YEAR 10 STUDENTS' PROOFS OF A STATEMENT IN NUMBER/ALGEBRA AND THEIR RESPONSES TO RELATED MULTIPLE CHOICE ITEMS: LONGITUDINAL AND CROSS-SECTIONAL COMPARISONS

Dietmar Küchemann and Celia Hoyles

Institute of Education, University of London

We found, in two separate studies (1996 and 2002), that high attaining Year 10 students in English schools tend to produce empirical proofs, though many of them seem able to appreciate some of the qualities of more powerful proofs. Students rate algebraic proofs highly, often for superficial reasons, though we found that in the second, longitudinal, study they were more discriminating in Year 10 than they had been in Year 9.

INTRODUCTION

Proof, where it involves deductive reasoning based on general relationships, distinguishes mathematics from science and from argumentation in daily life, where reasoning is more usually based on experimental evidence or analogy.

Our work (eg, Healy and Hoyles, 2000; Küchemann and Hoyles, 2001) suggests that even when school students are able to appreciate the qualities of a mathematical proof, their own explanations may be low in insight and instead consist mainly of empirical support for the statement they are trying to prove. It is possible to find abundant evidence (eg Bell, 1976; Balacheff, 1988; Coe and Ruthven, 1994) of school students having difficulty in providing mathematical explanations and who seem to adopt proof schemes that are empirical or external (Harel and Sowder, 1998) rather than involving general mathematical relationships - ie who at best use what Bills and Rowland (1999) call ‘empirical’ rather than ‘structural’ generalisations. There are also studies to suggest that some students, having learnt a mathematical procedure, may show little interested in why it works (eg Hiebert and Wearne, 1988). On the other hand, even young children seem able to engage in sophisticated forms of explanation and justification, given a classroom culture with appropriate sociomathematical norms (see eg Yackel, 2001).

THE STUDY

In this paper we look particularly at responses to two questions (A3 and HA4) which were devised by Healy and Hoyles (ibid) and which formed part of a written test that they gave to 2459 high attaining Year 10 students in 1996. The same questions were given to a similar sample (N = 1512) of high attaining Year 10 students in 2002, in research undertaken by the authors for the Longitudinal Proof Project, which ran from 1999 to 2003. The aim was to look for similarities and contrasts in patterns of student response.

Students' proof choices

Question A3 had a multiple choice format (see Figure 1, below). Students were presented with various 'proofs' of the statement "When you add any 2 even numbers, your answer is always even" and were asked to choose the proof which was nearest to their *own approach* and which would get the *best mark* from their teacher. In 2002 students were also asked which proof they *liked best*.

Aysha, Brian, Coby, Deon, Eric and Fiona were trying to prove whether the following statement is true or false:

When you add any 2 even numbers, your answer is always even.

Aysha's answer

a is any whole number.

b is any whole number.

$2a$ and $2b$ are any two even numbers.

$2a + 2b = 2(a + b)$.

So Aysha says it's true

Brian's answer

$2 + 2 = 4$ $\quad$ $4 + 2 = 6$

$2 + 4 = 6$ $\quad$ $4 + 4 = 8$

$2 + 6 = 8$ $\quad$ $4 + 6 = 10$

So Brian says it's true

Coby's answer

Even numbers are numbers that can be divided by 2. When you add numbers with a common factor, 2 in this case, the answer will have the same common factor.

So Coby says it's true

Deon's answer

Even numbers end in 0, 2, 4, 6 or 8. When you add any two of these the answer will still end in 0, 2, 4, 6 or 8.

So Deon says it's true

Eric's answer

Let x = any whole number, y = any whole number.

$x + y = z$

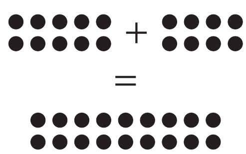
$z - x = y$

$z - y = x$

$z + z - (x + y) = x + y = 2z$

So Eric says it's true

Fiona's answer



So Fiona says it's true

a) Whose answer do you like best? . . . .

b) Whose answer is closest to what you would do? . . . .

c) Whose answer would get the best mark from your teacher? . . . .

Fig 1: Question A3 (2002 version)

In HA4, students were asked to produce a proof for a similar statement to the one in A3 (this time concerning odd numbers). It was placed immediately after A3 on the test, in the belief that the options in A3 might help students devise a proof in HA4.

Option A in A3 is a 'structural' proof, expressed in algebraic form. Option B is empirical, based on just 6 examples (albeit fairly systematic ones). C is structural, like A, but expressed in narrative form. D is an exhaustive proof. It says something about the properties of all even numbers (namely, that in our number system they happen to end in 0, 2, 4, etc), but is essentially empirical rather than structural. It describes how even numbers behave, but not why. E is a pseudo or nonsense proof but, like A, is expressed in algebraic form. Option F was intended to be a structural proof, like A and C, but expressed 'visually', with sets of dots representing generic examples of even numbers. However, in retrospect the option is perhaps too cryptic, since the sets of dots can easily be interpreted as representing specific even numbers, making it an empirical proof. In the event, F was not a popular choice, perhaps because of this ambiguity, and we do not discuss it further in this paper.

Options A, C and D are all valid proofs of the given statement, in that they verify that the statement is true. However, A and C might be thought to be more satisfying (and educationally more useful) in that they also illuminate the statement, ie explain why it is true. Option B confirms the truth of the statement, but does not prove it, while E is nonsense.

The frequencies of the Year 10 students' choices in 2002 are shown in Table 1, below. The table lists the six options, in decreasing rank order of popularity, for the three criteria of *like best*, *own approach*, and *best mark*. As is immediately apparent, there are some dramatic changes in order for the different criteria.

	Year 10 choices for A3		LIKE best		OWN approach		BEST mark	
A	ALGEBRA-structure		D	35%	B	41%	E	38%
B	EMPIRICAL-6 examples		B	17%	D	29%	A	24%
C	NARRATIVE-structure		C	17%	A	13%	C	20%
D	EMPIRICAL-exhaustive		A	13%	C	9%	D	9%
E	ALGEBRA-nonsense		F	10%	F	3%	B	3%
F	VISUAL-structure		E	6%	E	3%	F	1%
c9	miscellaneous		c9	2%	c9	3%	c9	5%

Table 1: Y10 students' choice frequencies for A3 in 2002 (N = 1512)

Looking first at the algebraic proofs, few students seem to *like* them, perhaps because they find them difficult (A, 13%) or impossible (E, 6%) to understand; even fewer claim that they are close to their *own approach*, perhaps for the same reasons; however, they are the two most popular choices for best mark, with option E (38%), which is the more algebraic-looking of the two, even more popular than A (24%). This latter result is perhaps not surprising since in the popular imagination high powered maths is commonly equated with algebra.

As far as the empirical proofs are concerned (B and D), these are the two most popular choices for *like best* and for *own approach*, perhaps in large measure because they are relatively easy to understand. Interestingly, D is the most popular choice for *like best* (35% compared to 17% for B) and B the most popular for *own approach* (41% compared to 29% for D). This suggests that many students can appreciate that D is a powerful proof, but admit that B is closer to their own approach, even though it is more limited. When it comes to *best mark*, these two proofs are ranked very low, perhaps in part because they are not algebraic, but perhaps also because students recognise their limitations (namely that D is not general and B is not illuminating).

Finally, many students seemed able to appreciate the structural quality of proof C: though few chose it for *own approach* (9%), a substantial minority chose it for *like best* (17%) and also for *best mark* (20%) despite it being in narrative rather than algebraic form.

In the Longitudinal Proof Project, the students were given a similar question to A3 in Years 8 and 9. However, there were only 4 options in Year 8, non of which were algebraic, and only 5 options in Year 9. Also, the content, though always involving number/algebra, changed from year to year. Thus it is not possible to make simple longitudinal comparisons, although one can discern some trends. For example, there are some interesting changes in the *best mark* frequencies for each year's narrative-structural proof. In Year 9, this proof has a frequency of just 6% and is swamped by the two algebra proofs (48% and 28%), despite its strong showing in Year 8 (53%). However, its popularity increases again in Year 10 (20%), suggesting that the students are beginning to judge algebraic proofs more critically.

Comparisons with 1996

The version of A3 used with Year 10 students in 2002 was the same as the one used in 1996 with Year 10 students in the predecessor project, except for the addition of the *like best* criterion in the later version. In both cases the sample consisted of students in top sets from randomly selected schools, and though nothing further was undertaken to produce comparable samples, the response frequencies shown in Table 2 below suggest that the samples were in fact remarkably similar.

Criterion for choice	Distribution of choices: algebra (A3)											
	D empirical-exhaustive		B empirical 6 examples		C narrative-structure		F visual-structure		A algebra-structure		E algebra-nonsense	
	1996 %	2002 %	1996 %	2002 %	1996 %	2002 %	1996 %	2002 %	1996 %	2002 %	1996 %	2002 %
like best		35		17		17		10		13		6
own approach	29	29	24	41	17	9	16	3	12	13	2	3
best mark	7	9	3	3	18	20	9	1	22	24	41	38

Note: Underlined frequencies are 'substantially' higher than their other-year counterparts

Table 2: Y10 students' choice frequencies for A3 in 1996 (N = 2459) and 2002 (N = 1512)

The *like best* criterion was added because we had noticed that without it (as for example in our Year 8 version of A3), there seemed to be a tendency, especially amongst boys (Küchemann and Hoyles, ibid), to choose an option that they liked for *own approach*, rather than one that was genuinely similar to what they would have constructed themselves. It is a moot point whether one should ‘improve’ questions in this way, as it makes comparisons more difficult - and it renders a detailed discussion of the frequencies in Table 2 beyond the scope of this short paper. However, it is interesting to note the large increase in the *own approach* frequency for the empirical proof B, which perhaps indicates a growth in a ‘pragmatic’, data-generating approach to mathematics.

Students’ constructive proofs

Question HA4, which asked for a proof of the statement “When you add any 2 odd numbers, your answer is always even”, appeared immediately after A3 on the 1996 and 2002 written tests. Table 3 gives an indication of the *type* (but not the quality) of proof that students constructed in 2002. The 1996 frequencies are broadly similar. What is immediately apparent is the popularity of the purely empirical (31%) and empirical-exhaustive (19%) approaches, which chimes with the *own approach* frequencies for A3. At the same time, a sizeable proportion of students (17%) embarked on narrative proofs in which the structure of odd numbers is described effectively. Very few students, though, attempted an algebraic proof, which again echoes the *own approach* (and *like best)* frequencies for the algebra proofs in A3.

HA4: Proof types

EMPIRICAL	1 example	2%		
	several examples	15%		
	'crucial' example	14%	30%	31%
EMPIRICAL-EXHAUSTIVE	Odds end in 1,3,5,7,9	13%		
	fairly exhaustive	3%		
	very exhaustive	4%	7%	19%
ALGEBRA	NO structure	3%		
	PARTIAL structure: a=even, a+1=odd	2%		
	FULL structure: 2n+1 = odd	3%		
NARRATIVE	NO structure	1%		
	PARTIAL structure: 'up in 2s'	1%		
	FULL structure: odd = even plus 1	17%	18%	
VISUAL	NO structure	0%		
	PARTIAL structure	0%		
	FULL structure: OOOO OOOOO = odd	6%	6%	
Other		16%		

Table 3: Frequency of proof types of Y10 students’ constructive proofs in 2002 (N = 1512)

CONCLUSION

Evidence from the *Longitudinal Proof Project*, and its predecessor, suggests that even high attaining students in English schools have a strong propensity to construct empirical rather than structural proofs. At the same time, many students seem able to appreciate some of the qualities of more powerful proofs, even if they cannot, or do not attempt to, construct such proofs themselves. This suggests that carefully designed teaching which helps students evaluate and characterise different kinds of proofs could have a marked impact on the quality of students' explanations, provided it is sustained and built upon over time. We are currently exploring ways of doing this in our new DfES-funded project, *Developing Research-Informed Materials in Mathematical Reasoning for Teachers.*

Acknowledgement

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

REFERENCES

Balacheff, N.: 1988, 'Aspects of proof in pupils' practice of school mathematics', in D. Pimm (ed.), *Mathematics, Teachers and Children*, Hodder and Stoughton, London, pp 216-235.

Bell, A. W.: 1976, 'A study of pupils' proof-explanations in mathematical situations', *Educational Studies in Mathematics*, 7, 23-40.

Bills, L. and Rowland, T.: 1999, 'Examples, generalisation and proof', in Brown, L. (ed), *Making Meanings in Mathematics*, Annual Proceedings of BSRLM.

Coe, R. and Ruthven, K.: 1994, 'Proof practices and constructs of advanced mathematics students', *British Educational Research Journal*, 20, 1, 41-53.

Harel, G. and Sowder, L.: 1998: 'Students' proof schemes: results from exploratory studies', in A. H. Schoenfeld, J. Kaput and E. Dubinsky (eds), *Research in College in Mathematics Education III*, American Mathematical Society, Providence, R. I., pp 234-283.

Healy, L. and Hoyles, C.: 2000, 'A study of proof conceptions in algebra', *Journal for Research in Mathematics Education*, 31, 4, 396-428.

Hiebert, J. and Wearne, D.: 1988, 'A cognitive approach to meaningful mathematics instruction: testing a local theory using decimal numbers', *Journal for Research in Mathematics Education*, 19, 371-384.

Küchemann, D. and Hoyles, C.: 2001, 'Identifying differences in students' evaluation of mathematical reasons', *British Society for Research into Learning Mathematics, Proceedings*, 21, 1, 37-42.

Yackel, E.: 2001, 'Explanation, justification and argumentation in mathematics classrooms', PME 25, 1, Freudenthal Institute, Utrecht, Holland, 9-23.