Sources of Certainty and Uncertainty in Mathematical Problem Solving

Kaye Stacey and Cecilia de Beato

Institute of Education, University of Melbourne, Australia

This study investigates the certainty and uncertainty that students feel as they work on a mathematical problem. It is hypothesised that the over-confidence in decisions that characterises reasoning in many fields of human endeavour is also exhibited in mathematical work and that it may partly explain why students generally are reluctant to check their work. Students who feel certain that their work is correct would see little reason to check it. In the problem used in this study, uncertainty arose in making a generalisation, but also from carrying out straightforward calculations . Students with wrong methods that gave easy arithmetic were, in the end, almost as certain that their answers were correct as students with the correct method. These observations may help to explain why students with "obviously" wrong answers do not check, why students more often check arithmetic than reasoning and the tendency for groups to choose a simple wrong solutions even when a correct solution been proposed.

Investigations into the strategies and methods students use to solve mathematical problems have found that generally students do not check their work spontaneously (Davis, Jokusch & McKnight, 1978; Kantowski, 1977; Lee and Wheeler, 1987; Stacey, 1989). When they do, it is often an incomplete check or simply a repetition, rather than a re-assessment, of what they have just finished. Indeed, Stacey and Groves (1985) have noted that with many students, a verbal instruction to check their work is interpreted only as an instruction to repeat it. Several reasons have been advanced in the literature to explain why students do not check their work. The "looking back" phase is generally recognised as the most neglected, both in actual problem solving and in teaching emphasis. Schoenfeld (1985, p 316) notes that "lack of plan assessment and absence of review" are major factors contributing to failure in problem solving. When an answer has been obtained to a problem the natural, but immature, response is to want to move on quickly to the next task, rather than reflecting on what has been achieved. There is also a substantial body of evidence that some students do not know how to check their work. The research reviewed by Bell, Costello and Kuchemann (1983) concerning understanding of proof showed widespread misunderstanding both of the role of counterexamples and of the inadequacy of examples for proving that a generalisation is true. Understanding of both content is also involved here. Lee and Wheeler (1987), for example, point out that students can only use substitution of numbers as a strategy to check algebraic manipulations when they have a reasonably clear understanding of the relationship between arithmetic and algebra.

In this paper we examine another factor which underlies students' checking behaviour, which was suggested to us by Fischbein's (1987) analysis of the role of intuition in mathematical thinking. A series of research studies, reviewed by Fischbein (1987), into the subjective evaluation of certainty has established that for a reasoning endeavour to continue, a person must feel confident about the decisions they make during it. A steady tendency has been found for people to be over-confident in the accuracy of their own knowledge, decisions, interpretations and solutions. This natural tendency to over-confidence in the decisions that they make whilst solving a problem may therefore contribute to students' lack of checking, particularly of their reasoning.

Fischbein brings together experimental findings from various sources to develop his theory that a high degree of certainty in a reasoning process, a fundamental need of the human mind, is produced by reliance on self-evident, intrinsically certain, persistent intuitions. "During the very course of our reasoning, of our trial-and-error attempts, we have to rely on representations and ideas which appear, subjectively, as certain, self-consistent and intrinsically clear" (p x). Although most of Fischbein's analysis concerns conceptual structures and clusters of beliefs, he also recognises "anticipatory" intuitions which are specific to the problem solving process. Observations of students solving mathematical problems by Galbraith (1986) also point to a desire for high levels of certainty. He noted that students wanted to "close upon a definite result rather than maintain and open mind in the absence of conclusive information" (p 430) and observed a tendency to redefine or amend data so that it supports a particular inference or conjecture.

Galbraith also noted that when students were asked to select the better of two explanations, many students made their choice of the grounds of simplicity. This observation has also been made by Stacey (1990) who observed that when groups of students were solving problems together, they often chose a simple incorrect solution, even when a correct solution had been proposed by a group member. The crucial factor in whether the group solution was correct seemed not to be having the ideas, so much as choosing between them. Misplaced confidence in a simple idea and lack of adequate checking strategies were common faults.

Aims

With the background outlined above in mind, data was collected to explore the following questions:

(i) What factors cause certainty and uncertainty in problem solving amongst high school students?

(ii) How does the certainty of students who answer a question correctly compare to the certainty of students who answer it incorrectly?

(iii) Is there a relationship between certainty and checking? In particular are students who do a problem in only one way more certain of their results than students who search for other approaches?

Method

One problem (see Figure 1) was given to 227 Year 8 students (average age 13 years) at two girls' schools during class time. After completing the problem, the students answered a questionnaire. About 15 students (exhibiting different responses) were interviewed within two days of the problem solving to further elucidate the reasons for the responses. The questionnaire asked students to rate their certainty in their answers to the three parts of the problem on scales from 0 to 10. The zero, five and ten positions on the scale were annotated with descriptive comments, such as "completely sure you answer is wrong" at zero. This measurement of certainty was adapted from calibration studies reviewed in Lichenstein, Fischhoff and Phillips (1982). Then students were asked to indicate, separately for the 10x10 and 50x50 blankets, whether they had done the question in only one way or in more than one way and whether they had obtained one or more answers. This was the criterion selected for judging whether students had checked their work. Finally each student was presented with one of six additional pieces of information and then asked to re-assess their certainty for the 50x50 blanket. The purpose of the additional information was to investigate how supporting or contradictory evidence affected certainty. Space precludes reporting the results from this part of the study, but comparison of the changes in certainty due to each version indirectly gave insight into the methods that students used to check their work. In particular, it was found that students tended to use the gross features of the additional information as a check of reasonableness of answers, rather than to directly compare predictions made from their rules with the independent, additional evidence. It is very difficult, even in interview, to establish what checking a student is actually doing (Lee and Wheeler, 1987) and so the further refinement of this as a research tool is indicated.

Imagine you are going to make a patchwork doll's blanket by sewing together some tiny squares of material measuring 1cm by 1cm. You want to know how much sewing you will have to do to make the blanket.

To make a square blanket measuring 3cm by 3cm, you need 9 squares of material and it takes 12 cm of sewing. (Diagram given of 3x3 blanket with sewing between the squares, but not along the outside edge of the blanket, clearly shown)

- *How much sewing is needed to make a square blanket measuring 5cm x 5cm?*
- *How much sewing is needed to make a square blanket measuring 10cm x 10cm?*
- *How much sewing is needed to make a square blanket measuring 50cm x 50cm?*

Figure 1: Making a Blanket

Results and Discussion

About half of all students and three quarters of the students who were correct checked their work in some way. Of those who did check, 42% were correct whereas only 18% of those who did not check were correct. This difference is statistically significant (chi-squared=5.3, d.f.=1, $p<0.02$). Students who checked their work and found the same answer in two different ways had significantly greater certainty than students who did not check ($p<0.05$). These in turn had higher certainty than students who checked but found two different answers.

In order to relate changes in certainty to mathematical behaviour, responses were classified according to the solution method used. Some students followed one general rule for all three answers. Some began by drawing the 5x5 blanket and counting the number of centimetres of sewing but used a general rule for the 10x10 and 50x50 blankets. Others used the generalisation only for the third blanket. The number of students using the most popular methods are given in Table 1, which uses algebraic notation to indicate the rules even though the students themselves only indicated their generalisations by the calculations they performed. Thus, for example, the third row of the Table indicates that 20 students found the answers to all the parts of the question by calculating n^2+n for n = 5, 10 and 50, although they did not write the algebraic notation. This rule (and also the perimeter rule) were popular because they fitted the given information that a 3x3 blanket required 12cm of sewing. Table 1 also shows the mean certainty on each part of the problem for the students using each rule. In all cases, students were most certain about their answers for the 5x5 blanket and their certainty dropped for the 10x10 blanket and again for the 50x50 blanket. The higher certainty for the 5x5 blanket was associated with use of drawing and counting to obtain the answer. The only two groups that did not count (the perimeter and the n^2+n rule) had the lowest certainties for this part. These groups of students seemed to settle immediately on a solution that happens to fit the given data without exploring the situation in any depth, but they did not do this because they were very certain of the answer. In Fischbein's terms, they did not have an anticipatory intuition of which they were very sure. Instead, they seemed merely to accept the uncertainty and not do anything about it - although a few of them drew the 5x5 blanket, they did not even count.

Table 1
Mean Certainty of Users of Different Rules for 5x5, 10x10 and 50x50 blankets.

Method	N	5x5		10x10		50x50	
		rule	mean	rule	mean	rule	mean
correct rule	46	count	8.89	2n(n-1) or count	8.17	2n(n-1)	7.20
perimeter	51	4n	7.10**	4n	6.94**	4n	6.82
$n^2 + n$ rule	20	$n^2 + n$	6.60**	$n^2 + n$	6.55**	$n^2 + n$	5.45**
$2n^2$ rule	3	$2n^2$	7.67	$2n^2$	6.33	$2n^2$	5.67
count and $2n^2$	2	count	10.00	$2n^2$	9.00	$2n^2$	8.50
count and $2n^2$	4	count	9.25	count	9.25	$2n^2$	7.25
multiples of previous answer	7	count	9.57	twice 5x5 answer	8.14	5 times 10x10 answer	7.29
multiples of previous answer	14	count	8.79	count	7.79	5 times 10x10 answer	6.43

** sig. difference between this mean and corresponding mean for the correct rule ($p<0 .01$)

Uncertainty on moving to a generalisation.

Naturally certainty dropped when students moved from counting to a general rule. The mean drop in certainty when students counted on two consecutive parts of the problem was 0.78 (averaged over 18 unambiguous instances from Table 1) and when students used a rule for two consecutive parts, the mean drop was 0.32 (157 instances). However, there was a much larger mean drop in certainty for students moving from counting to use of a general rule (drop of 1.47 averaged over 27 instances). Students with the correct rule were excluded from this analysis because it was sometimes unclear whether they had counted or used the rule or both in the second part of the problem.

Uncertainty when calculating.

Some of the interviews indicated that part of the drop in certainty from the 5x5 blanket to the 50x50 blanket could be attributed to uncertainty about the results of arithmetic calculations involving larger numbers. For the purpose of this investigation, a calculation was defined as hard if it involved multiplication of two numbers over ten, even though this skill is taught at least four years lower down the school than Year 8. The mean drops in certainty from the 10x10 to the 50x50 blankets for students using the same rule on these parts of the problem were calculated separately for methods involving hard calculations (n^2+n, $2n^2$ rules requiring 50^2) and easy calculations (perimeter, multiples requiring multiplication by 4 or 5). The mean drop for rules involving hard calculations was 0.96 (25 instances), which is significantly greater than the mean drop of 0.21 for rules involving only easy calculations (58 instances). The size of this effect had not been expected. Over three years after students have first learned how to perform multiplications such as 50x50=2500, the uncertainty involved with it is still comparable with the uncertainty they feel when making a generalisation. This may partly explain why students' checking concentrates so much on repeating arithmetical calculations, where teachers see a greater need to concentrate on reasoning.

Certainties of students with a simple or complex rule.

A comparison of students with the correct rule and the perimeter rule shows the practical implications of the drops in certainty due to generalising and harder calculations. The perimeter group began with an immediate generalisation, based on little evidence. They did not subsequently have to make the transition to a generalisation and their calculations were at all stages very simple. Thus, although the perimeter group were significantly less certain of their 5x5 blanket answers ($p<0.01$) than the correct group, there was no significant difference between their certainties for the 50x50 blanket and those of the correct group. On the challenging part of the problem, the certainties of students who are correct and those who have grabbed the simplest wrong rule apparently without any investigation, are comparable.

Conclusion

The use of certainty ratings has proved to be a promising research tool for understanding students' thinking during problem solving. Students who jumped quickly to generalisations based on little evidence did not seem to do so because of stron certainty, an intuition in Fischbein's sense. However, their certainty in the correctness to the final part of the problem is comparable to that of students who are correct. Thus, in a group discussion, both a simple wrong rule and the correct rule may be propounded with equal conviction, leading a group to choose the wrong answer over the right answer. Both making a generalisation and doing arithmetical calculations caused students to lose certainty in their work. Students who have made wrong assumptions at the start of their work which happen to lead to simple processing will be no more likely to check their work at the end than students who are correct. Their simple (wrong) rule has "proved itself" in the ease of calculation of the answers it produces.

References

Bell, A.W.,Costello, J. &Kuchemann , D. (1983). *Research on Learning and Teaching*. Windsor: NFER-Nelson.

Davis,R.B., Jokusch, E. & McKnight,C. (1978). Cognitive processes in learning algebra. *Journal of Children's Mathematical Behavior*, 2(1), 10 - 320.

Fischbein, E. (1987). *Intuition in Science and Mathematics.* Dordrecht: Reidel.

Galbraith, P.L. (1986). The use of mathematical strategies: factors and features affecting performance. *Educational Studies in Mathematics*, 17, 413 - 441.

Kantowski, M.G. (1977). Processes involved in mathematical problem solving. *Journal for Research in Mathematics Education*, May, 163 - 180.

Lee, L. & Wheeler, D. (1987). *Algebraic Thinking in high school students: their conceptions of generalisation and justification.* Research report. Montreal: Concordia University.

Lichenstein,S., Fischhoff, B. & Phillips, L.D. (1982). Calibration of probabilities: The state of the art to 1980. In D. Kahnemann, Slovic, P. & Tversky, A. *Judgement under uncertainty*. Cambridge: Cambridge University Press.

Schoenfeld, A. (1985) *Mathematical Problem Solving.* Orlando: Academic Press.

Stacey, K. (1989). Finding and using patterns in linear generalising problems. *Educational Studies in Mathematics*, 20,147 - 164.

Stacey, K. and Groves, S. (1985) *Strategies for problem solving*. Melbourne: Latitude Publications.

Stacey, K. (1990, June) Problem solving in groups: are two heads better than one? Paper presented at the Annual Conference of the Mathematics Education Research Group of Australia, Bathurst, NSW.