

Empirical Articles

Middle School Students' and Mathematicians' Judgments of Mathematical TypicalityCandace Walkington*^a, Jennifer Cooper^b, Olubukola Leonard^c, Caroline Williams-Pierce^d,
Chuck Kalish^e

[a] Department of Teaching and Learning, Southern Methodist University, Dallas, TX, USA. [b] Department of Psychology, Stonehill College, Easton, MA, USA. [c] Department of Curriculum and Instruction, University of Wisconsin – Madison, Madison, WI, USA. [d] Department of Educational Theory & Practice, University of Albany, Albany, NY, USA. [e] Department of Psychology, University of Wisconsin-Madison, Madison, WI, USA.

Abstract

K-12 students often rely on testing examples to explore and determine the truth of mathematical conjectures. However, little is known about how K-12 students choose examples and what elements are important when considering example choice. In other domains, experts give explicit consideration to the typicality of examples – how representative a given item is of a general class. In a pilot study, we interviewed 20 middle school students who classified examples as typical or unusual and justified their classification. We then gave middle school students and mathematicians a survey where they rated the typicality of mathematical objects in two contexts – an everyday context (commonness in everyday life) and a mathematical context (how likely conjectures that hold for the object are to hold for other objects). Mathematicians had distinct notions of everyday and mathematical typicality – they recognized that the objects often seen in everyday life can have mathematical properties that can limit inductive generalization. Middle school students largely did not differentiate between everyday and mathematical typicality – they did not view special mathematical properties as limiting generalization, and rated items similarly regardless of context. These results suggest directions for learning mathematical argumentation and represent an important step towards understanding the nature of typicality in math.

Keywords: typicality, proof, inductive generalization

Journal of Numerical Cognition, 2018, Vol. 4(1), 243–270, doi:10.5964/jnc.v4i1.70

Received: 2016-07-11. Accepted: 2017-06-03. Published (VoR): 2018-06-07.

Handling Editors: Anderson Norton, Department of Mathematics, Virginia Tech, Blacksburg, VA, USA; Julie Nurnberger-Haag, School of Teaching, Learning, and Curriculum Studies, Kent State University, Kent, OH, USA

*Corresponding author at: 6401 Airline Rd., Dallas, TX, 75205, USA. Tel.: 214-768-3072. E-mail: cwalkington@smu.edu



This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International License, CC BY 4.0 (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Educated adults are more likely to judge that an equilateral triangle is a triangle than that a scalene triangle is. Such adults also report that 400 is a “better” even number than is 798 (Armstrong, Gleitman, & Gleitman, 1983; Feldman, 2000; Lupyan, 2013). These judgments are often taken as evidence that people’s knowledge of mathematical objects involves more than strictly formal definitions (Lakoff, 1990). A large literature in cognitive science demonstrates that expertise involves non-formal reasoning structures (Bédard & Chi, 1992; Kahneman & Klein, 2009; Reyna, 2012), including judgments about the important characteristics of various objects in a domain. In mathematics, such judgments may be useful when considering which examples to test to explore mathematical conjectures or ideas. However, research suggests that K-12 students may only have partially

developed expertise in thinking critically about examples. The current paper explores middle school students' and mathematicians' judgments of important characteristics of mathematical objects. Our overarching goal is to consider how to support strategic reasoning with examples to promote activities related to mathematical justification.

The general explanation for why equilateral triangles are more often judged to be triangles than scalenes, and why 400 is said to be a better even number than 798, is that the former member of each pair is understood to be more *typical* than the latter. Typicality refers to how representative a given item is of a general class. Typical instances are highly similar to many other members of their class; atypical instances are dissimilar from most others (see [Murphy, 2002](#) for review). Typicality has important consequences for reasoning - people are more likely to generalize from typical instances than atypical instances. All else equal, learning that a dog has a particular property makes people more confident that most other mammals also have the property than does learning that a whale has it ([Osherson, Smith, Wilkie, Lopez, & Shafir, 1990](#)). Typicality can lead people to treat information as more or less valuable.

One of the most interesting features of mathematical typicality is not that people judge some numbers or shapes to be better instances of their classes than others, but how they do so. Mathematically, an equilateral triangle is atypical. Equilateral triangles have many distinctive properties, and facts about equilateral triangles may not generalize. For example, Angle-Side-Side congruence holds for equilaterals, but not for all triangles. Similarly, given our base 10 system of counting, 400 is a relatively special quantity. Notably, many calculations are much easier with 400 than with 497 (e.g., squaring). However, equilateral triangles and numbers that are multiples of 100 are likely more familiar and frequently encountered in daily life ([Lakoff, 1990](#)). One possibility, then, is that there may be different, even competing, senses of typicality in mathematics. Everyday typicality, based on familiarity may conflict with mathematical typicality, based on generalizability of mathematical properties.

We examine middle school students' and mathematicians' judgments of typicality in two contexts - how often numbers and shapes are seen in everyday life, and how generalizable numbers and shapes are in terms of their mathematical properties. These judgments have important implications for teaching K-12 students about the strategic use of examples in mathematical arguments, which aligns with current standards in mathematics education.

Literature Review

We begin our review in the area in mathematics education that we have been exploring issues of typicality – mathematical justification, generalization, and proof. We then discuss typicality relations in general, before moving to discuss their significance in mathematics.

Mathematical Justification, Generalization, and Proof

Justification and proof are important activities in mathematics education ([National Council of Teachers of Mathematics, 2000](#); [Stylianides, 2007](#); [Yackel & Hanna, 2003](#)). We follow [Harel and Sowder \(1998\)](#) in defining proving as “the process employed by an individual to remove or create doubts about the truth of an observation” (p. 241). This definition allows for a wider range of approaches and modalities to be considered than the traditional “two column” proof ([Herbst, 2002](#)). The development of proofs includes several integrated stages – identifying patterns, making conjectures, providing non-proof arguments, and providing proofs – and

these activities involve both making generalizations and providing support for mathematical claims (Stylianides, 2008). This perspective closely mirrors the *Common Core State Standards for Mathematical Practice* (Common Core State Standards Initiative, 2010) which state that K-12 students should construct mathematical arguments and critique the reasoning behind arguments. Specifically, K-12 students should “make conjectures and build a logical progression of statements to explore the truth of their conjectures” as well as “reason inductively about data, making plausible arguments.”

Research suggests K-12 students often struggle to learn how to construct viable and convincing mathematical arguments and provide valid generalizations (Dreyfus, 1999; Healy & Hoyles, 2000; Martin, McCrone, Bower, & Dindyal, 2005). However, research in cognitive science has revealed that children show a surprisingly strong ability to reason and construct arguments in domains like science (Gelman & Kalish, 2005; Gopnik et al., 2004), where inductive, example-based reasoning has long been considered appropriate. Here, inductive reasoning refers to using examples to provide support for understanding a conjecture, which is different from the more formal notion of “proof by mathematical induction,” as well as from traditional notions of formal, deductive proofs. In science, if one wishes to test a conjecture like “All birds have hollow bones,” there is no deductive way to prove this conjecture – reasoning from examples to create causal theories is often the most appropriate form of inference.

Deductive reasoning is, however, fundamental to human reasoning generally, and mathematical reasoning specifically (English, 1997). Piaget’s developmental model of cognitive development suggest that it is not until the formal operational stage (11-12 years and older) that students are capable of reasoning hypothetically in service of logical deduction (Piaget, 1928). Research since has challenged the notion that preadolescent children do not have access to deductive reasoning (e.g., Mody & Carey, 2016) although their understanding may be fragile (English, 1997). During middle childhood, children can distinguish between deductive inferences, inductive inferences, and uninformed guesses (Pillow, 2002). When students do reach adolescence, the production of deductive reasoning can depend on the perceived relevance or familiarity of the content structure (Ward & Overton, 1990).

Example Use During Justification and Proof

Examples are culturally-mediating tools that connect learners and mathematical ideas – they provide a specific and familiar context for exploration and understanding of nuances and constraints (Goldenberg & Mason, 2008). Learners consider how features of an example can be modified while still maintaining its membership to the class of mathematical objects it is intended to exemplify (Watson & Mason, 2005). Sometimes, learners’ notions of what can be changed about an example are unnecessarily restricted based on inappropriate perceptual features, like believing that a long skinny triangle is a “stick” and not a triangle (Goldenberg & Mason, 2008). A learner’s repertoire of available examples, along with their methods for constructing those examples and inter-relations between examples, constitutes their example space (Watson & Mason, 2005). Example spaces have structural characteristics like the density of available examples, the generativity potential for new examples using existing examples, the connectedness of examples, and “the extent to which a given example is specific or whether it is representative of a class of related examples” (Sinclair, Watson, Zazkis, & Mason, 2011, p. 302).

Both K-12 students (Cooper et al., 2011; Knuth et al., 2011) and mathematicians (Lockwood, Ellis, Dogan, Williams, & Knuth, 2012) use examples to explore mathematical conjectures. Mathematicians actively consider

examples and “it is probably the case that most significant advances in mathematics have arisen from experimentation with examples” (Epstein & Levy, 1995, p. 6). There is often a back-and-forth interplay between mathematicians’ example-related reasoning and their deductive proof activities (Alcock & Inglis, 2008; Lockwood, Ellis, & Knuth, 2013). For examples to be most useful, the prover may need to intentionally choose examples with a specific purpose in mind, while confronting a task that lends itself well to their use (Sandefur, Mason, Stylianides, & Watson, 2013).

The role examples play in the justification activities of K-12 students is quite different. K-12 students can be overly reliant on examples and mistakenly use examples alone as a form of proof (Harel & Sowder, 1998; Healy & Hoyles, 2000; Knuth et al., 2009; Koedinger, 1998; Porteous, 1990). There is also evidence that K-12 students may not analyze examples strategically in order to make sense of a mathematical statement and to develop insight into its proof (Cooper et al., 2011; Knuth, Kalish, Ellis, Williams, & Felton, 2011). The contrast between the role examples play in the work of mathematicians and K-12 students is not surprising given that these students typically receive very little instruction on how to analyze examples when developing, exploring, understanding, and proving conjectures. Although K-12 students may see confirmatory examples as sufficient for proof, questioning whether the confirmatory result was an accident because of a “special case,” or whether the example is “generic” and appropriately represents a general case, may trigger a need for proof (Buchbinder & Zaslavsky, 2011).

Typicality and Expertise

Research in other domains suggests that experts have a set of heuristics for judging which relations are important in their domain of expertise (Bédard & Chi, 1992). Experts effectively employ formulas and principles because they are able to identify which features of a problem are important ones to focus on and how to generalize from one problem to another. Novices also generalize, but they sometimes focus on irrelevant or non-productive features (Chi, Feltovich, & Glaser, 1981; Chi & VanLehn, 2012). Novices are often described as focusing on “surface” features, those that are immediately apparent but not necessarily important in the domain, while experts focus on “deep” features, which are more productive for reasoning about properties in a domain. Part of being an expert is having an educated sense of typicality, a sense that reflects real relations in the domain (Feeney & Heit, 2007). With more knowledge about the domain, ideas about relevant features and relationships may shift; a landscape architect, a horticulturalist, and a botanist focus on different features of trees (Medin, Lynch, Coley, & Atran, 1997).

While deductive inference appears to be a late-emerging skill, requiring explicit practice and instruction, inductive inference seems to be a cognitive primitive (Kuhn, 1989). In a very basic sense, any form of conditioning is a kind of inductive inference (e.g., learning to expect a more frequent event over a less frequent event). Preschool-aged children, and even infants, are sensitive to some principles of inductive inference (e.g., generalizations are more likely to be true for more homogeneous classes; Gelman, 1988). Research on scientific reasoning and critical thinking (see Kuhn, 1989) suggests that explicit, deliberate inductive argumentation or proof has an extended developmental and instructional timeline.

By 2 years of age, children understand that objects can belong to the same category, even if they are perceptually dissimilar, and use this to reason successfully about properties of those objects (Gelman & Coley, 1990). Kindergarteners can reason that a property is more likely to hold for one member of a class given that it holds for another member when those two members are similar (López, Gelman, Gutheil, & Smith, 1992). By

second grade, children can recognize typicality – how representative the member is of its class, and how that should impact judgments (López, Gelman, Gutheil, & Smith, 1992). Later, children gain the ability to also consider the diversity of examples – the idea that sets of examples that better cover the space of the domain are stronger for inferences. While 6-year-olds prefer sets of typical exemplars as evidence, 9-year-olds favor diverse samples, except when they are atypical, while adults consistently favor diverse samples (Rhodes, Brickman, & Gelman, 2008). The most sophisticated forms of inductive argumentation, a normative theory of induction, is formal statistics (Hacking, 2001).

Typicality and Mathematical Expertise

Although typicality relations are clearly important in domains like biology, little is known about what features are relevant within mathematics and whether mathematicians and K-12 students demonstrate different knowledge about mathematical features and relationships. For a mathematician, zero may be an extremely atypical (and important) number in the mathematical context of testing examples. However, people also have ideas about typicality relations in everyday life. Zero may be a fairly typical number as it is commonly encountered in stores (e.g., \$0 down payment), on devices (0% downloaded), etc.

The features of objects that matter for judgments of typicality in mathematical contexts are an open question, but guesses are available. For geometric shapes in particular, there is a clear hierarchy of important mathematical properties that may matter when testing conjectures, such as having congruent sides and angles, right angles, etc. The mathematical properties of numbers that may matter for typicality are less clear and may be dependent on characteristics of particular conjectures – although numbers like zero and one clearly have special properties. We next discuss how mathematicians consider examples when exploring conjectures.

How Do Mathematicians Think About Typicality?

Examining how mathematicians think about the examples they test when exploring conjectures offers insight into important elements of strategic thinking about examples. In an online survey distributed to 220 mathematicians at 27 universities around the country (Lockwood, Ellis, & Lynch, 2016), mathematicians were asked to respond to two open-ended questions: “If you sometimes use examples when exploring a new mathematical conjecture, how do you choose the specific examples you select in order to test or explore the conjecture? What explicit strategies or example characteristics, if any, do you use or consider?” An analysis of the responses to these items is reported elsewhere (Lockwood et al., 2016), so here we summarize the results.

Mathematicians would often cite that they used simple examples (e.g., “Easy ones! Start with toy cases and slowly build up the complexity.”) and complex examples (e.g., “I try to find examples that include all of the (foreseen) barriers to a proof. The “hardest” examples in the sense of what I’m trying to prove.”), as well as examples that made computation easy. Mathematicians also reported choosing counter-examples, examples with specific properties (e.g., “For number-based conjectures, I choose 0, numbers close to 0 - both positive and negative - very large and very small numbers.”), and examples that represent a boundary case (e.g., “I will next try to test some strange or pathological examples, to really push the boundaries of what might be possible in this situation.”). Some mathematicians also reported using generic examples (e.g., “Try the most general example which is still practical to test”), examples that are familiar, unusual examples (e.g., “Next, I try something slightly more obscure.”) and common examples (e.g., “Ones that are not special, ones that I judge to be typical.”), as well as random, exhaustive, and dissimilar examples.

It also is important to consider mathematicians' use of examples when exploring specific conjectures, particularly conjectures that would not be trivial to them with their expertise. In a follow-up study (Lockwood et al., 2016), six mathematicians were presented with four conjectures (three of which were Putman exam conjectures) during one-on-one interviews. Results reinforced and extended the findings from the survey – mathematicians were purposeful in their selection of examples and used domain knowledge about the context to choose and strategically use examples, with their knowledge of special mathematical properties being especially important. Lockwood et al. (2016) describe how “In the interviews, the mathematicians' domain-specific expertise was apparent as they spoke about the mathematical features of their examples, such as choosing a number with many factors or creating a set with no primes.” (p. 177). Further, mathematicians looked for patterns through multiple examples and used examples to gain insight into whether a conjecture was true and make progress towards a proof. Overall, the mathematicians' responses supported the idea that there may be a network of mathematical relations that they are considering when choosing examples.

How Do Middle School Students Think About Typicality?

Middle school students may have less well-developed notions about choosing examples. In a recent study, middle school students and STEM doctoral students were engaged in card-sorting tasks where they constructed sets of numbers, triangles, and parallelograms (Williams et al., 2011). Both groups tended to sort numbers by whether they were even/odd or prime, with middle school students focusing more on multiplicative factors of the numbers, and STEM doctoral students focusing more on square numbers. Middle school students were also more attuned to surface-level features of numbers, like how many digits a number had and whether it contained a certain digit. For shapes, both STEM doctoral students and middle school students were most attuned to number of sides. STEM doctoral students also considered whether shapes were regular and whether they were symmetrical; middle school students were more likely to consider whether shapes were mathematically similar (i.e., same angles but a different size). Both middle school students and STEM doctoral students were also attuned to features of shapes like size, orientation, and familiarity.

Although this study provided useful ideas about properties of mathematical objects that are salient to different groups, we wanted to more specifically explore ideas about *typicality*. We engaged in a series of interviews with middle school students where we probed their ideas about what made mathematical objects “typical.” This pilot interview study then led to a large-scale survey with middle school students and mathematicians about conceptions of typicality.

Pilot Study

For our pilot interview study, our research question was: What reasons do middle school students give for classifying numbers and shapes as typical or unusual?

Method

One-on-one interviews with 20 middle school students (11 female, 9 male; 8 sixth-grade, 7 seventh-grade, and 5 eighth grade or higher math courses) from a large suburban school district in a Midwestern state were conducted. Parental consent forms were distributed to students in 6 middle schools and students who returned consent forms were selected on a first-come-first-serve basis. The district used a reform-oriented curriculum,

Connected Mathematics 2 (Lappan, Fey, Fitzgerald, Friel, & Phillips, 2006). Students were engaged in a 1-hour video-recorded interview where they were presented with the tasks in Table 1. There were 2 interviewers who were math education graduate students, and they followed a semi-structured protocol for interactions with participants (Kvale, 1996). The first four tasks were specific conjectures the participant would attempt to prove or disprove (Eric, Amy, Lewis, & Bob in Table 1). An analysis of how participants proved these conjectures (e.g., empirical versus formal algebraic approaches) is reported elsewhere (Cooper et al., 2011); testing examples was a primary strategy.

Table 1

Tasks Used in Middle School Student Interview Pilot Study

Specific Conjectures
1. Eric came up with a new mathematical property. He thinks this property is true for every whole number. 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 out with.
2. Amy came up with a new mathematical property. She thinks this property is true for every even number. First, pick any even number. Second, add this number to half of itself. Your answer will always be divisible by 3.
3. Lewis came up with a new mathematical property. He thinks this property is true for every triangle. For any triangle, two of the sides added together are longer than the third side.
4. Bob came up with a new mathematical property. He thinks this property is true for every parallelogram. The angles inside any parallelogram add up to 360 degrees.
Generic Conjectures
5. Matt came up with a new mathematical property. He thinks this property is true for every whole number. If someone asked you to pick a very typical whole number to test if this property is true, what whole number would you pick?
6. Caro came up with a new mathematical property. She thinks this property is true for every whole number. If someone asked you to pick a very unusual whole number to test if this property is true, what whole number would you pick?

Note. There were 10 Generic conjectures, but only 2 are shown for illustrative purposes.

We focus on a series of questions the interviewer asked each participant *after* they had determined the truth of each conjecture. The interviewer pointed out each example participants had tested when exploring the conjecture and asked the student if each example was typical or unusual and why it was typical or unusual. Participants were also given 10 generic conjectures (such as Matt and Caro in Table 1) within 5 mathematical domains (whole numbers, even numbers, odd numbers, triangles, and parallelograms). Participants were either asked to pick either a “very typical” example to test the unspecified conjecture, or a “very unusual” example.

The examples that middle school students generated for specific and generic conjectures were coded based on (1) whether they deemed the example to be “typical” or “unusual” and (2) the reasons the middle school students gave for *why* their examples were “typical” or “unusual.” The reasons were coded using emergent categories derived from constant comparisons (Glaser & Strauss, 1967), with multiple codes possible for a single example. For the four specific conjectures, the 20 participants generated 278 total examples (181 typical, 72 unusual, and 25 not specified). For the ten generic conjectures, the 20 participants generated 182 total examples (they were instructed to generate one typical or unusual example per conjecture). For the generic conjectures, in some cases the interviewer ran out of time and not all participants received all tasks – this happened in 18 instances. For the specific conjectures, in some cases, the interviewer did not ask participants to provide a reason why their example was typical/unusual, either because of time considerations or because each generated example was not exhaustively cycled through – this happened in 35 instances. A second

trained coder coded a 10% subset of participants' evaluations of (1) whether numbers/shapes were typical/unusual and (2) the reasons given for why, and obtained Cohen's kappa reliability values of 1.0/1.0 (number/shape) and 0.81/0.73 (number/shape) respectively.

Results

Results showed that for the two specific conjectures related to numbers (Eric and Amy in Table 1), the most common reasons given for numbers being unusual (with category prevalence in parentheses) were: a number being uncommonly encountered in life (26% of instances), a number being prime (23%), and a number being odd (13%). The most common reasons given for numbers being typical were: a number being even (35% of instances), being commonly encountered in life (20%), numbers that were composite (17%), and numbers that were small (8%). For example, one student said that the numbers he tested were typical because "Well they're like even, and I seem to like even numbers more than I like odd 'cause I like- I like symmetrical things. That's just- that's just me."

For the two specific conjectures related to geometric objects (Lewis and Bob), the most common reasons given for shapes being unusual were: a shape having unequal sides (33% of instances), shapes that were elongated or in non-standard orientations (14%), shapes that were not commonly encountered in life (9%), and shapes with right angles (9%). The most common reasons given for shapes being typical were: a shape being commonly encountered in everyday life (22% of instances), a shape having equal sides (14%), a shape being labelled as a square or equilateral triangle (12%), and a shape being labelled as an acute triangle (8%). One student discussed the triangles he generated, saying "I guess I think of the typical ones are all equal. And I guess I just think of triangles that are typical that are all in- just- I don't know.... equilateral I guess. And then these that are all different lengths, I think they're just different. I don't think of them as typical." Another student said a parallelogram was unusual because "Probably because it's so like thin and long. Cause when I think of a parallelogram I think of like a house or a door."

Being encountered in everyday life was often mentioned as a reason for being typical, so the patterns for these justifications were examined more closely. Most often (37% of typical cases coded as "everyday life"), the reason would describe how frequently the object was encountered or used. Other common reasons within this category related to how the object was seen when working with money, packaging, school books or tasks, and when examining architecture or other real world structures/objects. Thus both the idea of frequent real world encounters as well as specific, salient real world connections seemed important for typicality.

The results for general conjectures were similar. The most common reasons for a number being typical were being even (45% of instances), often encountered in everyday life (44% of instances), Base-Ten properties (22%), divisibility properties other than 10s/2s (20%), or being small (11%) or an "origin" number (like 1; 11% of cases). The most common reasons for a number being atypical were not being encountered in everyday life (47% of instances), being odd (35%), being prime (22%), being large (16%), and being difficult computationally (16%). The most common reasons given for a shape being typical were being encountered in everyday life (62% of instances), being an equilateral triangle or square (21%), having equal sides (18%), or being symmetrical or not elongated (15%). The most common reasons for a shape being atypical were not being encountered in everyday life (32% of instances), having unequal sides (32%), and being elongated or in non-standard orientation (12%).

Discussion

This study suggested that middle school students were able to decide whether the examples they tested were “typical” or “unusual,” and give reasons for this classification based on mathematical and non-mathematical properties. There were characteristics of both number and shape that were consistently cited as either relating to typicality or atypicality, with frequency in everyday life being a major consideration for both. The ways in which these typicality relations interact with middle school students’ ideas about mathematical generalization is less clear, as are the ways in which their judgments relate to the judgments of mathematicians, who are more experienced with mathematical justification. We explore these questions next.

Survey Study

We frame our next study by discussing two aims that came out of prior research. Taken together, the card sort study (Williams et al., 2011), the mathematician questionnaire study (Lockwood et al., 2016), the mathematician interview study (Lockwood et al., 2016) and the pilot interview study generated interesting insights about the possible typicality relations in the domain of mathematical proof.

Our first aim was to explore which particular items mathematicians and middle school students consider typical or atypical, *without imposing any particular categorization* of these items or calling upon specific properties. The ways mathematicians and middle school students reason about objects may vary depending on whether the context is explicitly mathematical (e.g., conjecture-testing), or an everyday life context outside of canonical mathematical practices (e.g., walking through their neighborhood). We are interested in exploring typicality relations drawn upon in both mathematical and everyday contexts (which we refer to as “mathematical typicality” and “everyday typicality”). We hypothesize that mathematicians may differ from those less experienced with mathematics in that mathematicians have a sense of mathematical typicality that is distinct from everyday typicality.

Relating to our second aim, recall that in other domains, like biology, an awareness of typicality relations means understanding deep features, those that will impact your ability to generalize from examples, while disregarding surface features that do not impact important domain relations. In our prior studies, middle school students and mathematicians brought up a variety of mathematical properties and other characteristics that contributed to an object being deemed “typical” or “unusual” (e.g., being a multiple of 10, a small number, or an equilateral triangle). Our second aim was to compile a list of *specific properties* cited in these prior studies and examine how an object having each of these properties was associated with typicality ratings in mathematical and everyday contexts. This is useful because understanding how experts and novices construct typicality relations from mathematical properties has practical, instructional implications. In addition, a finding that distinctions between mathematical and everyday typicality correspond to an existing network of known mathematical properties would allow for a deeper understanding of how typicality structures operate in mathematics, compared to simply observing that there are differences in a series of isolated cases (as in RQ1). Table 2 describes the properties of numbers and shapes that were compiled, as well as which prior studies were drawn upon when considering whether to include that property. We hypothesize that middle school students and mathematicians will both attend to specific properties for judgments about typicality; however, whether a property makes an object typical or atypical may vary.

Table 2

Mathematical and Everyday Properties Considered for Entering Individual Properties Into Model (RQ2), With Justifications for Inclusion

Domain	Property	Justification for Inclusion
Number		
	Prime	Salient to middle school students in pilot interview study Salient to mathematicians in questionnaire/interview studies
	Perfect Square	Salient to STEM doctoral students in card sort study
	The number 0 or 1 (identity element)	Salient to middle school students in pilot interview study Salient to mathematicians in questionnaire/interview studies
	Multiple of 5/10 Power of 2/10	Salient to middle school students in pilot interview study
	Small/Large/ Extra-Large ^a	Salient to middle school students in pilot interview study, card sort study Salient to mathematicians in questionnaire/interview studies
	Odd Parity	Salient to middle school students in pilot interview study Salient for middle school students and STEM doctoral students in card sort study
Parallelogram		
	Square, rectangle, or rhombus	Salient to middle school students in pilot interview study Salient to STEM doctoral students in card sort study (regular, symmetric)
	Size (small, large)	Salient to middle school students and STEM doctoral students in card sort study
	Orientation (non-standard/left-leaning)	Salient to middle school students and STEM doctoral students in card sort study
	Skinny/elongated	Salient to middle school students in pilot interview study
Triangle		
	Isosceles, equilateral, right	Salient to middle school students in pilot interview study Salient to STEM doctoral students in card sort study (regular, symmetric)
	Scalene, Obtuse	Salient to middle school students in pilot interview study
	Size (small, large)	Salient to middle school students and STEM doctoral students in card sort study
	Skinny/elongated	Salient to middle school students in pilot interview study
	Orientation (non-standard)	Salient to middle school students and STEM doctoral students in card sort study

^aItems with this property were given to mathematicians only.

Research Questions

We investigate the following research questions (each related to one of our aims):

R1. *Is there a sense of “mathematical typicality” within mathematical contexts that is different from “everyday typicality” within everyday contexts, and what items are considered typical or unusual in these contexts?* We investigate whether mathematicians and middle school students distinguish

mathematical from everyday typicality by looking at how correlated their everyday typicality ratings were to their mathematical typicality ratings, as well as what specific objects they rate as more or less typical.

R2. *What specific characteristics or properties drawn from prior work make objects typical and atypical for mathematicians and middle school students in each context?* We investigate the overarching mathematical properties (congruency, divisibility, etc.) and everyday features (size, orientation) of mathematical objects that influence typicality judgments in mathematical and everyday contexts.

Method

Participants

Middle school student sample — A total of 475 middle school students (46% female) took the survey with pencil and paper. These students were from a school in the pilot study. The school was 48% Caucasian, 21% African American, 14% Asian, 11% Hispanic, and 1% Native American, with 37% free/reduced lunch and 10% English Language Learners. Participating students were distributed across grades 6 (144 students), 7 (160 students), and 8 (163 students).ⁱ

Mathematician sample — A total of 430 mathematicians began the web-based version of the survey; however, 104 did not supply any typicality ratings, for a final sample size of $N = 326$. Of the 326 who supplied some data, 280 reached the end of the survey; partial survey responses were used in the analysis. The high attrition rate among the mathematicians was likely due to the context of the survey – it was voluntary and the mathematicians received no incentive to complete it, and it was sent as a blanket invitation to their email accounts.

Of the 326 mathematicians, over half (185) were working on a Ph.D. in Mathematics or Applied Mathematics, while most of the others (106) already held a Ph.D. in Mathematics or Applied Mathematics. Twelve participants were working on a Master's degree in Mathematics or Applied Mathematics, while one was working on a Bachelor's in this field. The remainder ($n = 22$) reported working on Master's or Doctorate degrees across a variety of fields, including business, statistics, engineering, and mathematics education. Of the 326 mathematicians, 273 supplied gender data – there were 191 Males (70%) and 81 Females (29.7%). Country of origin was reported by 264 mathematicians (80.7% United States, 19.3% other). English was also the first language of the majority of participants (76.7% of 275).

Instruments

Middle school student survey instrument — Each survey form contained questions from two of four different domains: numbers, parallelogramsⁱⁱ, triangles, and birds (birds are omitted from this analysis). For each domain, middle school students were presented with mathematical objects or items in that domain (e.g., a small equilateral triangle or the number “6”) and asked to rate each item's typicality on a 1 (*not at all*) – 7 (*very much*) scale (Figure 1) in a mathematical context and in an everyday context. These contexts were accompanied by instructions that told the student to consider the survey item in the relevant way, referring the student to their experiences in math class or at home. Mathematical typicality pertained to whether conjectures that held for the mathematical object would generalize to most other objects of that type. Although this is a generic prompt (as it does not give a particular conjecture), given that we were administering a large-scale survey to make broad comparisons between contexts, objects, and groups, it made the most sense for our research purpose.

Everyday typicality asked about the typicality of the item in everyday life. The full text of these instructions is shown on the left side of [Figure 1](#); these instructions were only given once at the beginning of each section. In addition, each question, as shown in the second column of [Figure 1](#), re-iterated the context.

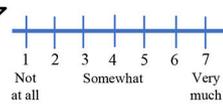
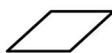
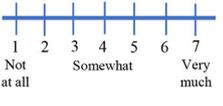
<p>Mathematical Context <i>“Now we want you to think about mathematical properties – that is, the kinds of things you learn about in math class.”</i></p>	<p>Imagine that we learned a new mathematical property that was true of this parallelogram. How likely is it that the property will be true of most other parallelograms?</p>		
<p>Everyday Context <i>“Now we want you to think about parallelogram shapes you see in everyday life outside of school, for example, in objects around your house or outside.”</i></p>	<p>How typical is this parallelogram of those you see in your everyday life?</p>		

Figure 1. Example of items from survey instrument.

Item selection — Mathematical objects to be placed on the survey were selected by the researchers to either cover the space of possible mathematical properties in the domain (e.g., the parallelogram in [Figure 1](#) is a rectangle; we also included squares, rhombi, etc. to include the mathematically special classes of parallelograms) or to have few properties that would distinguish the object mathematically (e.g., a long, skinny rhomboid with no 90-degree angles). See [Table 2](#). We also chose items that were common in everyday life (e.g., the number 10) and items that were uncommon in everyday life (e.g., the number 102), following our preceding discussion. See [Appendix](#) for a complete list of the items used on the survey.

Survey design — The number items were divided into three sets; there were two sets of parallelograms and two sets of triangles. There was also an additional set where middle school students were asked to judge the similarity or dissimilarity of two items, not considered here. Each student received two of these sets on their survey, and for each set they were asked to give ratings of the items in the set in both the mathematical and everyday contexts. Whether they rated items in a mathematical context or everyday context first was randomized. The order of the items (the 9 triangles, 9 numbers, or 10 parallelograms) within each context was randomized, as was the order of the 2 sets each student received.

Mathematician survey instrument — On Version 1 of the mathematician survey (given to 186 of the 326 participants), mathematicians would rate 7 number items in a mathematical context, and 7 number items in an everyday context. Then they would either rate 5 triangles or 5 parallelograms in a mathematical and everyday context. The instructions, prompts, and rating scales were all identical the middle school student survey to ensure strict comparability ([Figure 1](#)). The specific prompts were also identical – for everyday typicality, participants were asked “How typical is this number/triangle/parallelogram of those you see in your everyday life?” and for mathematical typicality participants were asked “Imagine that we learned a new mathematical property that was true of this number/triangle/parallelogram. How likely is it that the property will be true of most other numbers/triangles/parallelograms?” ([Figure 1](#)).

The items each mathematician received were randomly selected by the Qualtrics survey environment from a set of items that was twice as large (sets of 14 numbers, 10 triangles or parallelograms). A new version of the survey (Version 2) was introduced approximately halfway through the data collection period and was taken by 140 of the 326 participants. In this version the number and shape items were different, and the participants would rate a total of 8 number items in each context instead of 5. When collapsing both versions, the mathematician survey used the same number and shape items as the middle school student survey, except that one parallelogram and one triangle were left off of the mathematician survey. However, two additional number items (163452 and 3432984) were added to the mathematician survey to see if mathematicians rated the typicality of very large numbers differently than smaller numbers.

Data Collection Procedures

Middle school student data collection procedures — Middle school students completed the paper-based survey during their regular math class. They would first respond to a prompt about how typical it was to eat oatmeal for breakfast to familiarize themselves with the Likert rating scale and the idea of typicality. They were instructed to answer every question on the survey and were told that there were no “right” or “wrong” answers. Demographic data relating to gender, grade level, and teacher was requested.

Mathematician data collection procedures — Mathematicians were recruited by sending emails to mathematics departments at 39 universities in the United States. The email was sent to department personnel with the request that it be forwarded to mathematics faculty and doctoral students. The email invited the mathematicians to participate in an online survey in Qualtrics – there was no compensation offered and participation was voluntary. The email and the introduction to the survey stated that the purpose of this research was to better understand how middle school students use evidence in mathematics by examining mathematician data on example usage. Demographic data relating to gender, education, area of specialization within mathematics, language, and country of origin was collected from the mathematicians.

Results

R1: Mathematical Versus Everyday Typicality

The first part of the first research question asked whether participants distinguished between two contexts: everyday and mathematical typicality. We addressed this question by calculating and visualizing the correlation between everyday context and mathematical context typicality ratings for each group (mathematicians, middle school students) in each domain (numbers, parallelograms, triangles). We computed Pearson correlations by averaging for each group their ratings of each item within a domain, for everyday and mathematical contexts. We then looked at the correlation between those two sets of item averages.

Middle school students' everyday and mathematical typicality ratings were strongly correlated for the triangle domain, $r = .934$, $SE = 0.089$, the parallelogram domain, $r = .886$, $SE = 0.109$, and the number domain, $r = .856$, $SE = 0.103$. In contrast, mathematicians' everyday and mathematical typicality ratings were negatively correlated for triangles $r = -.841$, $SE = 0.140$ and parallelograms $r = -.834$, $SE = 0.137$, and weakly negatively correlated for number items, $r = -.407$, $SE = 0.176$. Middle school students tended to see numbers and shapes that were typical in their everyday lives as mathematically typical, and the numbers that they considered everyday atypical were also treated as mathematically atypical. In contrast, mathematicians tended to see highly familiar and common shapes (everyday typical) as mathematically atypical, and they rated unfamiliar,

uncommon shapes as mathematically more typical, resulting in a negative correlation. Additionally, for mathematicians there was a less reliable relation between everyday and mathematical typicality for numbers.

We explored these trends further by plotting the average mathematical typicality rating for each item versus the average everyday typicality rating for each item for both middle school students and mathematicians. If mathematical typicality and everyday typicality are not distinct, we would expect all items to fall on the line $y = x$, meaning that their everyday typicality rating (x coordinate) and mathematical typicality rating (y coordinate) are identical. This graphical approach highlights items whose typicality varies based on the context, allowing us to see which items are rated differently in a mathematical versus everyday context.

Figure 2 (6 panels) presents mathematicians' and middle school students' average mathematical and everyday ratings for number, triangle, and parallelogram items. The items in the upper half of each plot are items that were considered typical in a mathematical context, meaning that participants thought properties that hold for these items are likely to generalize. Items in the bottom half of each plot are items that were considered mathematically atypical, meaning that generalization from these examples to all items would generally be weaker. Items on the right side of the plot are considered to have high everyday typicality, with the items on the left side had been rated as relatively uncommon in everyday life. These plots reveal general trends for what objects are considered typical/atypical, the second part of our first research question. Middle school students found the numbers 1, 2, 10 and 25 to be typical in both mathematical and everyday contexts, while numbers like 83, 102, and 57 were atypical in both contexts. Mathematicians showed somewhat similar trends for everyday typicality, but only found 0 and 1 to be mathematically atypical. Middle school students found groups of triangles and parallelograms that could be described as part of well-known classes (e.g., equilateral, rectangle – see Figure 2) as typical in both contexts, while mathematicians found these same objects typical in an everyday context but atypical in a mathematical context.

Of interest in the number domain are a group of apparent outliers where everyday and mathematical typicality do not correspond. Mathematicians rated the additive and multiplicative identity elements (0 and 1) as less mathematically typical than other numbers, despite their high level of commonness in everyday life. For middle school students, there is one extreme outlier in the number domain – 13 – which is rated as less mathematically typical than other numbers of similar everyday typicality (e.g., 14, 11). This deviation likely reflects the status of 13 as a “superstitious” number. However, it is also possible that 13 is a salient example of the mathematical property of being a prime. A second outlier, 10,000, was rated by middle school students as more mathematically typical than other numbers, but less everyday typical. This rating may reflect the special significance of powers of 10 in our base ten counting system. Thus there was evidence that middle school students may distinguish the mathematical and everyday typicality of numbers. However, it is not clear that they are basing their judgments on mathematically significant features of the numbers. We see fewer obvious outliers in the shape domain – middle school students have a very consistent pattern where they rate shapes similarly in everyday and mathematical contexts, and mathematicians have a very consistent pattern where they rate shapes with high everyday typicality as having low mathematical typicality.

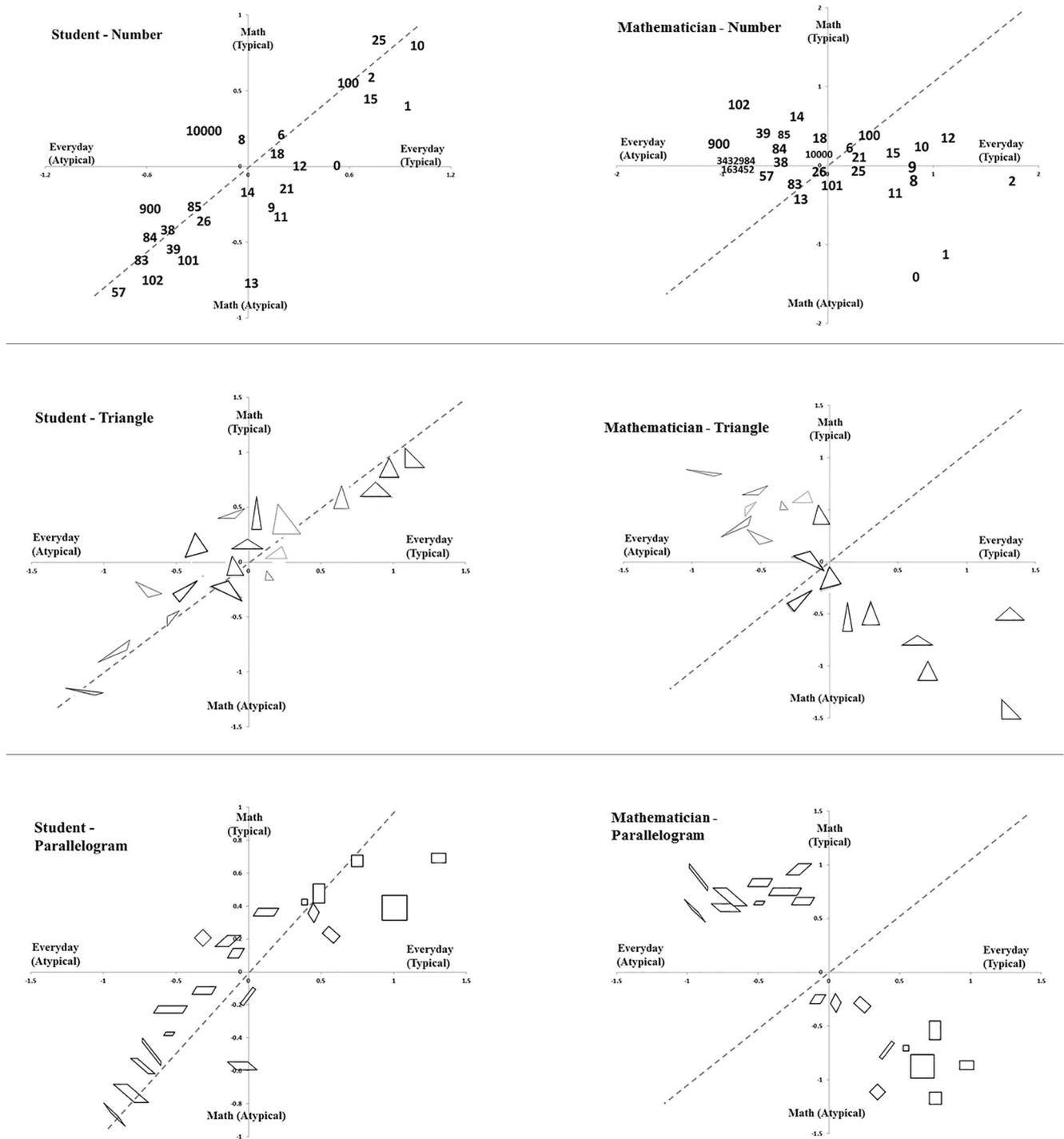


Figure 2. Middle school students' and mathematicians' average ratings of numbers (top), triangles (middle), and parallelograms (bottom), in a mathematical (y-axis) and everyday (x-axis) context, normalized.

R2: Properties That Make Objects Typical or Atypical

The second research question asked what specific properties make objects typical or atypical for mathematicians and for middle school students. The preceding analyses demonstrate that mathematicians and middle school students show very different judgments about mathematical typicality. Moreover, they suggest

that middle school students may use everyday typicality as the basis for judging mathematical typicality. To explore this issue further, we used a list of specific properties relevant to typicality, determined from prior research (Table 2).

Data were analyzed using mixed-effects linear regression models (Snijders & Bosker, 1999) run using the *lmer* command (Bates & Maechler, 2010) in the R software package (R Core Development Team, 2010). All regression models included participant ID and item ID as random effects and the typicality rating (1-7) as the dependent measure. Model selection was conducted using a chi-square test to examine significant reductions in deviance, and significance levels were computed using MCMC sampling (Baayen, 2008). We fit a separate model for each domain: number, triangle, and parallelogram, as each domain has different specific properties of import. We also fit separate models for middle school students and mathematicians. The models included context (mathematical, everyday) as a fixed effect, indicator variables for the specific mathematical and everyday properties (Table 2), and the interaction of context and the properties. Ratings in everyday and mathematical contexts was placed into a single model, rather than separate models, for several reasons – primarily, it allowed us to detect cases where, for instance, a property increased typicality in both contexts, but increased it *significantly more* in one context. In addition, keeping the data together allows for better estimation of participant and item random effects, and reduces Type 1 error.

Tables 3, 4 and 5 present regression results showing the contribution of various properties to typicality judgments. In each regression, one level of a factor was designated as the reference level. The coefficients indicate the magnitude of the effect with a change to the other level of the factor (all factors were binary, except magnitude). For example, in Table 3 (number) mathematical typicality is designated the reference level, as is “being a large number.” Other reference categories for the binary predictors are not explicitly labeled, but are the opposite of the level indicated – for example, the references levels include *not* being an Identity Element, *not* Ending in 5, etc. Rows that are blank indicate variables that were significant in one model (e.g., mathematician), but not significant in the other model (e.g., student).

In general, the analyses presented in Tables 3, 4 and 5 support the qualitative patterns suggested in Figure 2. For mathematicians, only Identity had a significant effect on judgments of mathematical typicality – it decreased mathematical typicality. For middle school students, Small Magnitude, Multiple of 5, Power of 10, and Even Parity increased mathematical typicality. Thus middle school students felt that conjectures were more likely to generalize if the numbers tested were small or even numbers, or multiples of 5 or powers of 10. For middle school students, magnitude had a significantly larger effect on everyday than mathematical typicality, but significantly increased typicality in both contexts. Notably, only middle school students’ everyday typicality judgments were influenced by a number being 0 or 1 (which we refer to as an identity element), not their mathematical typicality judgments. Mathematicians and middle school students both saw 0 and 1 as more everyday typical than non-identity numbers, but mathematicians saw Identity Elements as less mathematically typical than other numbers. Mathematicians made sharper distinctions between the mathematical and everyday typicality, rating primes, perfect squares, powers of 2, and even numbers as significantly more everyday typical than mathematically typical.

Table 3

Middle School Students' (Left) and Mathematicians' (Right) Regression Model Outputs for the Effects of Mathematical and Everyday Properties of Numbers on Typicality Ratings

Predictor	Middle School Student				Mathematician			
	B	SE(B)	t	p	B	SE(B)	t	p
(Intercept)	4.16	0.45			2.41	0.11		
Context Mathematical	(ref.)				(ref.)			
Context Everyday	0.03	0.17	0.16	.8644	0.19	0.07	2.72	.0056
Magnitude Large	(ref.)				(ref.)			
Magnitude Small	0.76	0.15	5.07	.0001	-0.02	0.12	-0.18	.8178
Ends in 5	0.96	0.25	3.81	.0016	-0.09	0.20	-0.43	.6662
Power of 10	1.25	0.21	5.91	.0001	0.04	0.17	0.24	.8572
Odd	-0.55	0.16	-3.39	.0036	-0.03	0.13	-0.26	.8640
Identity	-0.27	0.30	-0.91	.3792	-0.72	0.25	-2.82	.0088
Context Everyday: Magnitude Small	0.59	0.09	6.74	.0001	1.41	0.08	18.00	.0001
Context Everyday: Odd	0.31	0.08	3.64	.0002	-0.43	0.09	-4.76	.0001
Context Everyday: Identity	0.46	0.16	2.95	.0026	1.23	0.17	7.39	.0001
Magnitude Extra Large					-0.20	0.22	-0.88	.4340
Prime					-0.19	0.17	-1.14	.2588
Perfect Square					-0.09	0.15	-0.59	.5176
Power of 2					0.02	0.20	0.12	.8188
Context Everyday: Magnitude Extra Large					-1.04	0.15	-7.11	.0001
Context Everyday: Prime					0.55	0.12	4.69	.0001
Context Everyday: Perfect Square					0.42	0.10	4.13	.0002
Context Everyday: Power of 2					0.64	0.14	4.74	.0001

Note. Only mathematicians received numbers that had "Magnitude: Extra Large." Other blank rows indicate that predictors were not significant for that group.

Analyses of triangles and parallelograms showed similar patterns. Mathematicians judged Isosceles and Equilateral triangles as less mathematically typical than non-Isosceles and non-Equilateral triangles (respectively). However, the effects of these dimensions reversed for everyday typicality (e.g., Isosceles more typical than non-Isosceles). Right triangles were more everyday typical than non-Right triangles for mathematicians. Middle school students judged non-Skinny, Isosceles, and Equilateral triangles to be more typical in both contexts, although equilateral triangles were significantly more typical in an everyday context, compared to a mathematical context. Standard orientation and Right contributed significantly to middle school students' everyday typicality but not mathematical typicality judgments of triangles. Thus for triangles, mathematicians recognized that special mathematical properties (Isosceles, Equilateral) make triangles less mathematically typical, while middle school students had the opposite trend. In an everyday context, both groups recognized properties (Equilateral, Isosceles, Right) that make triangles more typical in an everyday context.

Table 4

Middle School Students' (Left) and Mathematicians' (Right) Regression Model Outputs for the Impact of Mathematical and Everyday Properties of Triangles on Typicality Ratings

Predictor	Middle School Student				Mathematician			
	B	SE(B)	t	p	B	SE(B)	t	p
(Intercept)	4.97	0.41			3.90	0.15		
Context Mathematical	(ref.)				(ref.)			
Context Everyday	0.15	0.20	0.75	.4572	-1.00	0.17	-5.90	.0001
Skinny	-0.74	0.17	-4.28	.0024				
Isosceles	0.63	0.17	3.64	.0042	-1.21	0.11	-11.30	.0001
Equilateral	0.75	0.33	2.28	.0374	-1.91	0.20	-9.48	.0001
Acute Triangle	(ref.)				(ref.)			
Obtuse Triangle	-0.05	0.19	-0.25	.7942	0.27	0.13	2.02	.0664
Right Triangle	0.20	0.27	0.75	.4530	-0.25	0.17	-1.48	.1874
Standard Orientation	0.59	0.18	3.19	.0096	0.06	0.13	0.44	.6762
Context Everyday: Equilateral	0.68	0.21	3.19	.0016	2.96	0.27	10.81	.0001
Context Everyday: Obtuse Triangle	-0.16	0.10	-1.56	.1242	-0.36	0.18	-2.00	.0562
Context Everyday: Right Triangle	0.46	0.16	2.91	.0042	0.90	0.23	3.95	.0001
Context Everyday: Isosceles					1.75	0.15	11.82	.0001
Context Everyday: Standard Orientation					0.34	0.18	1.91	.0744

Note. Blank rows indicate that predictors were not significant for that group.

For parallelograms, mathematicians judged Squares to be less mathematically typical than non-Squares, Rectangles tend to be less mathematically typical than non-Rectangles, and Rhombi less mathematically typical than non-Rhombi. However, in each case these effects reversed for everyday typicality (e.g., Squares more typical than non-Squares). Parallelograms with near Golden-ratio proportions were rated by mathematicians as more everyday typical than other parallelograms, but proportion had no significant effect on mathematical typicality. For middle school students, being Square and being Rectangular significantly increased mathematical typicality. These same features also increased everyday typicality, with squares being significantly more everyday typical than mathematically typical. Several other features, including Size, Orientation, and Skew, affected everyday but not mathematical typicality. For mathematicians, special mathematical properties like Rectangle and Square increased everyday typicality but decreased mathematical typicality; for middle school students, these properties increased both ratings. Middle school students and mathematicians also attended to other more everyday properties (Golden-ratio, Orientation) when considering everyday typicality.

Table 5

Middle School Students' (Left) and Mathematicians' (Right) Regression Model Outputs for the Impact of Mathematical and Everyday Properties of Parallelograms on Typicality Ratings

Predictor	Middle School Student				Mathematician			
	B	SE(B)	t	p	B	SE(B)	t	p
(Intercept)	2.87	0.61			4.63	0.11		
Context Mathematical	(ref.)				(ref.)			
Context Everyday	0.39	0.27	1.41	.1650	-1.59	0.11	-15.10	.0001
Square	0.88	0.34	2.58	.0150	-2.50	0.15	-16.57	.0001
Rectangle	0.63	0.31	1.99	.0548	-1.99	0.20	-9.87	.0001
Rhombus	0.57	0.40	1.44	.1538	-1.53	0.20	-7.81	.0001
Standard Orientation	0.43	0.24	1.78	.0798				
Size Large	(ref.)							
Size Small	0.19	0.55	0.34	.7104				
Leans Left	-0.51	0.34	-1.52	.1336				
Context Eday: Standard Orientation	0.49	0.11	4.49	.0001				
Context Eday: Size Small	-1.04	0.25	-4.20	.0001				
Context Eday: Lean Left	0.52	0.15	3.48	.0002				
Context Eday: Square	0.55	0.15	3.64	.0002	4.13	0.19	21.52	.0001
Context Eday: Rectangle	1.31	0.14	9.30	.0001	3.23	0.26	12.42	.0001
Context Eday: Rhombus	0.55	0.18	3.14	.0014	2.21	0.26	8.60	.0001
Golden					-0.05	0.26	-0.19	.9048
Context Eday: Golden					0.77	0.34	2.28	.0246

Note. Eday = every day. Blank rows indicate that predictors were not significant for that group.

General Discussion

Expertise is often characterized by distinctive judgments regarding typicality relations among objects in a domain. In mathematics, it is reasonable to assume that experts see different connections among objects than others. Psychological principles of inductive inference have been explored in other domains, such as living things (Osherson et al., 1990). However, little research has examined how these inductive principles may have relevance in mathematics. In addition, while deductive reasoning is generally thought to become fully available during adolescence (i.e., middle school), little prior research has examined when and how principles of strong versus weak *inductive* inference in mathematics tend to develop in children, and how understanding of such principles can be facilitated. Here, we explored the nature of mathematicians' and middle school students' judgments of mathematical objects with respect to *typicality relations* – judgments of how appropriate it is to generalize from a particular object to the domain. Results suggest that mathematicians have some special and somewhat consistent criteria for judging the typicality of objects in their domain. Middle school students, on the other hand, did not yet have fully developed ideas related to the significance of mathematically-specific organizations for examples. The appropriate use of principles of inductive inference, such as typicality, has been found to develop early in domains other than mathematics (e.g., López, Gelman, Gutheil, & Smith, 1992). Children can recognize which members are typical instances of their class, and use typicality as a basis to draw valid inferential conclusions. However, as mathematical domains have the problem of conflating everyday and mathematical typicality, it does not appear that development happens in the same manner.

Everyday Typicality

Mathematicians and middle school students both showed robust and consistent judgments about what we termed “everyday” typicality. Some numbers and shapes were consistently cited as more common in everyday life than others (e.g., small numbers and equilateral shapes). Middle school students’ reasoning often cited that these objects are encountered more often in their experience. People may have perceptual (Feldman, 2000) or experiential biases from interacting with the world (Lakoff, 1990) that lead them to represent some objects as more typical and then report that typical objects are most common. Further, items with high utility in everyday life (smaller numbers, “nice” numbers) tend to be used more often, and may thus be more salient in everyday contexts. We suspect that everyday typicality judgments reflect a mix of frequency, expectations, and utility. The numbers 1 and 10 were cited as being highly typical in everyday life, likely due to both the frequency with which they are encountered and their usefulness for making quick calculations within Base-Ten. In sum, mathematicians and middle school students generally agreed with each other on their ratings of everyday typicality.

Mathematical Typicality

Both studies showed interesting contrasts between this everyday sense of typicality, and a “mathematical” sense of typicality – i.e., the degree to which one mathematical object was a reliable basis for inferences about other objects. Middle school students showed robust and consistent judgments about mathematical typicality for both number and shape that did not differ from their judgments about everyday typicality. In contrast, mathematicians made sharp distinctions between everyday and mathematical typicality; these judgments were almost perfectly negatively correlated for geometric shapes. The most typical shapes in an everyday context, like squares and right triangles, were the least typical in a mathematical context. For numbers, mathematicians either did not have consistent judgments about the mathematical typicality or did not see a relevant network of typicality relations when considering a generic conjecture in this domain. Whatever this signals about the nature of expertise in the domain of numbers, the contrast with middle school students is important. Mathematical typicality is not the same as everyday typicality for mathematicians, for numbers or shapes.

Limitations

Our studies varied whether participants were given a specific mathematical conjecture to explore (like in the pilot interview study) or given a generic prompt about conjecture-testing (like in the survey study). Typicality judgments are certainly applied most productively when a specific mathematical task given. For example, when considering conjectures about additive properties, concerns relating to whether a number is positive or negative might be most paramount for generalizability. In contrast, when considering conjectures about multiplicative properties, properties relating to divisibility might matter most. The type of mathematical task may be especially important for numbers (which have many potential properties of importance) compared to shapes (which have a clearer hierarchical organization of relevant properties).

In addition, in the survey study the classes to which generalizations were invited were more specific for shapes than for numbers. Participants were asked whether facts about one number generalized to “most other numbers,” while facts shapes were generalized to “most other triangles/parallelograms” instead of the larger class of all geometric figures. There was also a difference in specificity – shapes were presented without side or angle measurements, whereas the number items were specific – “13” or “102.” Finally, the example numbers

were all one special subtype of numbers: whole numbers. If respondents interpreted the task as involving generalization to all types of numbers (rational, real, imaginary) then they may have viewed all whole numbers other than 0 and 1 as roughly equally typical.

Implications

A recent publication of the National Council of Teachers of Mathematics (Carter et al., 2009), *Focus in High School Mathematics: Reasoning and Sense Making*, argues that U.S. high school students are not being adequately prepared for the demands of a global, technological society. They argue that an increased focus on reasoning and sense-making in mathematics is critical, stating that “In addition to formal reasoning or proof, reasoning in mathematics often begins with explorations, conjectures, or false starts” (p. 1), thus allowing examples to play a critical role. Similarly, the *Common Core State Standards for Mathematical Practice* (Common Core State Standards Initiative, 2010) acknowledge the value of inductive inference for creating and critiquing mathematical arguments. Significant work has been conducted on inductive principles that people use when drawing inferences in other domains (Feeney & Heit, 2007). Many have argued that people’s intuitive principles for inductive inferences are generally congruent with normative principles of statistical inference (Oaksford & Chater, 1999). We explored one such principle, typicality, and found that middle school students’ mathematical inferences are influenced by typicality, but their notion of typicality is based on everyday rather than mathematical properties.

Implications for Typicality Relations in Mathematics

If middle school students are using their everyday senses of typicality to reason about mathematical relations, a central question these studies confront is whether there is an alternative network of typicality relations in the domain of mathematics that would serve them better. For this reason, we included Research Question 2, where we examined whether there are overarching properties of numbers and shapes considered important for generalizability by mathematicians. For the domain of Euclidean geometric shapes, it appears that such a network of properties may exist – mathematicians recognize a network of special geometric properties that limit the potential of an example shape for inductive generalization. Middle school students may have not yet adopted this lens for thinking about the mathematical properties of their examples, which may limit the quality of the inductive arguments they generate, and the utility of these arguments for allowing them to understand geometric properties and discover more general proofs.

The story for whole numbers is more complex – there does not seem to be a clear network of typicality relations that is recognized, but rather localized, micro-level properties of number (like prime-ness or parity) that become relevant when considering particular conjectures. Much of the situated work of mathematicians involves determining which properties are important to consider in the context of a conjecture, and then taking those properties into account, when possible, when exploring the conjecture (Lockwood et al., 2016). This type of expertise may be difficult to teach, and may be garnered only through significant experience in the domain. Middle school students have ideas about specific properties of number that limit inductive generalization; they identify that numbers common in everyday life that often have special properties – like Base-Ten numbers – are generally better candidates for generalization. Depending on the context of the conjecture they are working with, this could result in middle school students choosing examples that result in weakly-supported inferences. It may also be reflective of “ease of use” considerations which are not directly related to (but middle school students may confound with) generalizability. It is significant that for both domains, middle school students

generally noticed the same mathematical properties mathematicians did – they just struggled to apply this knowledge in a way that follows principles of inductive inference.

In other research, we have questioned middle school students on whether typicality considerations are important to take into account when they are proving conjectures (Cooper et al., 2011), or when they are evaluating the empirical proofs of others (Cooper, Dogan, Young, & Kalish, 2012). Results suggest that while middle school students do have good informal ideas about the relationship between typicality and justification through inductive inference, formally applying these principles in practice is challenging. For example, one student in Cooper et al. (2011) described how “If he didn't use unusual numbers, you know, you can never be sure if his property is correct,” while another said “If I used some ones that you people wouldn't normally use, besides 10, and if I did a little more maybe it wouldn't be or maybe it'd still be true.” In addition, middle school students in Cooper et al. (2012) recognized that testing a conjecture with more examples is better than fewer examples, and that using dissimilar examples is better than similar examples. Thus middle school students seem to have the building blocks of a more systematic sense of typicality as it relates to mathematical conjectures, but this relationship needs further development in the classroom.

Implications for Instruction

With generalization being framed as central to mathematics education, our results suggest that classrooms would benefit from exploration of the properties of the examples K-12 students choose, and discussion of how these properties interact with ideas relating to generalization in the context of the particular conjecture at-hand. Instruction on exploring conjectures could include discussions about the typicality of mathematical objects – i.e., which objects it makes the most sense to test as K-12 students first consider the conjecture, as they try to “break” the conjecture, and as they try to find a pattern to prove the conjecture, and why. Understanding that special mathematical properties of an object should be a consideration when determining whether a conjecture is likely to generalize to all other objects could be another instructional goal relating to the nature of inductive evidence in mathematical domains. Such discussions would also include what constitutes strong versus weak evidence in inductive reasoning. It is also important to consider how reasoning inductively can support the development of deductive proofs. Testing examples can be especially useful for understanding the structure of conjectures, or how they work, in order to reveal an underlying mathematical pattern. Finally, everyday aspects of typicality could be discussed – K-12 students could critically examine whether surface features like a number's magnitude or digit patterns or a shape's skinniness or orientation are important considerations in the context of their current mathematical exploration.

Concluding Remarks

Experts understand objects and ideas in their domain of expertise as a complex interwoven network of relationships based on important and useful properties. This network of relationships allows them to strategically confront and solve novel tasks through a consideration of relevant domain principles. Typicality is one such principle that is relevant to practices of inductive inference, but this principle has mainly been identified as being useful in fields other than mathematics. Here we provide evidence that mathematicians have a distinct sense of mathematical typicality that can guide their activities related to justification and generalization.

Recent reform movements in mathematics education have pushed for mathematical argumentation and generalization in elementary and middle school. A long tradition of research in math education has demonstrated that K-12 students do use empirical strategies of inductive inference to make sense of content they learn during mathematics instruction (Knuth et al., 2011). In addition, K-12 students have rich ideas about the characteristics and properties of the mathematical objects they interact with every day, like numbers and shapes (Cooper et al., 2011; Williams et al., 2011). Some of these characteristics relate to mathematical properties, while others rely on experience from their everyday activities. We argue that young students have the raw materials needed to build the sophisticated network of typicality relations that are useful when exploring mathematical ideas, if instructional approaches are used that support these activities.

Notes

i) Portions of this data were previously presented at the 2012 the Psychology of Mathematics Education – North America (PME-NA) Conference, the 2013 Annual Meeting of the American Educational Research Association, and the 2011 Annual Conference of the Cognitive Science Society.

ii) A parallelogram was defined for the middle school students on the survey: “A parallelogram is a 4-sided figure with opposite sides that are parallel.”

Funding

The research was supported in part by the NSF, Award DRL-0814710. Any opinions, findings, and conclusions or recommendations are those of the authors and do not necessarily reflect the views of the NSF. It was also supported through funding from the U.S. Department of Education Institute of Education Sciences (Award No. R305B100007) Postdoctoral Training Program in Mathematical Thinking, Learning, and Instruction.

Competing Interests

The authors have declared that no competing interests exist.

Acknowledgments

The authors wish to thank Eric Knuth, Amy Ellis, Elise Lockwood, Fatih Dogan, and Andrew Young, for their contributions to this work.

References

- Alcock, L., & Inglis, M. (2008). Doctoral students' use of examples in evaluating and proving conjectures. *Educational Studies in Mathematics*, 69(2), 111-129. doi:10.1007/s10649-008-9149-x
- Armstrong, S. L., Gleitman, L. R., & Gleitman, H. (1983). What some concepts might not be. *Cognition*, 13(3), 263-308. doi:10.1016/0010-0277(83)90012-4
- Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. New York, NY, USA: Cambridge University Press.
- Bates, D., & Maechler, M. (2010). lme4: Linear mixed-effects models using Eigen and Eigen++ (R package version 0.999375-35) [Computer software]. Retrieved from <http://CRAN.R-project.org/package=lme4>

- Bédard, J., & Chi, M. T. H. (1992). Expertise. *Current Directions in Psychological Science*, 1, 135-139. doi:10.1111/1467-8721.ep10769799
- Buchbinder, O., & Zaslavsky, O. (2011). Is this a coincidence? The role of examples in fostering a need for proof. *ZDM*, 43(2), 269-281. doi:10.1007/s11858-011-0324-7
- Carter, J., Forster, S., Howe, R., Kader, G., Kepner, H., & Valdez, P. (2009). *Focus in high school mathematics: Reasoning and sense making*. Reston, VA, USA: National Council of Teachers of Mathematics.
- Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5, 121-152. doi:10.1207/s15516709cog0502_2
- Chi, M. T. H., & VanLehn, K. A. (2012). Seeing deep structure from the interactions of surface features. *Educational Psychologist*, 47(3), 177-188. doi:10.1080/00461520.2012.695709
- Common Core State Standards Initiative. (2010). *Common Core State Standards (Mathematics Standards)*. Retrieved from <http://www.corestandards.org/the-standards/mathematics>
- Cooper, J. L., Dogan, M. F., Young, A. G., & Kalish, C. W. (2012). Stronger arguments within inductive generalization in middle school mathematics. In L. R. Van Zoest, J.-J. Lo, & J. L. Kratky (Eds.), *Proceedings of the 34th annual meeting of Psychology of Mathematics Education, North American Chapter (PME-NA)* (pp. 331-334). Kalamazoo, MI, USA: Western Michigan University.
- Cooper, J., Walkington, C., Williams, C., Akinsiku, O., Kalish, C., Ellis, A., & Knuth, E. (2011). Adolescent reasoning in mathematics: Exploring middle school students' strategic approaches to empirical-based justifications. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 2188-2293). Boston, MA, USA: Cognitive Science Society.
- Dreyfus, T. (1999). Why Johnny can't prove. *Educational Studies in Mathematics*, 38(1-3), 85-109. doi:10.1023/A:1003660018579
- English, L. D. (1997). Interventions in children's deductive reasoning with indeterminate problems. *Contemporary Educational Psychology*, 22(3), 338-362. doi:10.1006/ceps.1997.0939
- Epstein, D., & Levy, S. (1995). Experimentation and proof in mathematics. *Notices of the American Mathematical Society*, 42(6), 670-674.
- Feeney, A., & Heit, E. (2007). *Inductive reasoning: Experimental, developmental, and computational approaches*. Cambridge, United Kingdom: Cambridge University Press.
- Feldman, J. (2000). Bias toward regular form in mental shape spaces. *Journal of Experimental Psychology: Human Perception and Performance*, 26, 152-165. doi:10.1037/0096-1523.26.1.152
- Gelman, S. A. (1988). The development of induction within natural kind and artifact categories. *Cognitive Psychology*, 20(1), 65-95. doi:10.1016/0010-0285(88)90025-4
- Gelman, S. A., & Coley, J. D. (1990). The importance of knowing a dodo is a bird: Categories and inferences in 2-year-old children. *Developmental Psychology*, 26(5), 796-804. doi:10.1037/0012-1649.26.5.796

- Gelman, S. A., & Kalish, C. W. (2005). Conceptual development. In W. Damon & R. M. Lerner (Eds.-in-Chief), D. Kuhn & R. Siegler (Vol. Eds.), *Handbook of child psychology: Vol. 2. Cognition, perception, and language* (6th ed., pp. 687-733). New York, NY, USA: John Wiley & Sons.
- Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory*. New Brunswick, NJ, USA: Aldine.
- Goldenberg, P., & Mason, J. (2008). Shedding light on and with example spaces. *Educational Studies in Mathematics*, 69(2), 183-194. doi:10.1007/s10649-008-9143-3
- Gopnik, A., Glymour, C., Sobel, D., Schulz, L., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review*, 111(1), 3-32. doi:10.1037/0033-295X.111.1.3
- Hacking, I. (2001). *An introduction to probability and inductive logic*. Cambridge, United Kingdom: Cambridge University Press.
- Harel, G., & Sowder, L. (1998). Students' proof schemes. In E. Dubinsky, A. Schoenfeld, & J. Kaput (Eds.), *Research on collegiate mathematics education* (Vol. 3, pp. 234-283). Providence, RI, USA: American Mathematical Society.
- Healy, L., & Hoyles, C. (2000). A study of proof conceptions in algebra. *Journal for Research in Mathematics Education*, 31(4), 396-428. doi:10.2307/749651
- Herbst, P. G. (2002). Establishing a custom of proving in American school geometry: Evolution of the two-column proof in the early twentieth century. *Educational Studies in Mathematics*, 49(3), 283-312. doi:10.1023/A:1020264906740
- Kahneman, D., & Klein, G. (2009). Conditions for intuitive expertise: A failure to disagree. *The American Psychologist*, 64, 515-526. doi:10.1037/a0016755
- 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, USA: Routledge.
- Knuth, E., Kalish, C., Ellis, A., Williams, C., & Felton, M. (2011). Adolescent reasoning in mathematical and non-mathematical domains: Exploring the paradox. In V. Reyna, S. Chapman, M. Dougherty, & J. Confrey (Eds.), *The adolescent brain: Learning, reasoning, and decision making*. Washington, DC, USA: American Psychological Association.
- Koedinger, K. R. (1998). Conjecturing and argumentation in high-school geometry students. In R. Lehrer & D. Chazan (Eds.), *Designing learning environments for developing understanding of geometry and space* (pp. 319-347). Mahwah, NJ, USA: Lawrence Erlbaum.
- Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychological Review*, 96, 674-689. doi:10.1037/0033-295X.96.4.674
- Kvale, S. (1996). *InterViews: An introduction to qualitative research interviewing*. Thousand Oaks, CA, USA: Sage.
- Lappan, G., Fey, J. T., Fitzgerald, W. M., Friel, S. N., & Phillips, E. D. (2006). *Connected mathematics series* (2nd ed.). Boston, MA, USA: Pearson Prentice Hall.
- Lakoff, G. (1990). *Women, fire, and dangerous things: What categories reveal about the mind* (pp. 39-74). Chicago, IL, USA: University of Chicago Press.

- Lockwood, E., Ellis, A. B., Dogan, M. F., Williams, C., & Knuth, E. (2012). A framework for mathematicians' example-related activity when exploring and proving mathematical conjectures. In L. R. Van Zoest, J. J. Lo, & J. L. Kratky (Eds.), *Proceedings of the 34th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 151–158). Kalamazoo, MI, USA: Western Michigan University.
- Lockwood, E., Ellis, A., & Knuth, E. (2013). Mathematicians' example-related activity when proving conjectures. In S. Brown, G. Karakok, K. H. Roh, & M. Oehrtman (Eds.), *The Electronic Proceedings for the Sixteenth Special Interest Group of the MAA on Research on Undergraduate Mathematics Education* (pp. 16-30). Denver, CO, USA: Northern Colorado University.
- Lockwood, E., Ellis, A. B., & Lynch, A. G. (2016). Mathematicians' example-related activity when exploring and proving conjectures. *International Journal of Research in Undergraduate Mathematics Education*, 2(2), 165-196. doi:10.1007/s40753-016-0025-2
- López, A., Gelman, S. A., Gutheil, G., & Smith, E. E. (1992). The development of category-based induction. *Child Development*, 63(5), 1070-1090. doi:10.2307/1131519
- Lupyan, G. (2013). The difficulties of executing simple algorithms: Why brains make mistakes computers don't. *Cognition*, 129(3), 615-636. doi:10.1016/j.cognition.2013.08.015
- Martin, T. S., McCrone, S. M. S., Bower, M. L. W., & 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. doi:10.1007/s10649-005-6698-0
- 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*, 32, 49-96. doi:10.1006/cogp.1997.0645
- Mody, S., & Carey, S. (2016). The emergence of reasoning by the disjunctive syllogism in early childhood. *Cognition*, 154, 40-48. doi:10.1016/j.cognition.2016.05.012
- Murphy, G. (2002). *Big book of concepts*. Cambridge, MA, USA: MIT Press.
- National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA, USA: Author.
- Oaksford, M., & Chater, N. (Eds.). (1999). *Rational models of cognition*. London, United Kingdom: Oxford University Press.
- Osherson, D. N., Smith, E. E., Wilkie, O., Lopez, A., & Shafir, E. (1990). Category-based induction. *Psychological Review*, 97, 185-200. doi:10.1037/0033-295X.97.2.185
- Piaget, J. (1928). *Judgement and reasoning in the child*. London, United Kingdom: Routledge and Kegan Paul.
- Pillow, B. H. (2002). Children's and adults' evaluation of the certainty of deductive inferences, inductive inferences, and guesses. *Child Development*, 73(3), 779-792. doi:10.1111/1467-8624.00438
- Porteous, K. (1990). What do children really believe? *Educational Studies in Mathematics*, 21, 589-598. doi:10.1007/BF00315946
- Reyna, V. F. (2012). A new intuitionism: Meaning, memory, and development in Fuzzy-Trace Theory. *Judgment and Decision Making*, 7, 332-359.

- Rhodes, M., Brickman, D., & Gelman, S. A. (2008). Sample diversity and premise typicality in inductive reasoning: Evidence for developmental change. *Cognition*, *108*(2), 543-556. doi:10.1016/j.cognition.2008.03.002
- Sandefur, J., Mason, J., Stylianides, G. J., & Watson, A. (2013). Generating and using examples in the proving process. *Educational Studies in Mathematics*, *83*(3), 323-340. doi:10.1007/s10649-012-9459-x
- Sinclair, N., Watson, A., Zazkis, R., & Mason, J. (2011). The structuring of personal example spaces. *The Journal of Mathematical Behavior*, *30*(4), 291-303. doi:10.1016/j.jmathb.2011.04.001
- Snijders, T., & Bosker, R. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. London, United Kingdom: Sage.
- Stylianides, A. J. (2007). Proof and proving in school mathematics. *Journal for Research in Mathematics Education*, *38*(3), 289-321.
- Stylianides, G. J. (2008). An analytic framework of reasoning-and-proving. *For the Learning of Mathematics*, *28*, 9-16.
- Ward, S. L., & Overton, W. F. (1990). Semantic familiarity, relevance, and the development of deductive reasoning. *Developmental Psychology*, *26*(3), 488-493. doi:10.1037/0012-1649.26.3.488
- Watson, A., & Mason, J. (2005). *Mathematics as a constructive activity: The role of learner generated examples*. Mahwah, NJ, USA: Lawrence Erlbaum.
- Williams, C., Akinsiku, O., Walkington, C., Cooper, J., Ellis, A., Kalish, C., & Knuth, E. (2011). Understanding students' similarity and typicality judgments in and out of mathematics. In L. R. Wiest & T. Lamberg (Eds.), *Proceedings of the 32nd annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 1180-1189). Reno, NV, USA: University of Nevada, Reno.
- Yackel, E., & Hanna, G. (2003). Reasoning and proof. In J. Kilpatrick, W. G. Martin, D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 227-237). Reston, VA, USA: National Council of Teachers of Mathematics.

Appendix

Number	Triangle	Parallelogram
0		
1		
2		
6		
8		
9		
10		
11		
12		
13		
14		
15		
18		
21		
25		
26		
38		
39		
57		
83		
84		
85		
100		
101		
102		
900		
10000		
163452*		
3432984*		

Figure A1. Items used on surveys.

*These items were on the mathematician survey but not the middle school student survey.

**These items were on the middle school student survey but not the mathematician survey.