Understanding the Connections Between

Students' Ways of Reasoning In and Out of Mathematics

Amy Ellis, Charles Kalish, Eric Knuth (University of Wisconsin-Madison),

and Caroline Williams (Indiana University-Bloomington)

An enduring challenge in mathematics education is finding ways to help students understand the nature of evidence and justification in mathematics (Kloosterman & Lester, 2004). Consequently, mathematical reasoning – proof, in particular – has received increased attention in the mathematics education community. Researchers and reform initiatives alike are advocating a shift towards emphasizing proof and justification in the mathematics education of students at all grade levels (e.g., Ball, Hoyles, Jahnke, & Movshovitz-Hadar, 2002; National Council of Teachers of Mathematics, 2000; Knuth, 2002a, 2002b; RAND Mathematics Study Panel, 2002; Yackel & Hanna, 2003). Proof plays a critical role in promoting deep learning in mathematics (Hanna, 2000), but despite the growing emphasis on proof in school mathematics, research continues to paint a bleak picture of students' abilities to reason mathematically (e.g., Dreyfus, 1999; Healy & Hoyles, 2000; Knuth et al., 2002; Martin et al., 2005).

In contrast, cognitive science research has revealed surprising strengths in children's abilities to reason inferentially in non-mathematical contexts (e.g., Gelman & Kalish, 2006; Gopnik, et al., 2004). In particular, there is growing evidence that children are capable of developing sophisticated causal theories, and of using powerful strategies of inductive inference when reasoning about the natural world. This contrast raises a potential paradox: Why are children so adept at reasoning in non-mathematical contexts, yet seemingly so poor at reasoning in mathematical contexts?

The purpose of this study is to explore this contradiction by extending the cognitive science research into the domain of mathematics education. In the world outside the mathematics classroom, children typically develop facts and ideas via empirical generalizations and causal theories. Thus it is reasonable that students may rely on and make connections to their non-mathematical ways of reasoning as they encounter ideas and problems in mathematics. The main objective motivating this initial study was to examine the different types of inductive reasoning strategies students employ in both non-mathematical and mathematical contexts, and to elaborate the connections between those strategies.

Re-Conceptualizing the Challenge of Context: Linking Cognitive Science and Mathematics Education

Researchers acknowledge that a major source of students' difficulties with proof is represented by "the coexistence of formal and intuitive aspects, which materialise for example in the transitions from empirical to theoretical practices, from intuition to deduction, etc" (Sutherland, Olivero, & Weeden, 2004, p. 266). Research addressing this disconnect between formal and intuitive modes of justification, however, has primarily occurred within the domain of mathematics reasoning even while recognizing the highly contextual nature of students' understanding of proof (Fried & Amit, 2006). We propose a different approach: rather than viewing context as a source of difficulty, we consider it a potential source of knowledge that could help researchers better understand students' ways of reasoning both in and out of school mathematics. Indeed, lessons learned from the cognitive science literature suggest that empirical or inductive inference strategies can be viewed as a strength rather than a weakness – one that can be leveraged to foster the development of students' mathematical ways of reasoning.

Research in cognitive science supports the conclusions from the educational literature: People struggle with formal inference and deductive arguments. Studies of logical reasoning consistently show that people often fail to apply or even recognize formal inference strategies, instead using their empirical knowledge of the kinds of instances they have encountered (Evans, 2005; Markovits & Handley, 2005). From a traditional perspective, empirically-based responses in deductive reasoning tasks indicate a weakness or lack of understanding (Nisbett, 1993; Stanovich, 2004). However, this deficit interpretation is now being challenged. Researchers argue that people may be applying valuable, even normative, strategies but interpreting problems as empirical rather than logical (Oaksford, Chater, & Larkin, 2000). As a result, some advocate for a shift from viewing people as reasoning poorly about deductive problems to viewing them as reasoning well about inductive problems. This perspective dovetails with recent research emphasizing children's competence in inductive reasoning. Young children have powerful theories to predict and explain empirical phenomena, and have been shown to employ sophisticated learning strategies to develop those theories and evaluate evidence (Gopnik et al., 2004; see Gelman & Kalish, 2006 for review).

In mathematics education, inductive strategies have similarly been treated as a stumbling block to overcome rather than objects of study in their own right. Borrowing from recent work in cognitive science, we advocate for an alternative perspective that views inductive inference as a potentially powerful strategy that students bring to the task of learning mathematics. While the focus of past research has been distinguishing between empirical/inductive and formal/deductive justifications, the question of what makes one empirical justification better than another has not been well addressed. In a recent paper Christou and Papageorgiou (2007) argued that the skills involved in induction, such as "comparing" or "distinguishing", were similar in mathematical and non-mathematical domains. This prior research represents an important precondition for the research described in this study. Christou and Papageorgiou's work showed that students can identify similarities among numbers, distinguish non-conforming examples, and extend a pattern to include new instances. We focus on identifying how students may use those abilities when evaluating both mathematical and non-mathematical conjectures.

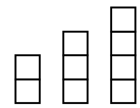
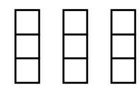
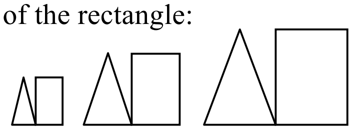
Methods and Data Sources

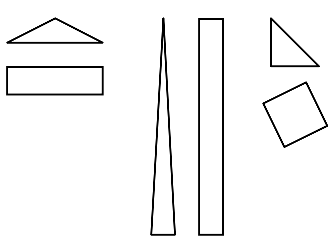
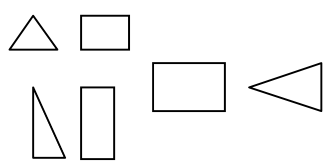
Forty-seven middle-school students (1 6th-grade student, 26 7th-grade students, and 20 8th-grade students) participated in a paper-and-pencil survey. The survey presented students with statements about number, geometry, and the natural sciences, each of which was paired with five different types of evidence as possible support for the statement. The survey items asked students to indicate how convincing each piece of evidence was by rating them individually and ranking them in relation to one another. Each rating occurred on a 4-point Likert scale ranging from "not at all convincing" to "fully convincing". Each survey contained irrelevant, random (arbitrarily picked examples), and deductive evidence items, as well as a single set of paired theme items: Similar versus dissimilar evidence; typical versus atypical evidence; and 1 example versus 5 examples. Based on past research, we expected typical examples, dissimilar examples, and more (5) examples to represent strong evidence; while atypical, similar, and few (1) examples would be weak evidence. Similarity and typicality was established in a pilot study with 22 undergraduate students. Participants in the pilot study rated the similarity and typicality of numerical, geometric and animal items used in the survey. Example survey items are presented below.

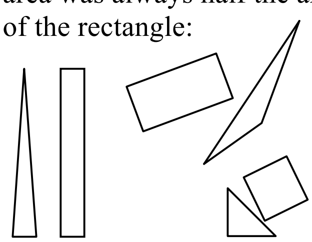
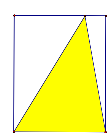
	Similar vs. Dissimilar	**Typical vs. Atypical**	**Deductive**
Number	Bennett says, "If you add any three consecutive numbers, the sum is equal to three times the middle number."		

<table>
<tr>
<td></td>
<td>Similar: Beth added three sets of consecutive numbers that she thought all looked pretty similar, and it worked each time:
$23 + 24 + 25 = 72$, and
$3 \cdot 24 = 72$
$12 + 13 + 14 = 39$, and
$3 \cdot 13 = 39$
$32 + 33 + 34 = 99$, and
$3 \cdot 33 = 99$

Dissimilar: Franz added three sets of consecutive integers that were different, and it worked every time:
$1 + 2 + 3 = 6$, and
$3 \cdot 2 = 6$
$4000 + 4001 + 4002 = 12003$, and $3 \cdot 4001 = 12003$
$1076 + 1077 + 1078 = 3231$, and $3 \cdot 1077 = 3231$</td>
<td>Typical: Franz added three sets of consecutive integers, and it worked every time:
$7 + 8 + 9 = 24$, and
$3 \cdot 8 = 24$
$113 + 114 + 115 = 342$, and
$3 \cdot 114 = 342$
$54 + 55 + 56 = 165$, and
$3 \cdot 55 = 165$

Atypical: Victor added three sets of consecutive numbers, and it worked each time:
$99 + 100 + 101 = 300$, and
$3 \cdot 100 = 300$
$299 + 300 + 301 = 900$, and
$3 \cdot 300 = 900$
$1999 + 2000 + 2001 = 6000$, and $3 \cdot 2000 = 6000$</td>
<td>Deductive: Jo explains that Bennett is right, and says she can show you with blocks. She starts with 9 blocks, and puts them in three stacks:

$2 + 3 + 4 = 9$
Then Jo takes one block from the 4 stack, and puts it on the 2 stack. Now all the stacks have 3 blocks:

$3 + 3 + 3 = 9$
Jo says that you can always take one block from the biggest of the stacks, and put it on the littlest to maket this work.</td>
</tr>
<tr>
<td>Geometry</td>
<td colspan="3">Timon says, “The area of a triangle is always exactly half the area of a rectangle with the same base and height.”</td>
</tr>
<tr>
<td></td>
<td>Similar: Rob looked at the following three pairs of triangles and rectangles and in each case the triangle’s area was always half the area of the rectangle:


Dissimilar: Courtney looked at the following three pairs of triangles and rectangles, and in each case, the triangle’s area was always half the area of the rectangle:
</td>
<td>Typical: Mike looked at the following three pairs of triangles and rectangles, and in each case, the triangle’s area was always half the area of the rectangle:


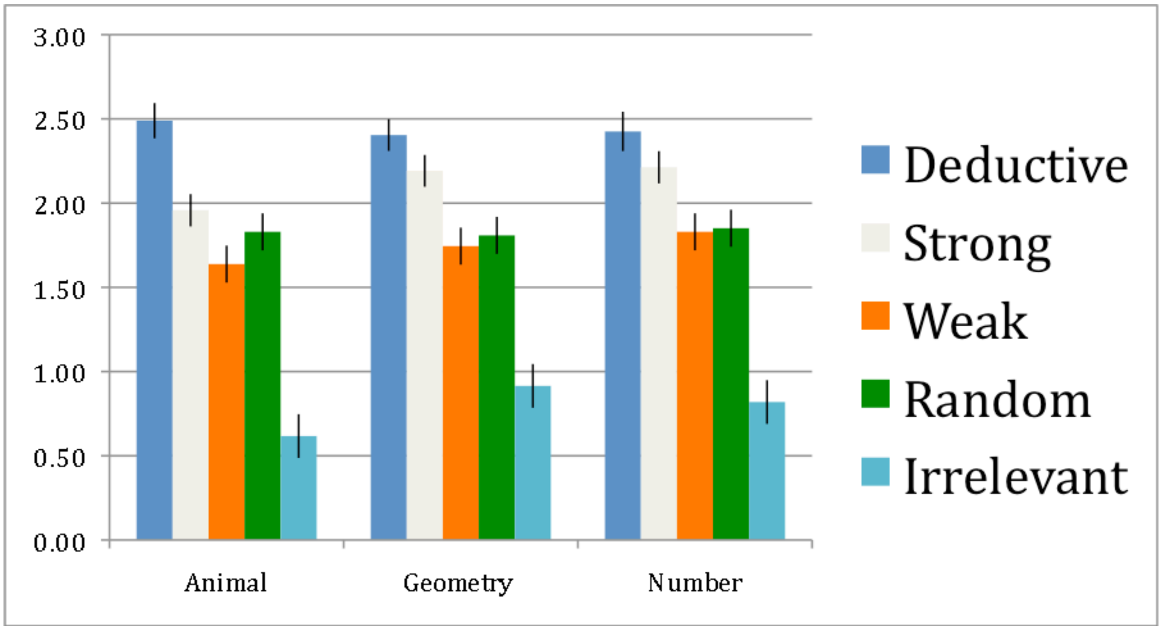
Atypical: Mike looked at the following three pairs of triangles and rectangles, and in each case, the triangle’s area was always half the area of the rectangle:
</td>
<td>Deductive: Valerie argued that a triangle has to have half the area of a rectangle with the same base and height, because you can take any triangle of any shape, and make a rectangle that has the same base and height like this:

Since the area of a rectangle is $b \cdot h$, and the area of a triangle is $\frac{1}{2} b \cdot h$, then any triangle will have an area that is ½ the rectangle.</td>
</tr>
</table>

Animal	Theres says, “All insects have chitinous skeletons.”		
	Similar: Alexis looked at ants, roaches, and beetles because, she said, they’re all pretty similar. She found that all three have chitinous skeletons. *Dissimilar*: Virginia looked at ants, wasps, and butterflies because, she said, they’re all different from each other. She found that they all have chitinous skeletons.	*Typical*: Juan looked at ants, beetles, and grasshoppers and found that they all have chitinous skeletons. *Atypical*: Ming looked at scale nymphs, mantises, and stickbugs and found that they all have chitinous skeletons.	*Deductive*: Chuck explained that chitin is the substance that forms a hard exoskeleton in all “arthropods” (or invertebrates with hard exteriors). Since all insects are defined as arthropods, all insects have chitinous skeletons.

The aggregate results] were examined across the three areas (number, geometry, and natural science) and the nine qualities of evidence (irrelevant item, similar items, dissimilar items, typical items, atypical items, 1 example, 5 examples, random evidence item, and deductive item) in order to ascertain general themes and reasoning tendencies.

Results

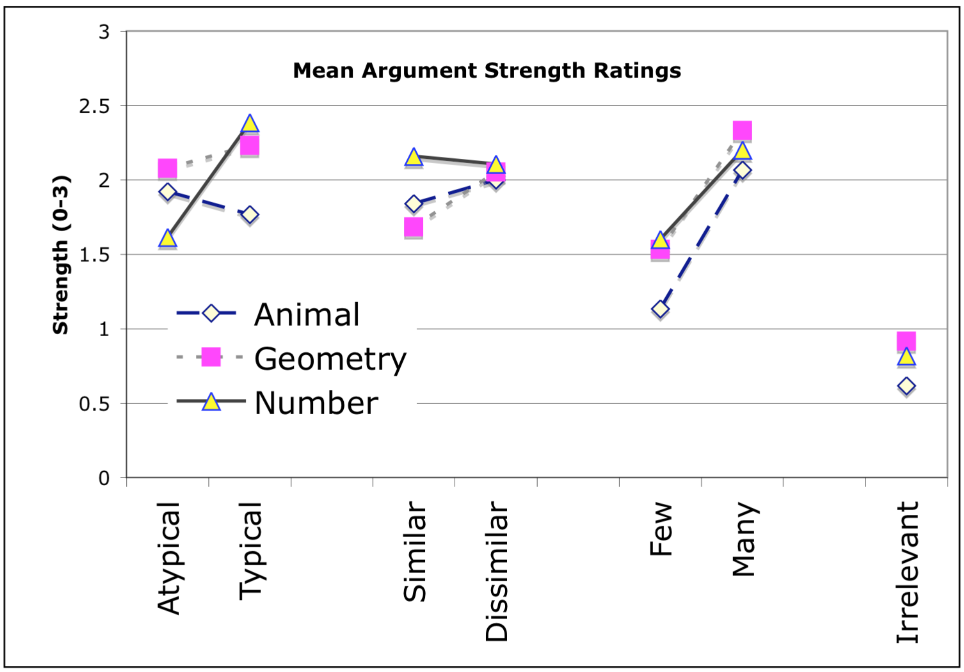
The tasks provided in the paper-and-pencil surveys were within the capacity of the middle school-aged students, as evidenced by the fact that irrelevant examples were judged to be weaker than any of the relevant ones. In addition, across all three domains the deductive arguments were judged to be the most convincing (see the figure below):



The first set of analyses considered the influence of strong and weak evidence on students' evaluations of arguments. In general, patterns across the mathematical and non-mathematical domains were similar: Example-based arguments were intermediate between deductive arguments and irrelevant ones. Arguments with stronger evidence (better examples) were rated higher than were those with weak evidence. However, there were some differences between mathematical and non-mathematical arguments. Only for Animal items were strong example-based arguments judged significantly less convincing than deductive arguments, $t(46) = 3.95$, $p<.001$. In neither the Number or Geometry cases were deductive arguments significantly stronger than the strongest example-based ones, both $t(46)<1.7$.

In general, students were more convinced by weak evidence in mathematical than in non-mathematical domains. Weak evidence for number items resulted in stronger arguments than did weak evidence for animal arguments, $t(46) = 2.9$, $p<.005$ (the difference between geometry and animal items did not reach statistical significance). Irrelevant examples resulted in stronger arguments for mathematical items than for animal items, $t(46) = 2.5$, $p<.05$.

The next step in analyses was to consider the different types of strong and weak evidence. One principle of evidence appeared clearly and consistently: More examples were better than fewer. In contrast, the results for similarity and typicality were mixed. As the table below demonstrates, students were more convinced by typical evidence for number and (weakly) geometry items, but they found typical evidence less convincing for the animal items. Moreover, for the geometry and animal items, dissimilar evidence was more convincing than similar evidence, but this was not the case for the number items.



We also found some evidence of domain differences. For the evidence quality we were able to manipulate with the most confidence (amount of evidence), fewer than 2/3 of the participants showed a distinction for mathematics items (rating arguments with more examples more highly than those with few examples), while 80% showed the distinction for animal items. This finding for the mathematics items was surprising, given our assumption that more evidence would be more convincing. In addition, the distance between the deductive and items-based evidence for the mathematics items was less than the distance between the deductive and item-based evidence for the non-mathematics items, which is the opposite of what one might hope given the deductive nature of mathematical reasoning, but is unsurprising given similar results from previous studies (e.g., Martin & Harel, 1989).

Finally, we identified one rather surprising element of similarity across domains: Randomly selecting examples to test was seen as a weak strategy in the mathematical domains, but a strong strategy in the animal domain. For example, arguments with three randomly selected examples were rated significantly higher than those with a single example for animals, $t(14) = 4.4$, $p<.001$, but not for number or geometry, both $t(14) < 1.6$. Random evidence resulted in significantly weaker arguments than did strong evidence for mathematical items, number: $t(46) = 3.2$; geometry: $t(46) = 2.8$, both $p<.01$. Arguments with random evidence and strong evidence were not different for animal items.

Discussion

Although researchers, educators, and reform documents all highlight the need to emphasize deductive proof and justification at all levels of mathematics, our participants' responses indicated that they did not view the mathematical items as a domain in which the deductive arguments held a particularly compelling role. Justifications based on many items and typical items were rated about equally strong as justifications based on deductive arguments. In fact, deductive justifications were rated as more compelling than examples-based justifications only in the non-mathematical domain – this result is the very opposite of what we would hope to see. One potential reason for this finding could be that the deductive arguments in the mathematics contexts were not accessible to the participants, whereas in the animal contexts they could make sense of those arguments. Critically, we did find that students could make distinctions in the value of different types of justification, employing principles defining stronger and weaker arguments. These principles are generally consistent with the inductive inference strategies observed outside of mathematical contexts. This suggests that students may employ powerful and useful reasoning strategies outside of mathematics that they could bring to bear in the mathematics contexts.

We found some differences in students' responses across the number and geometry contexts, which suggests that domain differences within mathematics may affect how students reason both inductively and deductively. Typically, a given subject area in mathematics is viewed merely as the background against which researchers examine students' understanding of proof. These preliminary data suggest that mathematical domain may play a critical role in influencing students' proof-based reasoning, and students may actually employ different inductive and/or deductive strategies when reasoning with, say, number theories than when reasoning about objects or relations in algebra or geometry.

Our results suggest that students may employ reasoning strategies outside of mathematics that they could bring to bear in mathematics contexts. The similarities and differences identified in students' reasoning about mathematics and non-mathematics items suggest that we may be able to identify productive ways of reasoning that are unique to, say, number contexts or natural phenomena contexts. The study reported here was an initial pilot study that provided preliminary data for supporting the design and implementation of a series of follow-up studies examining students' conceptions of what constitutes convincing evidence in a variety of domains. Better understanding what students find convincing in multiple contexts could enable teachers to help them transition to understanding, appreciating, and producing deductive arguments in mathematics. Students do have sensible ways of identifying stronger and weaker evidence-based arguments. Rather than dismissing these strategies, we intend to explore the ways in which instruction might build on students' intuitions to develop an appreciation of the unique value of proof.

References

Ball, D., Hoyles, C., Jahnke, H., & Movshovitz-Hadar, N. (August, 2002). *The teaching of proof.* Paper presented at the International Congress of Mathematicians, Beijing, China.

Christou, C., & Papageorgiou, E. (2007). A framework of mathematics inductive reasoning. *Learning and Instruction*, *17*, 55–66.

Dreyfus, T. (1999). Why Johnny can't prove. *Educational Studies in Mathematics*, *38*(1/3), (pp. 85–109).

Evans, J. S. B. T. (2005). Deductive reasoning. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning*, (pp. 169–184). New York, NY: Cambridge University Press.

Fried, M., & Amit, M. (2006). The Israeli classroom: A meeting place for dichotomies. In D. Clarke (Ed.), *Mathematics classrooms in 12 classrooms: The insiders' perspective*, (pp. 209–220). Rotterdam, The Netherlands: Sense Publishers.

Gelman, S. A., & Kalish, C. W. (2006). Conceptual development. In D. Kuhn, R. S. Siegler, W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology: Vol 2, cognition, perception, and language (6th ed.).* (pp. 687–733). Hoboken, NJ: John Wiley & Sons Inc.

Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: Causal maps and bayes nets. *Psychological review*, *111*, 3–32.

Hanna G. (2000). A critical examination of three factors in the decline of proof. *Interchange*, *31*(1), 21–33.

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

Kloosterman, P., & Lester, F. (2004). *Results and interpretations of the 1990 through 2000 mathematics assessments of the National Assessment of Educational Progress*. Reston, VA: National Council of Teachers of Mathematics.

Knuth, E. (2002a). Secondary school mathematics teachers' conceptions of proof. *Journal for Research in Mathematics Education*, *33*(5), 379–405.

Knuth, E. (2002b). Teachers' conceptions of proof in the context of secondary school mathematics. *Journal of Mathematics Teacher Education*, *5*(1), 61–88.

Knuth, E., Choppin, J., Slaughter, M., & Sutherland, J. (October, 2002). Mapping the conceptual terrain of middle school students' competencies in justifying and proving. *Proceedings of the Twenty-fourth Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education,* 1693–1700.

Markovits, H., & Handley, S. (2005). Is inferential reasoning just probabilistic reasoning in disguise? *Memory & Cognition*, *33*, 1315–1323.

Martin, T. S., McCrone, S. M. S., Bower, M. L., & Dindyal, J. (2005). The interplay of teacher and student actions in the teaching and learning of geometric proof. *Educational Studies in Mathematics*, *60*(1), 95–124.

Martin, W. G.,& Harel, G. (1989). Proof frames of preservice elementary teachers. *Journal for Research in Mathematics Education, 20*(1), 41-51.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics.* Reston, VA: Author.

Nisbett, R. E. (1993). *Rules for reasoning*. Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

Oaksford, M., Chater, N., & Larkin, J. (2000). Probabilities and polarity biases in conditional inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 883–899.

RAND Mathematics Study Panel Report. (2002). *Mathematical proficiency for all students: A strategic research and development program in mathematics education*. Washington, DC: U.S. Department of Education.

Stanovich, K. E. (2004). *The robot's rebellion: Finding meaning in the age of Darwin*. Chicago, IL: University of Chicago Press.

Sutherland, R., Olivero, F. & Weeden, M. (2004) Orchestrating mathematical proof through the use of digital tools. In M. Johnsen Høines & A. Berit Fuglestad (Eds.), *Proceedings of the 28th Conference of the International Group for the Psychology of Mathematics Education*, (Vol. 4 pp. 265–272). Bergen, Norway.

Yackel, E., & Hanna, G. (2003). Reasoning and proof. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A research companion to the Principles and Standards for School Mathematics* (pp. 227–236). Reston, VA: NCTM.