UNDERSTANDING STUDENTS' SIMILARITY AND TYPICALITY JUDGMENTS IN AND OUT OF MATHEMATICS

Caroline Williams (ccwilliams3@wisc.edu)
Olubukola Akinsiku (akinsiku@wisc.edu)
Candace Walkington (cwalkington@wisc.edu)
Jennifer L. Cooper (jcooper4@wisc.edu) Amy Ellis (aellis1@wisc.edu)
Charles Kalish (cwkalish@wisc.edu) Eric Knuth (knuth@wisc.edu)
Departments of Curriculum & Instruction and Educational Psychology
University of Wisconsin - Madison

Research in the field of mathematics indicates that many students struggle with justification and proof. However, in non-mathematical contexts, students are relatively strong at inferential reasoning. Our research presents two parallel lines of investigation—one focused on mathematical domains, the other focused on non-mathematical domains—in order to examine and compare the ways in which students reason in each. We report on three interrelated studies:(i) a small-scale interview study in which students were asked to sort numbers, shapes, and birds according to their characteristics;(ii) a small-scale interview study in which students were asked to determine whether various properties were true or false; and (iii) a large-scale survey to further elaborate the results from the first studies.

Justification and proof serve important roles in the field of mathematics, and consequently, there is increasing concern among mathematics educators regarding the difficulties students demonstrate in learning to reason mathematically (Healy & Hoyles, 2000; Kloosterman & Lester, 2004; Knuth et al., 2002). In non-mathematical domains (e.g., biology), however, the research is strikingly different: cognitive science research highlights considerable strength in students' causal reasoning (e.g., Gelman & Kalish, 2006; Gopnik et al., 2004). These seemingly conflicting findings serve as the driving force behind our research: why is it that students struggle to reason mathematically and yet are simultaneously skilled at reasoning outside of mathematics? In order to begin answering that question, we adopted a methodology often used in cognitive science (e.g. Kalish & Lawson, 2007) to examine students' reasoning in the mathematical domains of number and geometry. The main objective was to examine different types of inductive strategies students use in mathematical domains. Outside the mathematics classroom, children typically develop facts and ideas via empirical generalizations and causal theories (Chater & Oaksford, 2008), and students may therefore rely on and make connections to their non-mathematical ways of reasoning as they encounter ideas and problems in mathematics. This paper reports the results of the first phase of a multi-year study designed to explore the connections between students' reasoning in mathematical and non-mathematical domains.

Leveraging Inductive Reasoning

Researchers in mathematics education acknowledge the difficulties students have with formal proof (Dreyfus, 1999; Healy & Hoyles, 2000; Knuth, Choppin, & Bieda, 2009), and generally consider the inductive reasoning skills that students typically employ to interfere with their development of deductive means of reasoning and justifying. Researchers in cognitive science have found that students have difficulties *in general* with formal inference and deductive arguments, but they—unlike many mathematics educators—have continued to pursue a deeper examination of inductive inference. We agree with the latter that inductive reasoning is a powerful and useful tool, and are consequently beginning to extend cognitive science research on strategies of inductive inference to the mathematical content domains of number and geometry.

Cognitive science researchers have examined the inductive judgments people make and found that peoples' beliefs are frequently based upon category membership status, similarity, and typicality (see Feeney & Heit, 2007). For example, after being told a novel property that a robin possesses, people are more likely to extend that property to other robins than to all other birds in general. People are also more likely to extend that novel property to a blue jay than to a mouse, based upon perceptions of greater similarity between robins and blue jays than between robins and mice. In addition, people believe a robin to be a more typical bird than a penguin, and are consequently more likely to generalize from robins to all birds than from penguins to all birds (Osherson et al., 1990). In mathematics we lack a similar knowledge of students' understanding of the categories, levels of similarity between objects, and typicality of objects that may inform students' inferential reasoning. Therefore, the purpose of this course of study is to understand and map students' similarity and typicality judgments in the mathematical content domains of number and geometry.

This paper follows a three-part study design in which the results of each study were used to design and implement the following study, chaining together in such a way that the third study relied upon as well as further elaborated the results of both the first and second studies. Given this dependent relationship, as well as the quite distinct differences in methodologies and consequent analyses for each study, we outline each study and attendant results individually, focusing particularly on the ways that the studies build upon each other.

Study 1

In order to design a study that would elicit the typicality ratings and similarity relations for numbers and shapes (triangles and parallelograms), we needed to identify what features of each domain were considered relevant. When looking at triangles, for example, what do middle school students, undergraduates, and mathematics and engineering (hereafter STEM) graduate students see as important characteristics? Consequently, we designed a structured but open-ended methodology that prompted interviewees to utilize and share their own feature classifications.

Methods

We adopted the sort-re-sort procedure used by Medin et al. (1997). In one-on-one interviews, participants were presented with a set of cards that contained numbers or geometric shapes, the selections of which were based upon the varying of dimensions identified by previous research (Feldman, 2000; Miller & Gelman, 1983; see Figure 1).

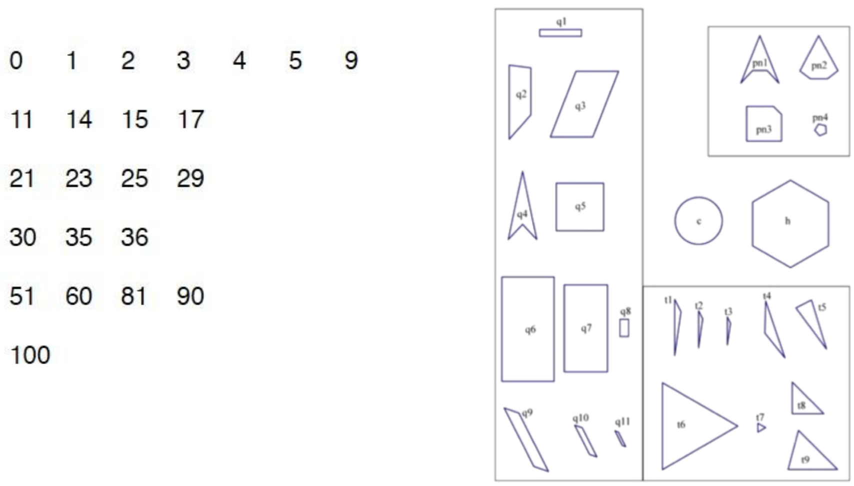


Figure 1: Number and shape cards.

Participants were asked to form groups with the cards, and then encouraged to further separate cards within those groups. Once sorting appeared to be complete, the cards were reshuffled and the participants were asked to form different groups. The sorting-re-sorting was continued until the participants had exhausted their different ways of grouping. The participants included 14 middle-school students, 14 undergraduate students (from a wide variety of majors with calculus the highest level of college mathematics completed), and 14 STEM graduate students).

Results

The results were striking because the different populations (middle-school students, undergraduates, and STEM graduate students) were not particularly different in their sorting and their sorting rationale. The following tables (see Tables 1 and 2) contain a list of several features used in the sorting task, along with the percentage of use by each population. In presenting the results, only those features noticed by at least 50% or more of the middle school students are included as that population is the primary focus of our research.

Table 1: Number Features.

Features	Middle Schoolers	Undergraduates	STEM Grads
Multiples	86%	93%	93%
Parity	79%	93%	87%
Prime	50%	71%	80%
Value of Digit	50%	29%	47%

In the domain of number, participants in all three groups were likely to remark upon features such as number parity, whether or not a number was prime, and the multiples of numbers. Comparing across populations (and not included in the table), the STEM graduate students were most likely to notice square numbers (53%), undergraduate students were most likely to remark upon shared digits between numbers (43%), and middle school students were most likely to mention factors of numbers (21%). Interestingly, in general undergraduate students did not appear, as might be expected, to be very different in their sorting than the middle school students—perhaps an indication that the salience of particular features seems to have a complex relationship with mathematics expertise.

The same trend appeared in the domain of geometry, such that there was not a clearly defined trajectory from middle school novices to STEM experts. Table 2 contains a list of the geometric features used to sort by at least 50% of the middle-schoolers, along with what percentage of each population sorted by that feature.

Table 2: Shape Features.

Features	Middle Schoolers	Undergraduates	STEM Grads
# of Sides	100%	100%	100%
Size	78%	85%	100%
Shape	64%	78%	93%
Angles	64%	85%	80%
Similar	50%	43%	40%

Once again, there were some particularly interesting similarities between the three groups, particularly that all of the participants remarked upon the number of sides. Middle school students, undergraduates, and STEM graduate students were all relatively likely to remark upon the size or shape of a geometric object. In addition, these two features are the only two that showed a progression from middle school students to undergraduates to STEM graduates. Other features was more erratically patterned. For example, 20% of the middle school students and about 50% of the STEM graduate students remarked upon the regularity of shapes, but none of the undergraduates sorted by regularity. In contrast, the similarity of shapes (i.e., the size and shape of geometric objects) was noted the more by middle school students than either undergraduates or STEM graduates, with the latter being least likely to use the feature to sort. (For more detail on Study 1, see Knuth et al., In press.)

Study 2

The results from Study 1 provided insight into the features of mathematical objects that are salient to middle school students, undergraduates, and STEM graduate students. Study 1 reveals something of the structure of students' representations of the domains of number and shape. Study 2 went on to explore whether this structure influences reasoning and proof strategies.

Methods

We designed a two-part semi-structured interview protocol in which we presented participants with both specific and general propositions and then asked them to determine whether the propositions were always true (see Figure 2).

> Eric came up with a new mathematical property. He thinks this property is true for every whole number.
>
> Eric's property:
> First, pick any whole number.
> Second, add this number to the number before it and the number after it.
> Your answer will always equal 3 times the number you started with.

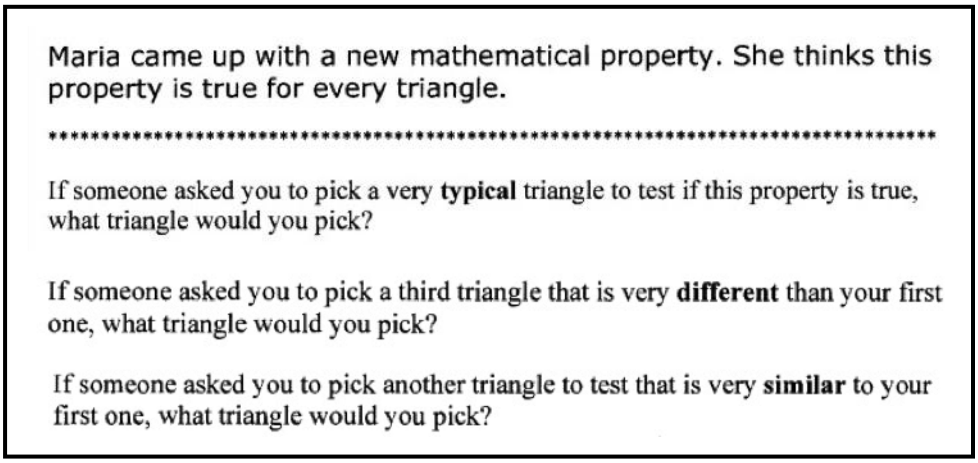
> Lewis came up with a new mathematical property. He thinks this property is true for every triangle.
>
> Lewis's property:
> For any triangle, two of the sides added together are longer than the third side.

Figure 2: Specific propositions.

If students used examples in attempting to determine the truth of a proposition, they were asked about their example choices, including their judgment of the similarity and typicality of each example. For the number proposition above, for example, one participant used the numbers 4, 6, 8, and 5 to check the proposition's validity. When the participant had determined to her satisfaction that the proposition was true, she was then asked whether any of the numbers were unusual or typical. The final series of questions for each proposition asked the participant to share their beliefs about the similarity of the numbers and shapes they tested:

> *Interviewer:* "So are all these numbers similar to each other or different than each other?"
>
> *Participant:* "Well, I guess I think of 4, 6, and 8 as similar. Because they're all even. And I think of 5 as different because it's an odd number."

Next, the participants were presented with general propositions (see Figure 3, as an example) and asked to supply examples that varied in typicality and similarity. In particular, participants were asked to generate three different examples to test: a typical example, an example similar to that typical example, and an example different from that typical example.


Maria came up with a new mathematical property. She thinks this property is true for every triangle.

****

If someone asked you to pick a very **typical** triangle to test if this property is true, what triangle would you pick?

If someone asked you to pick a third triangle that is very **different** than your first one, what triangle would you pick?

If someone asked you to pick another triangle to test that is very **similar** to your first one, what triangle would you pick?

Figure 3: General proposition.

The participants included 14 undergraduate students (not discussed in this paper) and 20 middle-school students. Further details about the methodology and results from the middle school students' responses to specific propositions are presented in Cooper et al. (2011).

Results and Discussion

Results from Part 2 suggest that the vast majority of middle-school students can use example-based reasoning as a method of proof, that more examples tended to be more convincing than few examples, and that the greater the variety of examples, the more convincing. Variety, here, seems to validate the numbers and shapes we used in Study 1, as students talked about varying numbers and shapes along the very dimensions we tried to systematically represent in the sorting task. Particularly, students frequently reported using typicality and similarity judgments to inform their selection of items as they tested examples. Importantly, students who discovered deductive proofs were less likely to vary examples overall, but on the problems where they did vary examples, they reported using variations in typicality and similarity in an intentional manner. While students who did not discover deductive proofs tested the greatest variety of examples, their variation was not intentional.

Students frequently referenced their everyday experiences—both inside and outside the mathematics classroom—during the interviews. For example, a respondent who talks about squares being typical because all the corners have 90 degree angles is quite different from the respondent who declared a tall rectangle to be typical because it was a common skyscraper shape. This result prompted us to reconsider the idea that typicality is a singular construct, and instead reframe typicality as context-relevant. In addition, as was also found in Study 1 with the non-linear trajectories of participants attending to particular mathematical features, in Study 2 middle school students' typicality judgments for number and geometry categories did not always correspond to expert-based dimensions. In particular, we became interested in how the student-based notions of typicality contrast with expert-based knowledge relating to special mathematical properties.

For instance, in a mathematical contexts, for the purposes of justification, experts may consider mathematically-special numbers like 0 and 1 to be highly unusual or atypical – mathematical properties that hold for these numbers may not be generalizable, and these numbers may be useful to employ in the search for counterexamples. However, in terms of numbers seen in everyday life, these numbers are very typical or common. The analysis of examples given by students and their associated typicality comments led us to tentatively map numbers and shapes by their everyday usual-ness (student typicality) and by their mathematical generic-ness (mathematical typicality, where non-generic indicates a high presence of mathematical properties, such as prime-ness in numbers or congruence in shapes). Figure 4 illustrates our placement of numbers according to these two dimensions.

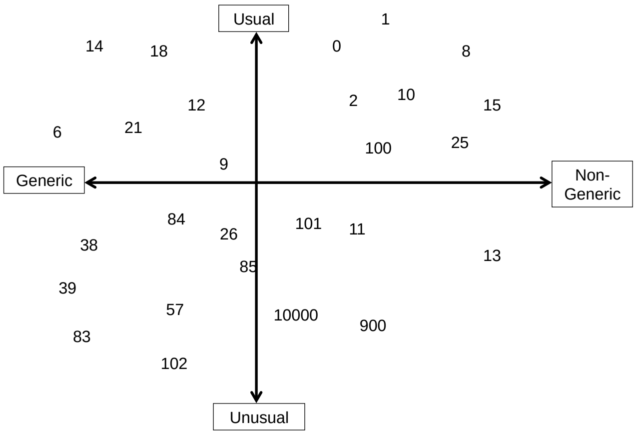


Figure 4: Everyday versus generic number placement.

Study 3

Given the intriguing results from Studies 1 and 2, our next step was to administer a survey to a large enough number of middle school students to validate our results. Given the need to consider the importance of context, we determined that cuing the students to the context of "math class" versus "everyday life" was a necessary nuance. That is, we wanted to find out if students consider the expert-based generic-ness dimension of number and shape typicality if sufficiently prompted to consider these objects in a mathematical context.

Methods

We designed a 7-point Likert-scale survey where students rated the typicality of items in three different mathematical domains (numbers, triangles, parallelograms) and three different contextual cues (neutral, mathematical, everyday) (see Table 3). Through careful variance of features, items covered a wide gamut of types—in triangle, for example, we included isosceles, equilateral, and scalene, and varied both the size and the orientation of the shape (some shapes had a side parallel to the bottom of the survey page, while others did not). Additionally, we always placed the neutral cues first, but otherwise varied the order of the cues and domains.

Table 3: Stimuli prompts in the number domain.

	Neutral Context	Mathematical Context	Everyday Context
Section Prompt	Now we want you to think about how typical different numbers are.	Now we want you to think about mathematical properties – that is, the kinds of things you learn about in math class.	Now we want you to think about numbers you see in everyday life outside of school, for example, around your house, in stores, or outside.
Individual Item Wording	Think of numbers. How typical is this number? [item displayed]	Imagine that we learned a new mathematical property that was true of this number. How likely is it that the property will be true of most other numbers? [item displayed]	How typical is this number of those you see in your everyday life? [item displayed]

A total of 474 middle school students, drawn from a suburban middle school in a Midwestern state, completed the survey. Students were distributed across grades 6 (144 students), 7 (160 students), and 8 (163 students), and mathematics classes used reform texts.

Results

Using the findings of Study 1 and Study 2, we identified a number of mathematical and non-mathematical properties that we hypothesized students may attend to in their typicality rankings (see Figure 5). We computed the mean typicality rating for items falling into each of these categories, depending on whether the item was presented in a mathematical or everyday context. The results suggest that while both mathematical (generic-ness) and non-mathematical (common-ness) properties can influence typicality ratings, the everyday notion of typicality is most salient to middle school students. The results also indicate that contextual cues do not significantly influence students' typicality judgments, and that even cueing the context "mathematics classroom" did not increase the likelihood that students would attend to particular mathematical properties.

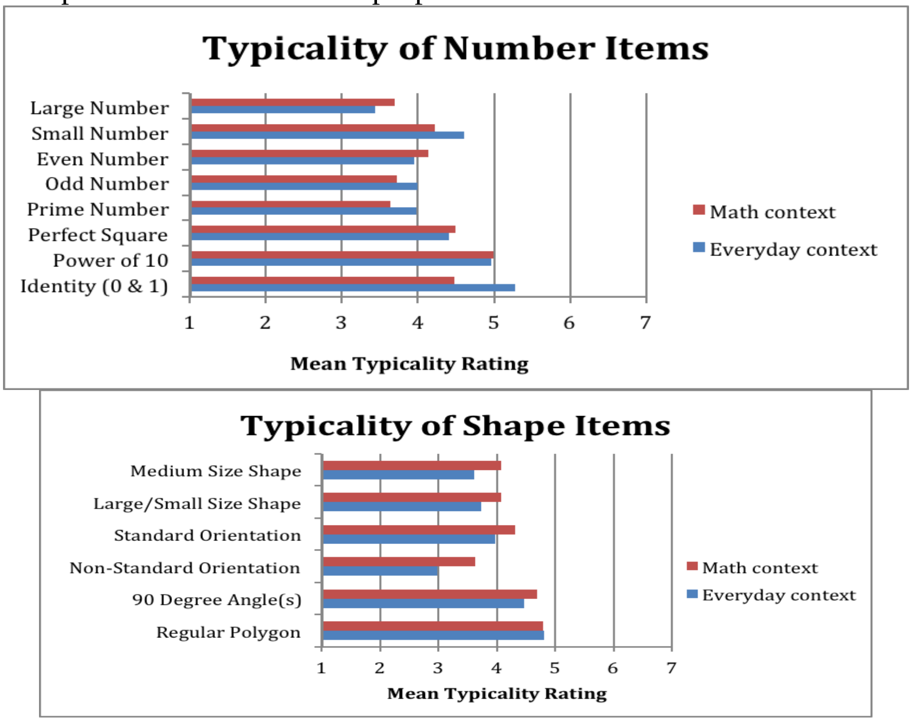


Figure 5: Mean typicality ratings for number and shape items.

Conclusion

Study 1 resulted in an identification of various sorting features of mathematical objects noticed by middle school students, undergraduates, and STEM graduate students, and furthermore, suggested that patterns across the three participant groups are not simple. In the domains of both number and geometry, an increase of mathematical expertise does not necessarily manifest itself in a higher likelihood of noticing particular features, and the three groups varied in the features to which they attended. The results of Study 2 suggest that typicality and similarity judgments may play a role in inductive and deductive reasoning, and that participants believed that varying the typicality and similarity of examples is a good strategy for testing propositions in both number and shape. In Study 3, the survey validated our initial smaller-scale interview studies, providing more confidence in our student typicality (usual-ness) identification of specific numbers and shapes. In the next phase of our work, we are further investigating the relationship between similarity/typicality and the nature of the empirical arguments student generate.

We believe that inferential reasoning can be leveraged to support the development of deductive reasoning, as it seems that applying strategies frequently used for the domain of living things—that is, thinking of examples, typicality, and similarity—can support the development of more sophisticated ways of reasoning in mathematics. Our research highlights the need to more closely and systematically examine inferential reasoning in mathematics, and instead of viewing students' use of examples as inappropriate or limiting, we view such use as a potential means of improving students' abilities to reason deductively. The importance of proof in mathematics, and the struggle of most students to justify and reason deductively, suggests that this examination is of paramount importance, and our research opens a new—and hopeful—line of study.

References

Chater, N., & Oaksford, M. (2008). *The probabilistic mind: Prospects for Bayesian cognitive science.* Oxford: Oxford University Press.

Cooper, J. L., Walkington, C. A., Williams, C. C., Akinsiku, O. A., Kalish, C. W., Ellis, A. B., & Knuth, E. J. (2011). Adolescent reasoning in mathematics: Exploring middle school students' strategic approaches in empirical approaches. *In Proceedings of the 33rd Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

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

Feeney, A., & Heit, E. (Eds.) (2007). *Inductive reasoning*. Cambridge: Cambridge University Press.

Feldman, J. (2000). Minimization of Boolean complexity in human concept learning. *Nature, 407,* 630-633.

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.

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

Kalish, C.W., & Lawson, C. A. (2007). Negative evidence and inductive generalization. *Thinking and Reasoning, 13* 394-425.

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., Choppin, J., & Bieda, K. (2009). Middle school students' production of mathematical justifications. In D. Stylianou, M. Blanton, & E. Knuth (Eds.), *Teaching and learning proof across the grades: A K-16 perspective* (pp. 153-170). New York, NY: Routledge.

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.

Knuth, E., Kalish, C., Ellis, A., Williams, C., & Felton, M. (In press). Adolescent reasoning in mathematical and non-mathematical domains: Exploring the paradox. To appear in V. Reyna, S. Chapman, M. Dougherty, & J. Confrey (Eds.), *The adolescent brain: Learning, reasoning, and decision making*. Washington, DC: American Psychological Association.

Medin, D.L., Lynch, E.B., Coley, J.D., & Atran, S. (1997). Categorization and reasoning among tree experts: Do all roads lead to Rome? *Cognitive Psychology, 12,* 496-522.

Miller, K. F., & Gelman, R. (1983). The child's representation of number: A multidimensional scaling analysis. *Child Development, 54,*1470-1479.

Osherson, D.N., Smith, E.E., Wilkie, O., Lopez, A., & Shafir, E. (1990). Category-based induction. *Psychological Review, 97,* 185-200.