

TRACING MIDDLE-SCHOOL STUDENTS' ARGUMENTS

Carolyn A. Maher
Rutgers University
cmaher@rci.rutgers.edu

Arthur B. Powell
Rutgers University
abpowell@andromeda.rutgers.edu

Keith Weber
Rutgers University
khweber@rci.rutgers.edu

Hollylynn Stohl Lee
North Carolina State
University
Hollylynn@ncsu.edu

The purpose of this paper is to deepen our understanding of how group discussion of middle-school aged urban students can facilitate the learning of mathematics. We illustrate how student challenges and counterarguments can lead them toward investigating the deep, underlying mathematical structure involved in the situation that they are investigating and the arguments that they are making. In particular, we investigated how students made judgments about fairness of dice by running computer simulations of die tosses. Our analysis revealed that although students initially offered unsophisticated justifications for their judgments on which dice were fair, challenges to students' justifications led to a lively debate on important mathematical principles such as the importance of sample size.

Introduction

Participating in discussions in which mathematical ideas are debated provide students with powerful opportunities for learning mathematics (Balacheff, 1991; Cobb et al., 2001). McCrone argues that discussions “allow students to test ideas, to hear and incorporate the ideas of others, to consolidate their thinking by putting their ideas into words, and hence, to build a deeper understanding of key concepts” (p. 111). For these reasons, influential organizations such as the NCTM (2000) and many researchers in mathematics education (e.g., Balacheff, 1991; Heibert & Wearne, 1993; Lampert, Rittenhouse, & Crumbaugh, 1996) recommend that discussion play a prominent role in reform-oriented mathematics classrooms. Analyzing students' discussions is also considered central in analyzing how students learn mathematics (Powell & Maher, 2002). Ongoing research is building an empirical base for the role of discussion in mathematical learning (e.g., Cobb et al., 1997; Manouchehri & Enderson, 1999; Stephan & Rasmussen, 2002; McCrone, 2005). The purpose of this paper is to contribute to this literature by presenting and analyzing a discussion in which African American and Latino middle school children from a poor, urban environment make, justify, and challenge statistical claims.

Our goal in this paper is to describe and illustrate one way that discussion can foster mathematical learning. In the mathematics education literature, several accounts for how discourse can contribute to mathematical learning have been proposed. Discussion can objectify students' previous experiences, thereby making these experiences objects that can be analyzed (Cobb et al., 1997), encourage students to take a more reflective stance on their mathematical reasoning (Manhouchehri & Enderson, 1999), require students to consolidate their thinking by verbalizing their thoughts (McCrone, 2005), and help students learn to communicate mathematically and participate in a wider range of mathematical argumentation (Lampert & Cobb, 2003). We propose that group discussion can also facilitate learning by inviting students to be explicit both about the ways in which they make new claims from previously established facts and about the standards they are using in deciding whether an argument is acceptable. Further, challenges from classmates can lead students to debate whether a particular method of argumentation is appropriate, provide students with the opportunity either to justify their methods when their reasoning is sound or to revise or abandon these methods when their

reasoning is flawed. In the discussion that we will analyze, students debated what statistical conclusions legitimately could be drawn from a set of existing data. In the discussion, students considered central statistical issues such as the importance of sample size in drawing warranted conclusions.

Theoretical Perspective

Krummheuer (1995) and others (e.g., Stephan & Rasmussen, 2002; Rasmussen & Stephan, in press) have argued that Toulmin's (1969) model of argumentation can be a useful analytical tool for understanding the progression of students' arguments during collective discourse. In Toulmin's model, an argumentation consists of three essential parts, called the core of the argument: the claim, the data, and the warrant. When an individual presents an argument to his or her community, he or she is trying to convince the audience of a particular *claim*. To support the claim that is being made, the individual typically presents evidence, or *data*. The audience may ask the individual presenting the argument to explain why one should deduce the claim being made from the data being presented. In Toulmin's scheme, such an explanation is referred to as a *warrant*. In mathematical discussions, students often will not state their warrant when they are presenting their argument to their classmates. Frequently, it is a challenge by a teacher or classmate that prompts the student to explicitly state the warrant that he or she had, until that point, only used implicitly (Rasmussen & Stephan, in press). It is possible for the audience to accept the data put forth by the presenter but reject the explanation for why the data support the conclusion. If this occurs, the presenter is required to provide *backing* for why the warrant is acceptable and the core of the argument is valid (cf., Stephan & Rasmussen, 2002).

A central argument advanced in this paper is that classroom discussion may foster mathematical learning by requiring students to be explicit about the warrants that they are using and requesting that they provide backing for why these warrants are legitimate, thereby establishing that their modes of reasoning are valid. As Rasmussen and Stephan (in press) note, "eliciting backings is crucial for supporting the evolution of increasingly sophisticated mathematical ideas". Without such elicitation, students may come to believe that all means of deducing new information are equally acceptable. It is also possible that the class may present an argument that illustrates why the warrant in question is not valid and should not be used to present a purported convincing argument. The arguments from our study differ from those presented in Rasmussen and Stephan in that our students were not trying to obtain a mathematical solution to a problem but were asked to examine whether various six-sided dice were fair based on data obtained from computer simulations. The claims students made were judgments about the fairness of each of the dice used in our study. The data (in Toulmin's sense) usually, but not always, consisted of data (in a statistical sense) obtained from running simulations in which the die under consideration was "rolled" multiple times. Similar to the classrooms described in Rasmussen and Stephan, what was most interesting to us as researchers was not whether a student believed a particular die was fair (the student's claim) or even the data the student used to justify this conclusion (the data) but what principles the student was using to draw conclusions from examining data (the warrant) and the ensuing debates about whether these principles were legitimate (the debate about warrants and the construction of backings).

Methods

Setting

This study is an adjunct of larger, ongoing analyses that emerge from a multi-prong, three-year research endeavor, “Informal Mathematics Learning Project” (IML), conducted in an after-school program in a partnership between the Robert B. Davis Institute of Rutgers University and the Plainfield School District in New Jersey, an economically depressed, urban area, whose school population is 98 percent African American and Latino students (1). Two primary goals of the IML project involve investigating (a) how middle-school students (11 to 13 years old) develop mathematical ideas and reasoning over time in an informal, after-school environment and exploring the relationship between agency and students’ learning as well as (b) how teachers facilitate IML sessions and attend to students’ ideas and reasoning. For two and a half academic years, including the intervening summers, we facilitated 30 sessions, 60- to 75 minutes each, with a cohort of approximately 25 students. During these sessions, students were asked to work on open-ended, well-defined tasks involving topics such as fractions, combinatorics, and probability. Throughout the study, collaboration, justification, and sense making were encouraged, and both researchers and students attended to students’ ideas and took them seriously.

This report focuses on a culminating probability and data analysis task, “Schoolopoly,” That has been used in prior studies (Stohl & Tarr, 2002; Tarr, Lee, & Rider, 2006). In this task, six hypothetical dice companies produced die that may or may not be fair and students are challenged to decide which company should supply the dice for a Schoolopoly game. Each pair of students was assigned two or three companies, and was asked to judge whether each company produced fair dice by simulating rolling their dice using *Probability Explorer* software. Each die company was explored by at least two groups. After their investigation, students produced a poster of their findings, including the data and graphical displays from their simulations and justifications for why they believed their judgment was correct. Students then explored the posters that each group constructed. Our analysis in this paper focuses on a 30-minute discussion that followed students’ analysis of posters as students debated which company to choose for the Schoolopoly game.

Analysis

All student activities were videotaped. The data were analyzed in a manner consistent with the first stage of the research methodology prescribed by Stephan and Rasmussen (in press). Members of the research team to get a strong sense of the data viewed the videotape repeatedly. The discussion was then transcribed. Next, a description (using the methods of Powell, Fransisco, and Maher, 2003) of the videotape was constructed. The description of the videotape is a relatively objective description of what transpired in which acts of interpretation are explicitly avoided. The data was parsed into specific student argumentations. Each argument that a student presented was coded according to Toulmin’s scheme as prescribed by Stephan and Rasmussen (in press); each argument was coded in terms of the claim being made, the data to support the claim, and, when given, the warrant for how the data implied the claim. In cases where no warrant was provided, we sometimes would make a note in our codings with our interpretation of the warrant that the student seemed to be implicitly using. Each coding and all interpretations were discussed within our research team until all disagreements were resolved. We also coded all challenges to a students’ argument by what part of the argument—the claim,

data, or warrant—was being challenged. The result of our analysis was a coded chronological account of the arguments and challenges that the students raised during the 30-minute discussion.

Results

A description of the discussion among students about which die was fair is organized by presenting an analysis of three episodes in which a student presented an argument that was challenged by his or her peers.

Episode 1.

The first episode occurs at the beginning of the discussion. The researcher begins by asking Chris to explain why he would choose to buy dice from Delta's Dice Company.

- Chris: Because, if you look at both of them [posters], they both like, like really explain the same thing. Like to me, I thought the first poster and the second poster were like about the same thing. They really explained it [...]
- R1: OK. Out of how many trials? What were they doing, do you remember?
- Chris: Uh... I think it was 600. But I don't know.
- R1: OK. So it seems to me that one of the reasons why you're picking Delta's Dice is because the two groups agree.
- Chris: Uh-huh. Because, like, in other ones [posters of other dice companies], like, one didn't agree and one did agree, or sometimes they didn't really explain enough.

We coded this excerpt in the following way:

Claim: Delta's Dice is fair

Data: Both posters agreed that Delta's Dice was fair and presented similar arguments. Posters for other dice did not agree, or did not give thorough explanations.

Warrant: None provided. We inferred Chris' warrant to be that if both posters for a die agreed, it was reasonable to conclude that die was fair.

Chris seems to be accepting that Delta's Dice was fair based on the authority of the posters that he examined, but not on the data or arguments contained in the posters. When the researcher asked Chris for details about the poster ("Out of how many trials? What were they doing?"), Chris could not recall details of the poster with confidence. He also agreed the agreement of the posters was one of the reasons he chose Delta's Dice. After Chris gave his argument, Chanel challenged it:

- Chanel: I just have a real quick question. Why does on the one scribbly and stuff, why does it say that one [outcome] is lower, one might be lower and the rest are higher, and why, how is that fair? [Referring to statement on poster "the dice are close to each other except for one might be low like number 5"] Yeah, I don't get it [...] that one, it might be lower, and the rest, the rest is just higher. So, how is, I don't get it, how is that fair?

In this excerpt, Chanel is questioning whether the data and claim made in one of the posters for Delta's Dice really is evidence that Delta's Dice is fair. Several other students raised similar questions. Chanel's challenge is arguably more sophisticated than Chris' argument since she

examines the data and argument presented in the poster and not just the conclusion expressed in the poster. However, it is also important to note that Chanel's challenge was to Chris' conclusion, but not to his reasoning. She did not question whether it was appropriate to conclude a die was fair because both posters claimed it was; rather she challenged Chris' claim because she reached a different conclusion from her examination of data and claims on one of the posters of Delta's Dice. The next two students after Chris presented similar arguments that Delta's Dice was fair, primarily relying on the fact that the two posters evaluating Delta's Dice both concluded that it was fair.

Episode 2.

Tiffany deviated from the previous students by arguing that Calibrated Cubes would be the die most likely to be fair:

R1: OK. Tiffany, what do you think?

Tiffany: Um, I picked Calibrated Cubes.

R1: OK. Can you tell us why?

Tiffany: Because, um, I think it's fair because all the numbers were even, 'cause when I looked at the charts, all the numbers had 11, I think.

Here, Tiffany is referring to a table on one of the posters that showed that on one of the samples of 80 trials, 1, 2, and 3 each occurred exactly 11 times. The table on the poster was arranged in such a way that the number of times a 4, 5, and 6 occurred were not shown. We coded this argumentation as follows:

Claim: Calibrated Cubes is a fair die.

Data: The table for Calibrated Cubes shows that 1, 2, and 3 each occurred exactly 11 times.

Warrant: None given. We inferred the warrant to be that if three sides of the die were rolled the exact same number of times in a particular simulation, the die is likely to be fair.

Note that Tiffany differed from the previous students in that the data with which she supported her claim did not concern the judgment of the posters that she inspected, but rather their contents. Chanel immediately challenged Tiffany's argument:

Chanel: I think that, um, like on Calibrated Cubes it just showed three 11's. It didn't show all of the cubes. 'Cause there were three more cubes, and those could have been 12, 13, or 14, or any other number. And it didn't show all the, um, numbers, it showed the three 11's. How do we know it wasn't like 34 or something on the other ones [outcomes 4, 5, and 6]?

Student: (off camera) Or 117.

Chanel: You agree with it!

Student: (off camera) I know but still.

Tiffany: But I have a question. Whoever that is, what was the other numbers? You don't have to lie.

In this excerpt, Chanel is challenging Tiffany's implicit warrant. It is possible that in the table that Tiffany looked at, the missing numbers could have been large, such as 34. Chanel appears to be implicitly arguing that if a 1 appeared 11 times and a 4 appeared 34 times, a

plausible possibility given the data Tiffany alluded to, then it would be unreasonable to call a die fair based on this evidence. Another student, apparently one who initially agreed with Tiffany that Calibrated Cubes was fair, chimed in that one of the missing numbers could have occurred 117 times, an impossibility since the total number of trials was 80. It is not clear whether Chanel's challenge caused this student to change his mind that Calibrated Cubes was fair, but this student clearly was attending to Chanel's challenge and building upon her reasoning. Tiffany's request to know the missing values of the table suggests that she too is attending to and appreciates the merits of Chanel's challenge. Finally, note that Chanel's challenge here differs from the one that she and others posed to Chris. Chanel did not simply challenge Tiffany's claim because she arrived at a different judgment; here she is challenging the validity of Tiffany's reasoning.

Episode 3.

Jerel challenges Tiffany's argument for a different reason than Chanel.

Jerel: Well look. They only ran it 80 times. You'll never know if another number is gonna come up and pass it. Even though it was even, they ran it a small amount of times. You need to run it a lot of times. Because

Terrill: Why, why? I don't understand why.

Jerel: I didn't say you had to, I said you *need* to [...] Because, you, all right, just like when we were doing Delta's Dice, we had ran it, um, I think a hundred times, and one number won by a lot. But when we ran it like one thousand times and all that, other numbers won....Because, like, um, other numbers won, but they were close to each other still, and the reason they got that all is because they had a little bit amount of numbers that they ran it, but when you like, I guarantee you if you ran it like 500 times, it would have been different. You ought to say it was unfair.

In this exchange, Jerel is arguing against a different aspect of Tiffany's warrant. Tiffany is drawing a conclusion based on the results of rolling the die 80 times. Jerel is questioning whether one can legitimately draw such conclusions from such a limited data set. In fact, he puts forth a counterclaim that one needs to run a large number of trials in order to reach a reliable conclusion, citing his experience in seeing discrepancies in his data when he ran simulations with 100 and 1000 trials. Jerel's counterclaim goes beyond the data Tiffany referred to and the argument that Tiffany made; it focuses on a central statistical issue—the importance of sample size. Jerel's counterclaim became the subject of intense debate among the students, with students offering arguments and counterarguments for why they did or did not need to examine data with a large sample size when judging whether a die was fair.

Summary

The first three students who offered their decision on which die to buy chose Delta's Dice, primarily because the two groups that inspected Delta's Dice both found Delta's Dice to be fair and no other dice company shared a similar level of agreement. The students making these arguments did not discuss the content of the posters. Only one of the three arguments was challenged. The challenges posed were not based on the reasoning that the student used, but on the conclusion that he reached. Tiffany was the first student to present an argument based on the data presented in the posters that she inspected. Challenges to Tiffany's arguments were based on the warrants that she appeared to be using to draw her conclusions. A central challenge to

Tiffany's data occurred when Jerel questioned whether it was appropriate to draw conclusions from a simulation that used a relatively small sample size. The issue of sample size subsequently led to a lengthy and lively debate. Subsequent challenges to students' arguments were based on the warrants that the students were using, or appeared to be using implicitly. In particular, arguments based on the conclusions of the posters but not the data presented in the posters were challenged by other students.

Endnote

(1) This research is a component of the National Science Foundation funded project, Research on Informal Mathematical Learning (REC-0309062). Any opinions, findings, conclusions and recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

- Balacheff, N. (1991). Benefits and limits of social interaction: The case of mathematical proof. *Mathematical knowledge: its growth through teaching*. A. Bishop, S. Mellin-Olsen and J. Van Dormolen. Dordrecht, Kluwer: 175-192.
- Cobb, P., Boufi, A., McClain, K., & Whitenack, J. (1997). Reflective discourse and collective reflection. *Journal for Research in Mathematics Education*, 28, 258-277.
- Cobb, P., Stephan, M., McClain, K., & Gravemeijer, K. (2001). Participating in classroom mathematical practices. *The Journal of the Learning Sciences*, 10, 113-163.
- Hiebert, J. & Wearne, D. (1993). Instructional tasks, classroom discourse, and students' learning in second-grade arithmetic. *American Education Research Journal*, 30, 393-425.
- Krummheuer, G. (1995). The ethnography of argumentation. In P. Cobb & H. Bauersfeld (Eds.), *The emergence of mathematical meaning: Interaction in classroom cultures*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Lampert, M. & Cobb, P. (2003). Communication and language. In J. Kilpatrick, G. Martin, & D. Schifter (eds.) *A research companion to principles and standards for school mathematics*. NCTM: Reston, VA.
- Lampert, M., Rittenhouse, P., & Crumbaugh, C. (1998). Agreeing to disagree: Developing sociable mathematical discourse. In D.R. Olson & N. Torrance (eds.), *Handbook of education and human development*. Oxford: Blackwell.
- NCTM (2000). *Principles and Standards for School Mathematics*, Reston, VA: NCTM.
- Manouchehri, A. & Enderson, M. (1999). Promoting mathematical discourse: Learning from classroom examples. *Mathematics Teaching in the Middle School*, 4, 216-222.
- McCrone, S.S. (2005). The development of mathematical discussion: An investigation in a fifth grade classroom. *Mathematical Thinking and Learning*, 7(2), 111-133.
- Powell, A. B., Francisco, J. M., & Maher, C. A. (2003). An analytical model for studying the development of mathematical ideas and reasoning using videotape data. *The Journal of Mathematical Behavior*, 22(4), 405-435.
- Powell, A. B., & Maher, C. A. (2002). Inquiry into the interlocution of students engaged with mathematics: Appreciating links between research and practice. In D. S. Mewborn, P. Sztajn, D. Y. White, H. G. Wiegel, R. L. Bryant & K. Nooney (Eds.), *PME-24 (Athens, GA)* (Vol. 1, pp. 317).
- Rasmussen, C. & Stephan, M. (in press). A methodology for documenting collective activity. In A. E. Kelly & R. Lesh (Eds.) *Design research in education*. Mahwah, NJ: Erlbaum.

- Stephan, M., & Rasmussen, C. (2002). Classroom mathematical practices in differential equations. *Journal of Mathematical Behavior*, 21, 459-490.
- Stohl, H., & Tarr, J. E. (2002). Developing notions of inference with probability simulation tools. *Journal of Mathematical Behavior* 21(3), 319-337.
- Tarr, J. E., Lee, H. S., & Rider, R. (2006). When data and chance collide: Drawing inferences from empirical data. In G. Burrill (Ed.), *Thinking and reasoning with data and chance: 2006 yearbook of the NCTM* (pp. 139-150). Reston: VA: National Council of Teachers of Mathematics.
- Toulmin, S. (1969). *The uses of arguments*. Cambridge: Cambridge University Press.