response for G1

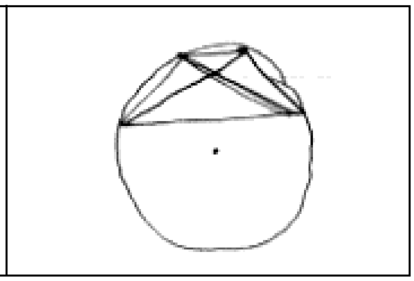


Figure 6: An 'absolutely decisive' code 3 response for G1

Some students gave what we called 'analytic' responses (coded 4) where, rather than simply demonstrating that the conjecture is wrong by means of a counter-example, the focus was narrowed onto certain features of cyclic quadrilaterals which showed that the conjecture *had* to be wrong.

The solid black columns in Figure 7 shows the frequencies of the codes for question G1 for the total Year 8 sample. As can be seen, nearly half the students agreed with the false conjecture (code 11), which is far greater than we expected, though it should be said that nearly as many students were able to provide clear counter-examples (code 3). There were few code 4 responses. Figure 7 also compares the responses for the total sample with those for classes P1, P2, Q and R. Here there are less obvious differences between classes P1 and P2 than for question A1; also the group Q frequencies are much closer to what one might have expected. Perhaps

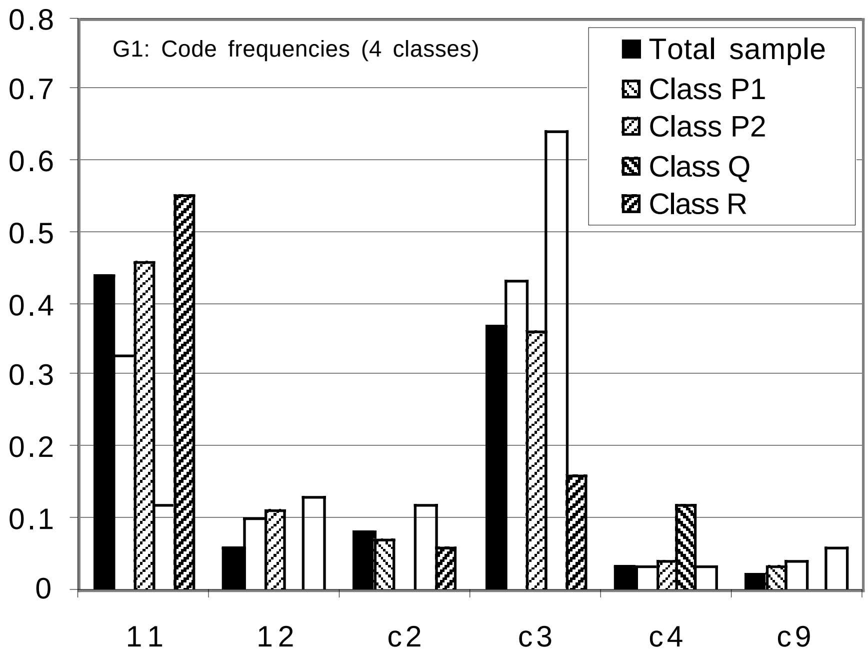


Figure 7: G1 code frequencies for total sample and for four groups

the most notable feature of the graph is the relatively poor performance of class R and this would be worth investigating further.

Discussion

A full scale statistical analysis of these data together with those from surveys of the same students in years 9 and 10 will enable us to map the way students' mathematical reasoning develops over time and to identify factors that are optimal in this development. In the meantime, the descriptive statistics presented here suggest, amongst other things, that performance in mathematical reasoning is not necessarily consistent across topics (algebra and geometry), is not necessarily similar in comparable classes and is not necessarily closely related to overall mathematical attainment; and whereas many high attaining Y8 students appreciate the value of a counter example, there is also widespread use of pattern spotting in data regardless of mathematical structure.

Acknowledgement

We gratefully acknowledge the support of the Economic and Social Research Council (ESRC), Project number R000237777.

Notes

1. Most categories were divided further, giving 22 subcategories in all.
2. All the students in the survey sat a general baseline test of mathematics attainment a few weeks before taking the proof test.
3. The Year 8 book devotes several pages to number sequences, and these are presented in a fairly open way; however, the setting is nearly always purely numerical, rather than involving spatial patterns as in A1.
4. As with A1, categories were divided further, though not to the same extent (there were 14 subcategories in all, compared to 22 for A1).

References

Healy, L. & Hoyles, C.: 2000, ‘A study of proof conceptions in algebra.’ *Journal for Research in Mathematics Education*, 31(4), 396-428.

Lehrer, R. & Chazan, D.: 1998, *Designing Learning Environments for Developing Understanding of Geometry and Space*. London: LEA.

Mason, J.: 1966, ‘Expressing generality and roots of algebra.’ In N. Bednarz, C. Kieran and L. Lee (eds.), *Approaches to Algebra* (pp. 65-86). Dordrecht: Kluwer Academic Publishers.