Journal for Research in Mathematics Education
1989, Vol. 20, No. 1, 41–51

PROOF FRAMES OF PRESERVICE ELEMENTARY TEACHERS

W. GARY MARTIN, *University of Hawaii*
GUERSHON HAREL, *Northern Illinois University*

This study asked 101 preservice elementary teachers enrolled in a sophomore-level mathematics course to judge the mathematical correctness of inductive and deductive verifications of either a familiar or an unfamiliar statement. For each statement, more than half the students accepted an inductive argument as a valid mathematical proof. More than 60% accepted a correct deductive argument as a valid mathematical proof; 38% and 52% accepted an incorrect deductive argument as being mathematically correct for the familiar and unfamiliar statements, respectively. Over a third of the students simultaneously accepted an inductive and a correct deductive argument as being mathematically valid.

The concept of proof is of great importance in the study of mathematics. Smith and Henderson (1959) stated, for example, that "the idea of proof is one of the pivotal ideas in mathematics. It enables us to test the implication of ideas, thus establishing the relationship of the ideas and leading to the discovery of new knowledge" (p. 178). Research has explored this important topic with elementary and high school students (Bell, 1976, 1979; Galbraith, 1981; Lester, 1975; Williams, 1980). Two studies are most pertinent to our research. Fischbein and Kedem (1982) investigated whether high school students understand that mathematical proof requires no further empirical verification. They verified empirically their assumption that "students, after finding or learning a correct proof for a certain mathematical statement, will continue to consider that surprises are still possible, that further checks are desirable in order to render the respective statement more trustworthy" (p. 128). Vinner (1983) focused on the question, What makes a given sequence of correct mathematical arguments a mathematical proof in the eyes of high school students? He asked students to give their preference for proving a particular case of a previously proved statement. He found they preferred using a particularization of the deductive proof rather than the general result. The general proof was viewed as a method to examine and to verify a particular case. Vinner further observed that students judged a mathematical proof on its appearance, relying on ritualistic aspects of proof.

In this study, we investigated the views of proof of a different population, preservice elementary school teachers. Further, we investigated a different aspect of proof, one related to inductive and deductive reasoning.

The views of proof held by preservice elementary school teachers are important. Because proof receives very limited attention in the elementary school curriculum, the main source of children's experience with verification and proof is the classroom teacher. Classroom teachers' understanding of what constitutes mathematical proof is important, even though they do not directly teach that topic. If teachers lead their students to believe that a few well-chosen examples constitute proof, it is natural to expect that the idea of proof in high school geometry and other courses will be difficult for the students.

In everyday life, people consider "proof" to be practically synonymous with "what convinces me" (Smith & Henderson, 1959). Conviction, however, is a result of experience; Bell (1976) suggested that "conviction arrives most frequently as the result of the mental scanning of a range of items which bear on the point in question, this resulting eventually in an integration of the ideas into a judgment" (p. 24). Anderson (1985) identified this kind of proof as an "inductively valid argument," an argument whose conclusion is not necessarily true but only highly probable. People in everyday life form or evaluate hypotheses by estimating the probability of hypotheses with respect to their relevant individual experience (Anderson, 1985). Relevant individual experience includes evidence that supports or refutes the statement whose validity is in question.

The mathematical meaning of proof can be clearly distinguished from this everyday meaning. Anderson (1985) identified this second kind of proof as a "deductively valid argument," where the conclusion must be true if the premises are true.

The viewpoint that a mathematical proof must be a deductive argument is certainly held by mathematically sophisticated persons. However, our experience suggests that persons with limited experience in mathematics often hold the point of view that an inductive argument can also be a mathematical proof. It seems likely that the use of inductive arguments for proof in everyday life is translated by such persons to the acceptance and production of inductive arguments as proofs for mathematical statements; they accept and provide examples as a legitimate process of mathematical proof. Furthermore, this viewpoint may be reinforced by instruction in earlier grades, which frequently uses examples to verify mathematical statements. As these students encounter higher mathematics, at the high school and university level, instructors present deductive arguments as mathematical proofs and stress (at least implicitly) that inductive arguments do not constitute mathematical proofs. Our question is what conclusions preservice elementary school students draw about the role of inductive arguments and deductive arguments in mathematical proof.

The following questions guided our investigation:

1. Do preservice elementary school teachers accept inductive arguments as proofs of mathematical statements? Are their evaluations of inductive arguments dependent on their familiarity with the statement?

2. Are preservice elementary teachers more convinced by some types of inductive arguments than others?

3. Do preservice elementary school teachers accept that a deductive argument constitutes a mathematical proof? Are their evaluations of deductive arguments dependent on their familiarity with the statement?

4. Are students' judgments of an argument influenced by its appearance in the form of a mathematical proof—the ritualistic aspects of proof—rather than the correctness of the argument?

5. How do students view deductive arguments presented in the particular case, that is, mathematical proofs in which the parameters are changed to specific numbers?

6. Is the acceptance of inductive arguments and deductive arguments as mathematical proofs mutually exclusive?

A partial investigation of the first and third questions was reported in Harel and Martin (1986). In this report, we present our complete findings.

METHOD

Subjects

The views of proof held by 101 preservice elementary school teachers (hereafter referred to as *students*) enrolled in a required sophomore-level mathematics course at Northern Illinois University were assessed in the 10th week of the semester. The students had two substantial sources of contact with the idea of mathematical proof, a prerequisite high school course in geometry, and explicit attention to proof throughout the required mathematics course.

Instrument

Students were asked to judge verifications of a familiar and an unfamiliar mathematical generalization. The familiar generalization, along with its proof, had been discussed in the mathematics course 3 weeks before the instrument was administered. Although the unfamiliar generalization was appropriate for inclusion in the course, it was not explicitly discussed. The two generalizations follow:

- *Familiar generalization:* If the sum of the digits of a whole number is divisible by 3, then the number is divisible by 3.
- *Unfamiliar generalization:* If a divides b, and b divides c, then a divides c.

Among many possible verifications related to inductive arguments (Anderson, 1985), four common verifications related to inductive arguments were included in the instrument:

Examples. Two particular instances of the generalization involving small numbers were presented. One of the examples is given in Table 1.

Pattern. A chart containing a sequence of instances of the generalization was presented. A total of 12 instances were included in the chart; examples are included in Table 1. The pattern verification type suggests the message that one can generate as many examples as wanted in support of the general statement. This verification type was used only with the familiar statement; the unfamiliar would require the coordination of three sets of values to form a pattern consistent with the familiar case.

Big number. A particular instance of the generalization involving large numbers was presented. (See Table 1.) The big-number verification type suggests the

Table 1
Inductive Verification Types[a]

	Familiar context
Example[b]	The sum of the digits of 48 is 12, which is divisible by 3. The number itself is divisible by 3.
Pattern[c]	*Numbers divisible by 3* / *Sum of the digits of the number* / *Is sum divisible by 3?* 3 / 3 / YES
Big number	We will pick any number so the sum of its digits is divisible by 3, say 721234182. Is this number divisible by 3? *(Computation shown to left.)* The answer is affirmative.
Example and nonexample	31 is not divisible by 3, and we see that the sum of its digits is 4, which is not divisible by 3. On the other hand, 36 is divisible by 3, and the sum of its digits is 9, which is divisible by 3.
	Unfamiliar context
Example[b]	12 divides 36. 36 divides 360. 12 divides 360.
Pattern[c]	n.a.
Big number	Let's pick any three numbers, taking care that the first divides the second, and the second divides the third; 49 divides 98, and 98 divides 1176. Does 49 divide 1176? *(Computation shown to left.)* The answer is yes.
Example and nonexample	3 does not divide 5, and 5 does not divide 7. We see that 3 does not divide 7. On the other hand, 3 divides 6, and 6 divides 12. In this case, 3 also divides 12.

[a]Each item was presented on a separate page, double spaced, with each sentence on a separate line.
[b]Two examples were given for this verification type.
[c]The table contained verifications for 3, 6, 9, 12, 15, 18, 102, 105, 108, 1002, 1005, and 1008.

message that if the statement is true for an arbitrarily chosen large number, then it is probably true for other numbers.

Example and nonexample. Students were presented with an example supporting the general statement and a nonexample. (See Table 1.) The nonexample differed in form in the familiar and unfamiliar contexts. In the familiar, it was a nonexample of the general statement in the sense that if the conclusion does not hold, neither does the condition. In the unfamiliar, it was a nonexample of the general statement in the sense that if the condition does not hold, neither does the conclusion.

Three verifications related to deductive arguments were included in the instrument:

- *General proof.* A correct general proof of the statement was presented, including statements justifying each step. In the familiar case, we used the proof that had been presented in the mathematics class, which was limited to three-digit numbers. See Table 2.

- *False proof.* A fallacious proof of the generalization, including statements purporting to justify each step, was presented. (See Table 2.) Although not a deductive argument, it may be incorrectly viewed by students as a deductive argument based on its ritualistic aspects, as suggested by Vinner (1983).

• *Particular proof.* Students were presented with a correct proof of the generalization, including statements justifying each step, in which particular numbers were substituted for each of the variables. See Table 2.

Table 2
Deductive Verification Types[a]

	Familiar context
General proof	Let a be any three-digit number, with digits x, y, and z. By the place value concept, $a = 100x + 10y + z$. This equality can be written $a = (99x + x) + (9y + y) + z$. By the commutative and associative properties, we get $a = (99x + 9y) + (x + y + z)$. Notice that the expression $99x + 9y$ is always divisible by 9, and therefore also by 3. Now if the second expression, which is the sum of the number's digits, is divisible by 3, then, by the "sum property," we get that the number itself is divisible by 3.
False proof	Let a be any whole number such that the sum of its digits is divisible by 3. Assuming its digits are x, y, and z, then $a = xyz$. Since $x + y + z$ is divisible by 3, also xyz is divisible by 3. Therefore, a is divisible by 3.
Particular proof	Consider 756. This number can be represented as follows: $756 = 7{\times}100 + 5{\times}10 + 6$. This can be rewritten $756 = (7{\times}99 + 7) + (5{\times}9 + 5) + 6$. By the commutative and associative properties, we get $756 = (7{\times}99 + 5{\times}9) + (7+5+6)$. Notice that the expression $7{\times}99 + 5{\times}9$ is always divisible by 9, and therefore also by 3. Now if the second expression, which is the sum of the number's digits, is divisible by 3, then, by the "sum property," we get that the number itself is divisible by 3.
	Unfamiliar context
General proof	a divides b; this means there exists a number k, such that $k \times a = b$. Also, b divides c, which means there exists a number n, such that $n \times b = c$. Now, substitute for b in the last equation, and we get $n \times (k \times a) = c$. By the associative property, $(n \times k) \times a = c$. Therefore, a divides c.
False proof	Let a, b, and c be any whole numbers such that a divides b, and b divides c. Since a divides b, the last digit of a must be divisible by 3. Since b divides c, the last digit of b divides c. From the last two statements, we get that a divides c.
Particular proof	Take 4, 8, and 12. 4 divides 8, which means there must exist a number, in this case 2, such that $2 \times 4 = 8$. 8 divides 24, which means there must exist a number, in this case 3, such that $3 \times 8 = 24$. Now substitute for the 8 in the previous equation, and we get $3 \times (2 \times 4) = 24$. So we found a number, (3×2), such that $(3 \times 2) \times 4 = 24$. Therefore, 4 divides 24.

[a]Each item was presented on a separate page, double spaced, with each sentence on a separate line.

Procedure

Students were presented with each of the verifications for the two generalizations in a 30-minute test. Verifications for a given generalization were presented within a section. The order of related verifications was consistent between the two sections, but the order of presentation of the two sections was randomized. Each verification was presented along with the statement of the generalization it purported to prove correct. Students were instructed to rate whether each verification of each statement was a valid mathematical proof on a four-point scale, where 4 indicated that they considered a verification to be a mathematical proof and 1 indicated that they did not consider it a mathematical proof. The four-point scale allowed us to split responses for later analyses into high and low categories, where responses of 3 and 4 were scored as a high level of acceptance and responses of 1 and 2 were scored as a low level of acceptance.

RESULTS

Frequencies of ratings for each of the verifications of the two statements are presented in Table 3. We consider the results relating to each of the research questions in turn.

Table 3
Frequencies of Ratings for Verifications of the Familiar and Unfamiliar Generalizations

Verification	Rating				χ^2 (df = 1)[a]
	1	2	3	4	
	Inductive verification types				
Familiar					
Example	13	23	25	40	8.33**
Big number	25	20	22	34	1.20
Nonexample	16	16	32	37	13.55**
Pattern	11	15	30	45	23.77**
Unfamiliar					
Example	16	20	26	39	8.33**
Big number	18	18	27	38	8.33**
Nonexample	18	25	26	33	2.23
	Deductive verification types				
Familiar					
General proof	3	23	22	53	23.77**
False proof[b]	26	22	27	25	0.09
Particular proof	12	30	25	34	2.86
Unfamiliar					
General proof	14	24	35	28	6.19*
False proof[b]	33	29	31	7	6.19*
Particular proof	19	27	33	22	0.80

[a] χ^2 computed on the high-low split.
[b] One student did not provide a rating to this item.
*$p < .05$. **$p < .01$.

Acceptance of inductive arguments. The distribution of the categories for the inductive verifications for both unfamiliar and familiar contexts is shown in Table 3. Each of the inductive verifications was rated high (i.e., as 3 or 4) by more than 50% of the students. In five of the seven cases, a significantly larger number of students rated the verification high than rated it low.

In both familiar and unfamiliar contexts, 80% of the students gave a high validity rating (3 or 4) to at least one inductive argument, and over 50% gave a very high validity rating (4) to at least one inductive argument. Fewer than 10% gave a very low validity rating (1) to all four inductive arguments.

As seen in Table 4, no significant differences were found in student acceptance of inductive arguments for the familiar statement and the corresponding arguments for the unfamiliar statement.

Table 4
χ^2 Statistics Comparing Frequencies of Acceptance of Verifications of Familiar and Frequencies of Acceptance of Unfamiliar Statements

Verification type	χ^2 (df = 1)
Inductive arguments	
Example	0.00
Big number	1.33
Nonexample	1.95
Deductive arguments	
General proof	2.77
False proof	2.87
Particular proof	0.15

Differences in acceptance of the different kinds of inductive arguments were not significant for the unfamiliar statement, χ^2 (2, N = 101) = 1.22, n.s. These differences were significant for the familiar statement, χ^2 (3, N = 101) = 8.95, $p < .05$. However, when the Pattern verification, unique to the familiar statement, was omitted from the analysis for the familiar statement, the differences were no longer significant, χ^2 (2, N = 101) = 3.22, n.s.

Acceptance of deductive arguments. Many students rated deductive arguments high (3 or 4) in both the familiar and unfamiliar contexts, as can be seen by student acceptance of the general-proof verification type in Table 3; significantly more students were categorized as High Deductive than as Low Deductive in both contexts. The level of acceptance of the general proof was not significantly different between the two statements; see Table 4.

Ritualistic aspects of proof. We explored the ritualistic aspects of proof by examining responses to the false-proof verification. On the one hand, if students reacted to the form of the proof, they should have rated false proofs high; we call such students High Ritualistic. On the other hand, students reacting to a logical evaluation of the proof should have rated false proofs low; we call such students Low Ritualistic. Because we are attempting to better understand how students view deductive arguments, we limit our report of results concerning the ritualistic aspects of proof to students who are categorized as High Deductive. Frequencies of categorizations for each statement are given in Table 5. Significantly more High Deductive students were rated Low Ritualistic than High Ritualistic for the unfamiliar context. Equal numbers of High Deductive students were rated High Ritualistic and Low Ritualistic in the familiar context.

Table 5
Frequencies of Categorizations of Ritualistic Aspects of Proof in High Deductive Students, in Familiar and Unfamiliar Contexts

	Categorization of ritualism		
Context	High	Low	χ^2 ($df = 1$)
Familiar	38	37	0.01
Unfamiliar	21	41	6.45*

*$p < .05$.

The role of particular proof. We further evaluated students' views of the role of deductive arguments in mathematical proof using their ratings of the particular-proof verification type. We limited our analysis to those students who are High Deductive and Low Ritualistic, since the other students had demonstrated a limited understanding of the semantics of deductive arguments. Furthermore, this allowed us to be comparable with previous studies. As seen in Table 6, more students accepting correct deductive arguments also accepted an argument presented in the particular case than rejected it. This result is statistically significant in the familiar case. Although the results in the unfamiliar case are not statistically significant, the trend is in the same direction.

Table 6
Percents of Categorizations of Particular Proof Found in High Deductive, Low Ritualistic Students

	Particular proof categorization		
Context	High	Low	χ^2 ($df = 1$)
Familiar	26	11	6.08*
Unfamiliar	23	18	0.61

*$p < .05$.

Relationship of acceptance of inductive and deductive arguments. Many students simultaneously rated both inductive arguments and deductive arguments high, in both the familiar and unfamiliar case, as seen in Table 7. Over 46% simultaneously rated general proof and at least one of the inductive-argument verification types high.

DISCUSSION

In this section, we will answer the six research questions and suggest an interpretation using frames (Minsky, 1975).

We found that many students accepted inductive arguments as proofs of mathematical statements, and this acceptance was not dependent on the familiarity of the context. Moreover, they were not more convinced by some types of inductive arguments than others. Similarly, we found that many students accepted deductive arguments as proofs of mathematical statements, and this acceptance also was not dependent on the familiarity of the context. This led us to postulate that in-

Table 7
Relationship of Acceptance of Inductive and Deductive Arguments for Familiar and Unfamiliar Statements

Rating of deductive argument	Rating of inductive arguments[a] High	Rating of inductive arguments[a] Low
Familiar statement		
Low	2	24
High	19	56
Unfamiliar context		
Low	5	34
High	16	46

[a] Subjects were rated in the high category for inductive arguments if they rated any of the inductive arguments as 3 or 4. Subjects were rated in the low category only if they rated all inductive arguments as 1 or 2.

ductive arguments and deductive arguments represent two proof frames, constructed by students as the result of experience in everyday life and in the mathematics classroom.

Acceptance of inductive and deductive arguments as mathematical proofs was not found to be mutually exclusive. This suggests that the inductive frame, which is constructed at an earlier stage than the deductive frame, is not deleted from memory when students acquire the deductive frame. Moreover, the everyday experience of forming and evaluating hypotheses by using evidence to support or refute them serves to reinforce the inductive frame. Thus, as our results indicate, inductive and deductive frames exist simultaneously in many students. The findings of Fischbein and Kedem (1982) suggest a further relationship between these two frames. They found that many students who were convinced by deductive proof still wanted further empirical verification. This suggests that the activation of both the inductive and the deductive proof frames may be required for students to believe a particular conclusion.

Many students who correctly accepted a general-proof verification did not reject a false-proof verification; they were influenced by the appearance of the argument—the ritualistic aspects of the proof—rather than the correctness of the argument. We can interpret this finding in terms of frames by postulating the existence of different levels of the deductive frame. Such students appear to rely on a syntactic-level deductive frame in which a verification of a statement is evaluated according to ritualistic, surface features. Alternatively, relatively few students have a conceptual-level deductive frame in which a judgment is made according to causality and purpose of the argument.

A further aspect of our results relates to the role of a particular proof. We found that students who correctly accepted a general-proof verification also showed high levels of acceptance of a particular proof. Two explanations can be suggested for this result. First, students may be interpreting a particular proof as an inductive argument, in which case it is seen as an instantiation of the inductive argument frame. This seems unlikely; the same pattern of high levels of acceptance of a particular proof was found in students rating inductive arguments low. Second, a particular proof may be viewed as an instantiation of the deductive process used in the general proof. In this case, they are "replaying" the general deductive argument in the particular case. Further evidence for this interpretation can be found in Vinner (1983). Vinner asked students to give their preference for proving a particular case of a previously proved statement. He found they preferred using a particularization of the deductive process used to prove the statement rather than the general result. We thus suggest there are two subframes of the conceptual-level deductive frame, a generalized-results subframe and a generalized-process subframe.

Our results suggest the model of the frames of proof held by students shown in Figure 1. The solid arrows denote a frame-subframe relationship, and the broken arrows denote coexisting frames.

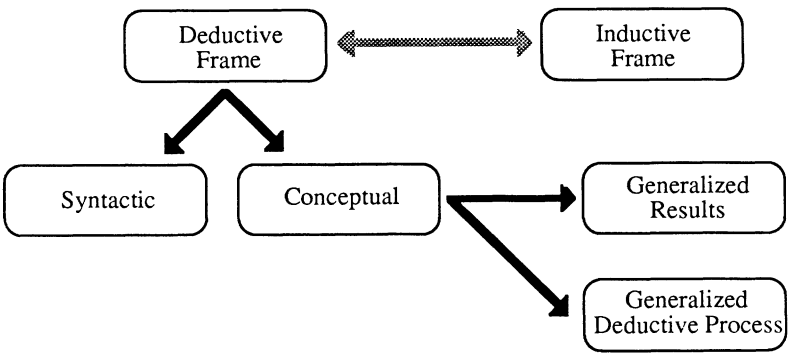

Figure 1. A schematic of students' proof frames.

CONCLUSIONS

Limitations and suggestions for further research

Several limitations of this study follow. First, our interpretations of prospective teachers' views of proof are inferred from written responses. Individual interviewing may allow deeper understanding of these phenomena. We suggest that studies in this mode be undertaken to validate our findings and to explicate more fully the explanations we have suggested. Second, our questions concern the evaluation of existing verifications rather than the production of verifications. Students may not act the same in a production mode as in an evaluation mode. It may well be that different knowledge structures are activated in the two settings. A full understanding of how students think about the role of empirical evidence and the logical justification of mathematical proofs requires attention to this crucial aspect.

Implications

The prospective teachers in the course received extensive and explicit instruction about the nature of proof and verification in mathematics. The message received was not that presumed to have been transmitted. We suggest that preservice teachers may require more experience with the place of empirical evidence in mathematics. Perhaps more examples of the limitations of empirical evidence and more examples of the power of mathematical proof need to be provided. Further attention to the similarities and differences of mathematics to everyday life may also be helpful.

REFERENCES

Anderson, J. R. (1985). *Cognitive psychology and its implications* (2nd ed.). New York: W. H. Freeman.

Bell, A. W. (1976). A study of pupils' proof-explanation in mathematical situations. *Educational Studies in Mathematics, 7,* 23–40.

Bell, A. W. (1979). The learning of process aspects of mathematics. *Educational Studies in Mathematics, 10,* 361–387.

Fischbein, E., & Kedem, I. (1982). Proof and certitude in the development of mathematical thinking. In A. Vermandel (Ed.), *Proceedings of the Sixth International Conference for the Psychology of Mathematics Education* (pp. 128–131). Antwerp: PME.

Galbraith, P. L. (1981). Aspects of proving: A clinical investigation of process. *Educational Studies in Mathematics, 12,* 1–28.

Harel, G., and Martin, W. G. (1986). The concept of proof held by preservice elementary teachers. In *Proceedings of the Tenth International Conference for the Psychology of Mathematics Education* (pp. 386–391). London: University of London Institute of Education.

Lester, F. K. (1975). Developmental aspects of children's ability to understand mathematical proof. *Journal for Research in Mathematics Education, 6,* 14–25.

Minsky, M. (1975). A framework for representing knowledge. In P. H. Winston (Ed.), *The psychology of computer vision.* New York: McGraw-Hill.

Smith, E. P., and Henderson, K. B. (1959). Proof. In P. S. Jones (Ed.),*The growth of mathematical ideas, grades K–12* (24th Yearbook of the NCTM, pp. 111–181). Washington, DC: NCTM.

Vinner, S. (1983). The notion of proof: Some aspects of students' view at the senior high level. In R. Hershkowitz (Ed.), *Proceedings of the Seventh International Conference for the Psychology of Mathematics Education* (pp. 289–294). Rehovot, Israel: Weizmann Institute of Science.

Williams, E. (1980). *An investigation of senior high school students' understanding of the nature of mathematical proof.* Unpublished doctoral dissertation, University of Alberta, Edmonton..

AUTHORS

W. GARY MARTIN, Assistant Professor, College of Education, University of Hawaii, Honolulu, HI 96822

GUERSHON HAREL, Assistant Professor, Department of Mathematical Sciences, Northern Illinois University, Dekalb, IL 60115