"YOU HAVE TO PROVE US WRONG": PROOF AT THE ELEMENTARY SCHOOL LEVEL[1]

Vicki Zack

St. George's School and McGill University, Montreal, Quebec

In solving a variant of the 'chessboard' task, a team of fifth grade elementary students are convinced ***that*** *their pattern works. They use what they know of the pattern to refute an argument by peers. There is evidence of conjecture, refutation and generalization, and aspects of proving. It is the teacher's contention that in order for an argument to be considered a proof, the students need not only* ***convince****, but also to* ***explain****. Thus teacher involvement and personal inquiry are in this instance necessary to provoke thought concerning* ***why*** *the pattern works as it does .*

In my current research I am endeavoring to see how the learning of mathematics is interactively accomplished within my fifth grade classroom. I consider what individual children and the teacher contribute to this collective activity. In this paper I will be concerned specifically with proof in this context. I will show how in work with one task which was not at first assigned with any intention of attending to proof, I found a number of the elements identified by Hanna (1995), namely "assumption, conjecture, example, counterexample, refutation and generalization" (1995, p. 48). I will focus on three of these aspects -- conjecture, refutation, and generalization -- and show that aspects of proving arose spontaneously during the activities.

There is little in the research literature on proving in relation to young children, with the exception of the seminal work being done by the Maher team at Rutgers (for example, Maher & Martino, 1996) and Lampert (eg., 1990), and preliminary work by Jones (1994). Maher, and Lampert propose that involvement in inductive and deductive reasoning which leads to the construction of proofs should begin at the elementary school level. Gardiner suggests that "the 'form' and 'language' of the reasoning changes as learners grow older, but <u>not</u> the requirement that mathematical reasoning be: (a) general (that is, valid for all possible examples in the universe under consideration), and (b) completely convincing" (1992, p. 4). Gardiner also highlights the essential ingredient, that of the notion of infinity, saying: "To tame infinity we need proof" (1992, p. 10). De Villiers has spoken of the "reasoning which young children exhibit in situations which are real and **meaningful** to them" (1991, p. 254, boldface in the original). I take as my starting point the definition of 'proof as a convincing argument' (Hanna, Balacheff, & Pimm, 1991, p. xxxiii). I will show how the children at times exhibit careful reasoning, as they build their arguments and attempt to convince. With reference to Mason's (1982) statement that when you

[1]In E. Pehkonen (Ed.), *Proceedings of the Twenty-First International Conference for the Psychology of Mathematics Education (PME 21)* (Vol. 4, pp. 291-298), Lahti, Finland, July 14-19, 1997

prove, first you convince yourself, then convince a friend, and then convince an enemy, I will show instances within the children's interaction of 'convince yourself', and 'convince a friend' that you are correct. My goal as a teacher is nurturing a higher level of agency and autonomy for learners, and the ways in which the students ask their own questions, direct their own inquiry and engage in sustained conversation about generalizations and proving is the prime focus of the paper. The children were convinced ***that*** their pattern worked. It is my contention that in order for an argument to be considered a proof, the students have to not only *convince*, but also to *explain*. I will indicate, briefly, ways in which my involvement and personal inquiry were in this instance necessary to provoke thought concerning ***why*** the pattern works as it does .

The school community and classroom setting, and assigned tasks

St. George's is a private, non-denominational school, with a middle class population of mixed ethnic, religious, and linguistic backgrounds; the population is pre-dominantly English-speaking. The homeroom class size in the 1995-1996 year was 26; the work, however, is always done in half-groups (13 children in each group) of heterogeneous ability. Problem-solving is at the core of the mathematics curriculum in this classroom. The school and classroom learning site is a community of practice which Richards (1991) has called *inquiry math*; it is one in which the children are expected to publicly express their thinking, and engage in mathematical practice characterized by conjecture, argument, and justification (Cobb, Wood, & Yackel, 1993, p. 98). Of interest here is the intersection between the last-mentioned items, and proof.

Mathematics class periods are 45 minutes, and twice a week are extended to 90 minutes. In addition to the in-class problem-solving sessions, each week the children also work on one challenging problem at home. They are expected to record their work and reflect on their strategies in a Math Log which serves as the initial basis of their group discussions in class. In class much of the session is conducted by the children as they discuss the problem first with a partner, then in a group of four or five, and finally with the entire group of thirteen students.

The children are videotaped throughout the school year on a rotating basis as they work in their groups. In addition to the videotape records, data sources include focused observations, student artifacts (math logs), teacher-composed questions eliciting opinions (written responses), and retrospective interviews.

The mathematical context of the problem/discussion

The COUNT THE SQUARES task is a variant of the 'Chessboard' problem (see Mason et al, 1982; Anderson, 1996). The work was assigned as follows:

Task #1 (April 29, 1996) :

Find all the squares . (A four by four grid is on the left)

Can you prove that you have found them all?

After the conclusion of the discussion on Task #1 (discussion was held May 1), **Task #2** was given **(May 1)**:

What if . . . this were a 5 by 5 square? How many squares would you have?

The students wrote their explanation for their answer to Task #2 in their Math Logs at home (as with Task #1), and then in class I asked them discuss their answers to the 5 by 5 grid, to think about and discuss the questions below, and to go as far in their exploration as they were comfortable. The following extensions were posed:

What if this were a 10 by 10 square? How many squares would there be?

What if this were a 60 by 60 square? How many squares would there be?

I did **not** assign the tasks with a view to provoking discussion on proof. However, the children's interaction in that 1995-1996 year, and the focus on proof at a number of conferences (Canadian Math Ed Study Group, May 1996; PME 20 and ICME 8, July 1996) aroused my interest in proof, which in turn led to this analysis.

Children's perceptions of the term 'prove'

The word 'proof' has a wide range of meanings, from everyday usage to the idea of formal rigorous mathematical proof. Amongst the children within the mathematics classroom there are as well subtle and important differences in how they interpret the term 'prove'. You will note in Task #1 the request: '**Prove that you have found all the squares**.' In conversation with me about the ways in which the children responded, Tommy Dreyfus suggested that there were two kinds of *assertions* in the data (personal communication, Dec. 10, 1996). The first is in regard to a single case; a number of students prove by *checking* that their answer for the specific question is correct. For Task #1, the correct answer is 30 squares. *Point finale*--no need for proof. Others are seen to refer to a pattern which has been discerned and which a number of the students contend will continue forever. This second kind of assertion is a general statement, and thus in principle requires proof.

'Convince yourself'

Will's pattern, his conjectures and his testing of them drove much of the work in his team of three and in the group of five. At the very outset, when the task was first assigned, Will seemed to have an intuition of a pattern, as he is seen to look and ponder and suddenly say: "I know what to do, I know exactly what to do"; he then proceeds to write the 'criss-cross' pattern on the Math Log page (Figure 1). When meeting with his partners two days later, Will says to Lew and to Ross, his partners that day: "I was pretty sure there would be a pattern, so I was keeping my eyes open and I found one." He then says he hasn't tested it on a different size square yet, but "the chances are if it works for those it works for others." Lew, in response to seeing Will's pattern on the page, is very impressed. Will checks out the answer for the next size square (5×5), and finds that his conjecture is correct, and says: "So, it's basically the same. I never realized it [i.e., the pattern] would be so helpful." Will checks his work up to the 5×5, and then assumes that the pattern which has worked for up to 5 will continue to work in the same way; he is subsequently seen to act in accordance with this assumption.

criss-cross **differences**

Figure 1: Will's two patterns

At one point in the first day's discussion, Will notices a 'pattern of differences' (Figure 1), hence a second pattern. There were other approaches voiced by the group members such as for example Gord's sum of squares, and Ross's sum of the little squares.

It is important to emphasize that the 'criss-cross' pattern is one I rarely see; over the past 3 years, only two children (of the 75 students) have discerned it, realized that it might be significant, and then pursued it (Will, and Alan). Will's criss-cross pattern is the one I assumed would be focal. It lends itself to the summing of the squares. However, Will himself is seen to spend much time using the pattern of differences to arrive at the numbers which will be the addends for the answer to the 60×60 task. He is adding to get the next number, and adding to get the total. Will does not at that point seem to be attending to the 'squares' component. It is Gord who from time to time refers to the squares. It is Gord and Lew who see that the solution for the 60 by 60 could be arrived at by multiplying (deriving the squares) and then adding: 60×60 + 59×59 + 58×58 and so on. They declare with excitement: "We're a genius!" The "we're a genius" speaks both to thrill of discovery and to their acknowledgement of the fruitfulness of the collaborative endeavor.

'Convince a friend': An argument and three counter-arguments

In order to 'convince a friend', Will, Lew and Gord proceed by saying, and showing, that they have a pattern and that it works. They use what they know of the pattern to refute the proposal presented by another pair, Ross and Ted. My caution to the reader is that the summary which follows makes it all sound far too straightforward; the unfolding of the argumentation, the challenges posed therein by the 5 individuals, the discoveries and reconfirmations which occur during the 15-minute interaction are not represented here.

We take up the May 3rd discussion at the point at which Will, Lew, and Gord (the group of 3) meet with Ross and Ted (group of 2), and form the group of 5. Although their approaches to arriving at the answers have varied, the five peers have up to this point all been in agreement with the answers for the tasks up to the 10 by10. The answer for the number of squares in a 4 by 4 is 30, in a 5 by 5 is 55, and in a 10 by 10 is 385. It is during the discussion of the answer to the 60 by 60 question that there is a disagreement, and one group is challenged by the other to refute: "You have to prove us wrong."

As they begin their group of five discussion they agree that the answer for the 10 by 10 is 385 squares. The group of three states that they-- Will, Lew and Gord -- had not yet completed getting the answer for the 60 by 60. Ross and Ted feel that they have the solution for the 60 by 60, which is to take the 385 (the answer for the 10 by 10 square) and multiply it by 6 to get the number of squares for a 60 by 60; the resulting answer is 2310. [Of interest is to note that 10 of the 26 children in the class used this strategy.] Will and Lew are very sure that Ross and Ted are wrong:

L: I'll make you a bet.

W: I'll make you a bet.

L: I'll bet you anything in the world.

R: I'm not betting. You have to prove us wrong.

Will and Lew and Gord then proceed to use three counter-arguments, all based upon their generalizations, to refute what Ross and Ted have said. The first argument (Will, Lew) is that 3600 (the result of multiplying 60 by 60) is already bigger than 2310. Lew adds: "And that's just the little squares." One needs only one counter-example to disprove, but Ross and Ted are not seen at that point to concede. For the second argument Lew and Will create a generalization. They propose that if what Ross and Ted are saying is true, then it should work in general; and they then proceed to give a counter-example. Lew and Will consider the answers for the 4 by 4, and then the 8 by 8 square. They use the information to show that the answer for the 8 by 8 -- 204 squares-- is not simply double of the number of squares in the 4 by 4 -- 30 squares. The point they make is that just as one cannot multiply by 2 to get the answer, one cannot multiply 385 by 6 to get the answer for a 60 by 60. The third and last argument of the three arguments put forward is given by Will, and supported by

Gord; he points out that there is a pattern at work, and that doing a move such as taking 385 and multiplying that number times 6 means that one is not allowing the pattern to continue to grow, but that rather one is 'restarting' the whole pattern.

In their counter-arguments to Ross and Ted, Will and Lew (with Gord supporting Will and Lew) are in the end successful in refuting, and in convincing. Ted is heard to say: "Yeah, they're-, you guys are right, I go along with you guys." It occurs after the second argument has been presented, but it is the idea of the 3600 (the first argument) to which Ted directs his attention, and which he now assets is correct. Ross does not voice his agreement explicitly, as did Ted, but is later seen to support Lew and Will when, in the larger group, Lew and Will work to refute another team's presentation of the 385 times 6 strategy.

The findings seem to suggest that the students who succeed in convincing their peers (Will, Lew and Gord, in this instance, and others) are those whose justifications are based upon the generalizations. Will, Lew and Gord assert that the pattern must be adhered to. Will insists that it will continue forever: "You can use the pattern to calculate any number, even a googol times a googol." What Will, Lew and Gord do and say reflects their certainty that the pattern is correct in all cases. There is evidence of conviction prior to proving; their arguments are based upon their conviction *that* their pattern works in all instances. The pattern of summing the squares, i. e., $1^2 + 2^2 + 3^2 + \ldots$, is indeed correct. To be taken up next with the children, then, is proving in the sense of *explaining* the mathematical basis of the generalizations.

The teacher's role in pushing to explain *why it works*

One element which was not pursued by the students was that of explaining why the pattern works as it does. In investigating other patterns such an interest was present. (as in Zack, 1995; Graves & Zack, 1996). The absence might in part be due to the challenge inherent in this task. I myself was absorbed by the questions the task evoked. In follow-up interview and presentation sessions conducted seven months after the assignment was done, in December 1996, ten of the students re-immersed themselves in the task. What I was able to do then was to suggest that questions of *why* had to be addressed, and that I myself had many queries. I shared with them my own questions and my own search for illumination (Note 1). Thus, I provided a model of inquirer, of teacher as student (Freire). I built upon the elements which they had discovered and introduced to me, and showed how it related to some of what I had learned from other mathematics educators who had served as intermediaries for me: Bill Nevin, Sept. 24, 1996, and David Reid, Oct. 13, 1996 (personal communications) and John Mason (1982, pp. 18-21). I presented demonstrations of the explanations I had encountered, and also provided them with a "non-obvious expression" which I had found in a work by Anderson (1996, p. 35). I told the children that I myself did not understand how Anderson had arrived at that expression, nor why it worked. I did indicate to the children that there was a way to derive the formula, but that neither I nor they had the tools to do so. The discussion which ensued allowed a preliminary investigation of the children's criteria for proof.

All agreed that the pattern of summing the squares was intensely time-consuming. During the May 3 class time, one child had spent much time with his partner seeking a formula (Alan with Keiichi), with no success. (Please note that in the past he had had great success deriving algebraic expressions, as had others in the group as well). When shown the Anderson formula, n(n+1)(2n + 1)/6, the ten students interviewed felt that the Anderson formula would be useful, and economical. However, perhaps due to the emphasis in our classroom work that we have put upon explaining oneself, in their emergent definitions of what they felt proof ought to be, the students emphasized that their criteria for proof included: (a) a need for evidence, (b) that the proof must make sense, and (c) that the person presenting must say why it works. It was in written responses to the prompt "What do you think of Johnston Anderson's rule?" that the children expressed their positions. Ross, for example, stated that Johnston Anderson's rule was "brilliant, but he should explain why it works." Lew commented: "I think that if the Johnston rule had evidence, if Johnston himself explained why it worked it would be more convincing." Rina felt that Anderson's expression was "a great way to figure out the problem but it doesn't make sense . . . I think a mathematical proof is when you say why it works and if it works for everything show why." Only one child, Sanjay, did not voice a need for further elaboration, saying of the rule: "It's sort of like pi, it just works." Thus, despite finding Anderson's formula expedient, the majority of the students stressed that one ought to know why it worked as it did. In revisiting the problem, one of my objectives was to make the students aware of the importance of seeking to explain the mathematical structure. Hanna has asserted that in education proofs that explain should be favoured over those that merely prove (1995, p.48); the children as well are seen to seek proofs which explain. The criteria the children stated represent a healthy 'habit of mind' in our push to have learners think meaningfully about proof.

Note 1: My questions were: Why does it go from 9 2by2's in the 4×4 to 16 2by2's in the 5×5, etc.? Are there only 25 little squares added on when one moves from the 4×4 to the 5×5? Nevin showed me that it was 25 squares of different sizes, while Reid showed me that one could consider that only 25 1by1 squares appear which have never been there before; all the rest are expansions. Mason (1982, p.18-21) explained to me the derivation of the general formula $1^2 + 2^2 + 3^2 + \ldots$, relating it to the number of lines touched by the squares top to bottom, and side to side.

Acknowledgment: I am deeply indebted to colleagues who pursued various aspects with me: Barbara Graves, Tommy Dreyfus, David Reid, and Bill Nevin. This research was supported by a Social Sciences and Humanities Research Council grant from the Government of Canada #410-94-1627.

References:

Anderson, J. (1996). The place of proof in school mathematics. Mathematics Teaching 155, pp. 33-39.

Cobb, P., Wood, T., & Yackel, E. (1993). Discourse, mathematical thinking, and classroom practice. In E. Forman, N, Minick, & C. A. Stone (Eds.), Contexts for learning: Sociocultural dynamics in children's development (pp. 91-119). N. Y.: Oxford University Press.

De Villiers, M. (1991). Pupils' needs for conviction and explanation within the context of geometry. In Furinghetti, F. (Ed.), Proceedings of the Fifteenth Conference of the International Group for the Psychology of Mathematics Education (PME 15), Assisi, Italy, June 29-July 4, 1991, pp. 255-262.

Gardiner, T. (1992). Infinity: Recurring themes in school mathematics . UK: UK Mathematics Foundation, UK SMC, School of Mathematics, University of Birmingham, B15 2TT.

Graves, B., & Zack, V. (1996). Discourse in an inquiry math elementary classroom and the collaborative construction of an elegant algebraic expression. In Puig, L., & Gutiérrez, A. (Eds.), Proceedings of PME 20, Valencia, Spain, July 8-12, 1996, pp. 27-34.

Hanna, G. (1995). Challenges to the importance of proof. For the Learning of Mathematics, 15 (3), pp. 42-49.

Hanna, G., Balacheff, N., & Pimm, D. (1991) Theoretical and practical aspects of proof. (Working Group abstract). In Furinghetti, F. (Ed.), Proceedings of PME 15, Assisi, Italy, June 29-July 4, 1991, p. xxxiii.

Jones, L. (1994, November). Reasoning, logic and proof at key stage 2. Mathematics in School, pp. 6-8.

Lampert, M. (1990). When the problem is not the question and the solution is not the answer: Mathematical knowing and teaching. American Educational Research Journal 27 (1), 29-63.

Maher, C., & Martino, A. (1996). The development of the idea of mathematical proof: A 5-year case study. JRME, 27 (2), 194-214.

Mason, J., with Burton, L., & Stacey, K. (1982). Thinking mathematically. Addison-Wesley.

Richards, J. (1991). Mathematical discussions. In E. von Glasersfeld (Ed.), Radical constructivism in mathematics education. Kluwer.

Zack, V. (1995). Algebraic thinking in the upper elementary school: The role of collaboration in making meaning of 'generalisation'. In Meira, L., & Carraher, D. (Eds.), Proceedings of PME 19, Recife, Brazil, July 22-27, 1994, pp. 106-113.