

It's a Home Run!

Using Mathematical Discourse to Support the Learning of Statistics

Kathleen S. Himmelberger and Daniel L. Schwartz



The Standards developed by the National Council of Teachers of Mathematics (2000) state that instructional programs should enable all students to communicate mathematical ideas. The Standards indicate that good communication includes the ability to express organized and precise ideas as well as the ability to analyze and evaluate the mathematical thinking of others. Learning mathematics goes beyond procedural fluency; it also includes learning to discuss mathematical ideas. For this purpose, small groups have become a frequent configuration in the mathematics classroom. When combined with a suitable exercise, small-group discussions can have positive effects on mathematical understanding.

Just putting students in groups, however, does not guarantee productive discussion. Students may not recognize what they should be talking about, may talk about the wrong things, or may not have a driving question to keep their discussions lively and forward moving. Exceptional mathematical exercises and knowledgeable teachers are needed to elicit effective forms of discourse—discourse that naturally moves toward precision, clarity, and generality and gives students the foundation to learn related concepts more easily. Here we present an activity that has worked well for generating productive discussions when teaching statistics.

DISCOURSE THAT PREPARES STUDENTS FOR LEARNING

The Algebra 1, Algebra 1.1 (first half of Algebra 1 in one year's time), and Algebra 1A (Honors Algebra 1) students and teachers at our high school were asked to participate in a university study, "Inventing to Prepare for Future Learning," designed by Schwartz and Martin (2004). Students invented their own ways to measure variability by completing specially designed activities. These "inventing activities" were unique, because the goal was not for students actually to discover the canonical solutions for measuring variability (though that would have been a fine outcome). Instead, the goal was for students to have productive discussions that prepared them to understand in depth the standard solutions. Letting students invent their own solution methods was very helpful to us as teachers; it relieved us of the difficult task of guiding students to invent correct methods without undermining their own production and discussion of mathematical ideas.

The "Pitching Machines" Activity

To clarify how inventing tasks promotes student discussion and learning, we use the example of an activity called "Pitching Machines." (We have included this activity as a worksheet on page 256.) Students are given the grids shown in **figure 1**. Each black circle indicates where a pitch landed when aimed at the X in the center. The students' task was to find a method of computing a reliability index for each machine; the method should produce a score that ranks each machine on a reliability continuum so that people can decide which one they would like to buy. When students asked the meaning of *reliabil-*

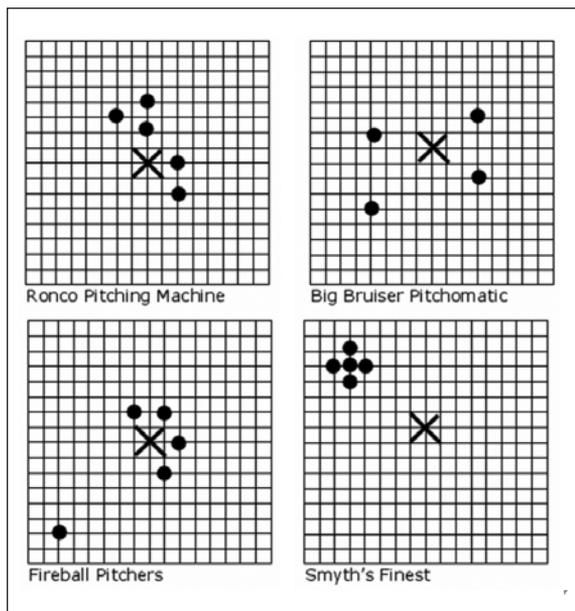
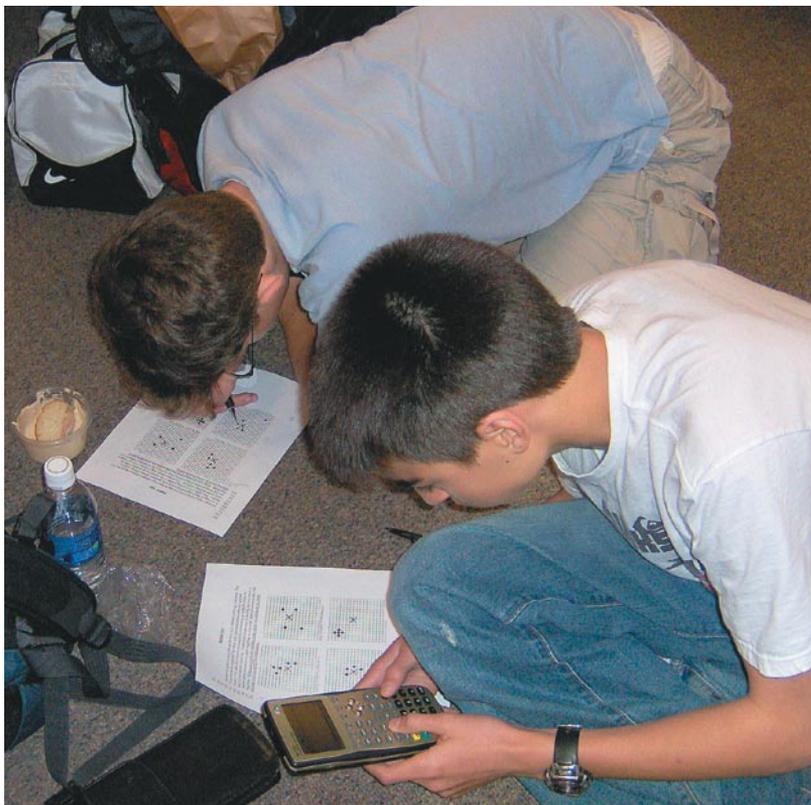


Fig. 1 Grids for the "Pitching Machines" activity



PHOTOGRAPH BY SUZANNE ANTINK; ALL RIGHTS RESERVED

ity, the teacher encouraged them to create a definition based on characteristics that a baseball coach would look for when purchasing a pitching machine.

The grids used an instructional technique called *contrasting cases*. Contrasting cases, like tasting glasses of wine placed side by side, can help people notice aspects of a situation they might otherwise overlook. In the context of statistics instruction, the contrasting cases helped students notice important quantitative properties. For example, the grids held different quantities of pitches, so students noticed that their solution method needed to handle different sample sizes. This observation prepared them to understand why variability formulas often divide by n (to take the average of deviations from the mean). Another example involves the pitching machine that has a tight cluster of pitches far away from the target. When contrasted with the others, this machine helped students notice that variability is not the same thing as inaccuracy, a source of early confusion common among students.

The teachers' responsibility was to circulate among the groups, clarify the task, and ask questions. Some teachers found it uncomfortable to resist giving answers, but this awkwardness wore off quickly when they saw the results of their restraint. There is more than one way to measure reliability, a fact that provided students increased latitude in generating feasible solutions.

The students worked in small groups (numbering two to five) for approximately forty-five minutes to

invent their reliability index. **Figure 2** shows some of the creative ways in which they tried to solve the problem. Group A found the area covered by the pitches; this formula is equivalent to a range formula because only the far points affect the answer. Group B found the perimeter using the grid marks and then used the Pythagorean theorem to compute the length of each line segment; this approach is similar to summing the distances between consecutive numbers in a data list, except that it often ignores interior points. Group C found the average distances between pairs of points, although pairings are selected haphazardly; this method uses the average instead of summing. Group D found the average distance from a haphazardly chosen point to all other points; if students had chosen the distance from the mean of the points instead of an arbitrary starting point, this would be equivalent to the mean deviation. Group E found the frequency of balls in each of the four quadrants; the use of a frequency-based solution is uncommon among students. Group F found the average distance between all pairwise points using a ruler, a general but tedious solution.

These ninth-grade students had studied measures of central tendency and graphical representations extensively in middle school but had little background in other statistical concepts. We were amazed by the level of sophistication shown in their discourse—not necessarily the words they used but the ideas and problems they considered. Each group across all the class levels was able to derive some measure of variability. The solutions themselves were not critical; only one group came up with a generalizable solution. Instead, it was the discussion among group members concerning how to handle aspects of variability that was important. As students devised a method for measuring reliability, they noted most of the features highlighted by the contrasting cases and worked to find methods that could accommodate those features.

The following transcript captures a representative discussion within one group of students. Students often pursued their individual ideas, but they also vigorously commented on one another's ideas and the overall state of their problem-solving process. The visual nature of the problem made it easy for students to point to particular features and be sure that they were all looking at the same thing (although this aspect makes it harder for the reader to follow along in the transcript).

Early on, students noticed that the grids held different numbers of data points (pitches). They debated the value of a solution method that includes all the pitches versus the value of a method that uses a rangelike formula.

Juan. See, you start at one point and count to others from it.

Michael. Yeah, but the other thing . . . Do you realize that over here there are only 4 balls that they tested?

Juan. Sure.

Michael. Over here there are 5.

Juan. Yeah, but it doesn't matter. You only have to use the number of balls in the calculation, or some[thing] that follows the [inaudible].

Michael. So?

Juan. The problem is, for example, . . . here, if you start counting from here, you'll get a very different answer than if you start counting from here.

Michael. Exactly.

Lorenzo. Yeah.

Juan. So I would find something that includes all of them. Like distance from the target.

Lorenzo. Yeah, like [inaudible].

Juan. Shortest distance from the target over longest distance from target is something I'd consider . . . sorry, longest over shortest.

One advantage of student groups, when they work well, is that they generate more ideas to work

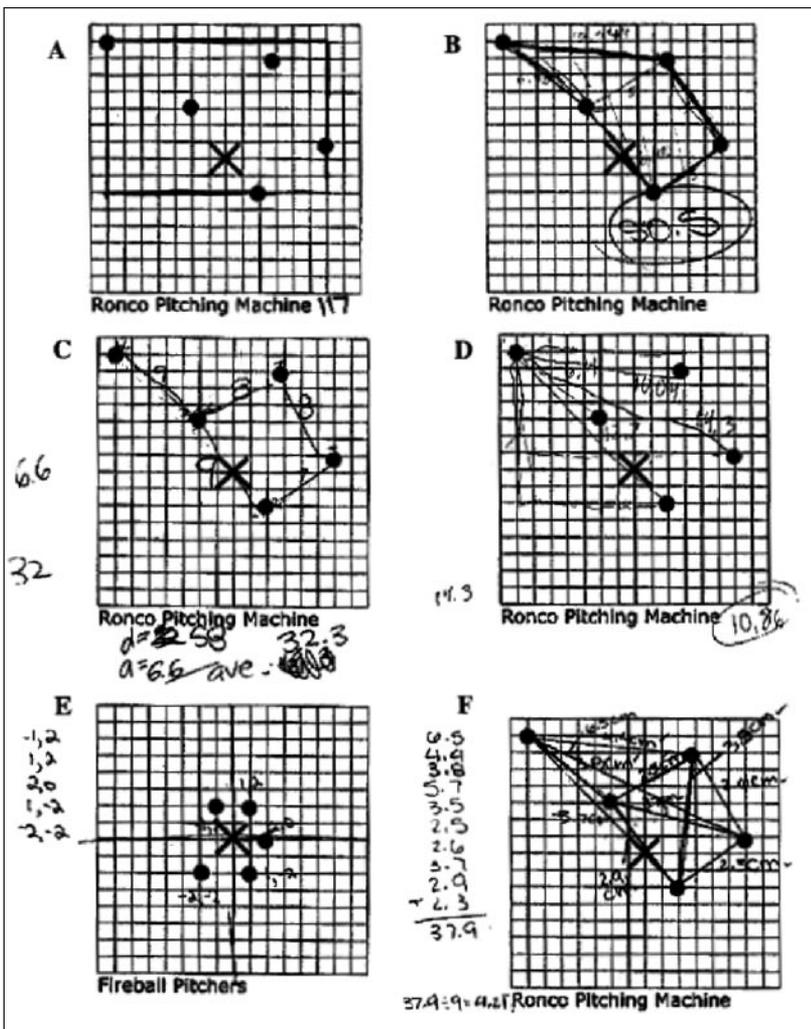


Fig. 2 A few sample student solutions

with than do students working alone. The fractured nature of the conversation is a natural feature of the early stages of moving through many different ideas. The following exchange provides an example of how students moved productively among ideas. Lorenzo, who had been relatively quiet, found an opportunity to contribute by noticing that one pitching grid had a tight cluster of pitches and another grid had an outlier. This observation led to an important discussion in which the students worked with one another's ideas. Ultimately, Michael articulated a set of important issues for Lorenzo's proposed solution method and for subsequent methods they might have generated.

Lorenzo. Right here [pointing to the Smyth's Finest grid], they're all grouped together.

Juan. Yeah.

Lorenzo. But this outlier . . . so we just . . .

Michael. The closest is 2.

Lorenzo. I know.

Juan. The largest is . . .

Lorenzo. I know, but you have . . .

Juan. The longest over the shortest distance.

Lorenzo. Yeah, or we could just eliminate . . . just eliminate that one.

Juan. That will give you the most reliability . . . The problem is . . . Then you'll say this [Big Bruiser Pitchomatic] is very reliable because the distances [shortest and longest] are the same. I was trying . . .

Michael. Although this one [Ronco Pitching Machine] would be very reliable because all of them are closer to the target. Like, for this one [Smyth's Finest], we can always move the target this way, so that you know every single ball . . .

At this point, the teacher arrived at this group's table and helped the students understand the task. In particular, the teacher emphasized that the solution method had to yield values that ranked the machines in accordance with the students' own intuitions of reliability. This constraint prevented the students from generating an arbitrary solution method. Notice that the teacher did not answer the students' questions about how to solve the problem; the students had to assume responsibility for their solution.

Teacher. What is your conclusion? Which one is the most reliable?

Michael. Smyth's Finest.

Teacher. Which one is the least?

Juan. Big Bruiser Pitchomatic.

Lorenzo. Yeah.

Michael. Ronco.

Juan. Big Bruiser Pitchomatic. This one is. This one!?

Michael. That one is less . . . less reliable, and this one is most.



PHOTOGRAPH BY SUZANNE ANTINK; ALL RIGHTS RESERVED

Teacher. And what about these two?

Juan. Oh, you want us to rank them.

Michael. These two are in the middle.

Teacher. But your rule should reflect your ranking.

Juan. Well, sure, we have to now come up with a rule that affects our predefined bias.

Teacher. So, if you say this is the most reliable, and your rule only comes up with this . . . is the highest number somewhere in the middle then?

Juan. The problem is now that what we have to do—now that we are mathematically [biased]—we have to justify it.

Teacher. That's right.

Juan. So how should we go about doing this?

Teacher. That's an interesting question.

Juan. So what you're going to say now is, "Figure it out for yourself."

Teacher. That's right. You've got it.

Juan. I've figured how this class works already.

At our high school we allocate block periods (ninety minutes) twice a week; therefore, we were able to fit the group discussion and whole-class presentations of the solutions into a single block. After all the groups had devised their methods, each group recreated on an overhead projector its solution for one of the grids. It was important that each solution be transparent and precise, because the method would be explained to the class by a student volunteer from a different group using only what was written on the overhead. This requirement elicited even more valuable discussion among group members about possible misinterpretations and ways to clarify and simplify the graphics. During the presentations, students continued to question one another about the rationale and generality of their methods. The teacher remained an interested bystander, helping only to clarify what students said, if necessary.

The “Trampoline” Activity

During the regular fifty-minute mathematics period the following day, the teacher gave a second, briefer task to make sure that the students could connect what they had learned from the original visual problem to numerical data presentations. In this task, the student groups received pairs of small data sets, one pair at a time; for example, they first received {1 3 5 7 9} versus {3 4 5 6 7}. They were told that each number in a set represents how high a ball bounced when it was dropped on a particular trampoline. Their task was to invent a way to compute which of the two trampolines was more consistent. They repeated the process with the pairs {1 1 1 1 9} versus {1 3 5 7 9} and {1 3 5} versus {1 1 3 3 5 5}.

Then the teacher gave a brief (five- to ten-minute) lecture about the mean deviation, one solution method that mathematicians invented to solve these kinds of problems. The teacher applied this method to the trampoline data to show how it worked. Students were interested to see how their solution method fared against that of the mathematicians. Following the lecture, they spent ten to fifteen minutes practicing with a new small data set. Our hope was that the students’ use of discourse and their experiences with the “Pitching Machines” and “Trampoline” activities better prepared them to understand variability and learn how the mean deviation formula does such a good job of handling it.

EVIDENCE THAT STUDENTS LEARNED

A week after learning about variability and standardized scores (e.g., grading on a curve), ninety-five students took a written test. One question asked students to compute a measure of variability. Eighty-six percent accurately computed the mean deviation. A year later, we retested a random subset of thirty students; 57 percent remembered how to compute the mean deviation, even though they had practiced for only ten minutes the year before and had had no intervening practice. We compared these students with a random sample of college students from a public university ranked in the top twenty nationwide (according to *U.S. News and World Report*) who had taken a semester of college statistics within the past two years. None of the college students remembered how to compute a measure of variance.

We also included an item in the posttest to see if the students understood the rationale behind the structure of the mean deviation formula, which we thought the invention process had prepared them to learn from the brief lecture. We asked them why the formula

$$\frac{\sum |x - \bar{x}|}{n}$$

has an n in the denominator. Sixty-four percent of the students indicated that the division addressed the



PHOTOGRAPH BY SUZANNE ANTINK; ALL RIGHTS RESERVED

problem of different sample sizes. For comparison, we also asked them why the slope formula, which they had learned a few weeks earlier without a discourse-driven curriculum, subtracts x_1 from x_2 in

$$m = \frac{y_2 - y_1}{x_2 - x_1}.$$

For this problem, only 32 percent of the students could explain why, although they could use the formula. Discourse-based activities, when done well, can increase students’ conceptual understanding of statistics and formulas.

KEY COMPONENTS OF ELICITING MEANINGFUL MATHEMATICAL DISCOURSE

In our experience, the key components behind this successful classroom experience included an engaging activity that helped students focus on the important things they should discuss, students working in groups of two or more, and a teacher comfortable with and proficient at facilitating classroom discourse. We discuss each component in turn so that teachers may be able to find or design their own lessons.

The activity must be approachable. The students should be able to try things out quickly and easily revise their initial efforts as they begin to notice more aspects of the problem. For example, the visual format of the “Pitching Machines” activity plus its comprehensible context and contrasting cases allowed students to start sketching answers quickly, before they started to compute specific answers. Problems having several solution paths are also desirable because they make it possible for many students in the group to bring different yet legitimate ideas to the table and keep the conversation going. Problem-solving books abound; the major task for the teacher is to identify relevant, enriching problems that can lead to good discussion and a better understanding of the curriculum. We hope our example will help other teachers identify likely candidates. (Additional examples of statistics



activities may be found in the appendix of Schwartz and Martin [2004] at AAALab.Stanford.Edu.)

Students working in groups have the opportunity to engage in exciting, sometimes heated discussions about mathematics. Ultimately, the goal is for individual students to adopt the group habits of questioning, checking, and explaining, which are critical for meaningful learning and mathematical communication. Schoenfeld (1995) noted in his studies of mathematical problem solving that students working alone tended to search for familiar, formal mathematics they could apply to novel problems quickly, with little regard for the accuracy and feasibility of their response. In the “Pitching Machines” task, we tried to solve this problem by having students work on a novel inventing task in groups, where their shared ignorance of how to begin, instead of causing a rushed solution, freed them to analyze all aspects of the problem.

In addition to the valuable collaboration among students, an added bonus of student discourse is that it reveals student thinking for the teacher, who can take note of misconceptions or ideas to be discussed later. The teacher must not jump in immediately to point out either erroneous or correct solutions but be willing to let the groups hash out their ideas. A simple question about their process or how it applies to a different data set can help students move forward. At the same time, teachers need to be able to adapt quickly to unorthodox or creative solutions they might encounter as they circulate among groups. For example, for the “Pitching Machines” activity, many students found the area (fig. 2). Several teachers questioned whether the area solution would work if all the dots were in a straight line. By generating the new case on the fly, the teachers kept the student discussion alive. At the same time, the teachers modeled an important aspect of mathematical thinking, namely, generating cases that test the generality of a solution method. In summary, for these types of activities, the teacher assumes the role of a baseball team manager, who must decide whether to let the ball players keep playing, interrupt to give the

pitcher advice or support, or even set up a play that the players can try out.

TRADE-OFFS

Having students engage in invention activities and making group presentations does require more class time than do more traditional tell-and-practice lessons and may not elicit correct responses during the invention session. An additional class period may be required to present the correct solution. Most mathematics teachers allow one day per topic or section presented in the text, plus a review day before the test. The invention process would add a day of instruction. However, our experience shows that students who have a chance to discuss their own inventions are more prepared to learn and retain subsequent related material, thus potentially allowing them to save time in the future. Schwartz and Martin (2004), for example, also conducted a formal study that compared a more traditional tell-and-practice style of instruction with the discourse-over-inventing instruction. They found that the two styles of instruction looked the same when students were tested directly on the computational aspects of the lessons. The big difference showed up later. The students trained in discourse learned more than twice as much from later related lessons.

We recommend the periodic use of invention as good topics and problems present themselves, especially at the beginning of a unit or main idea in mathematics. If a student hits the ball correctly, it will go out of the park.

REFERENCES

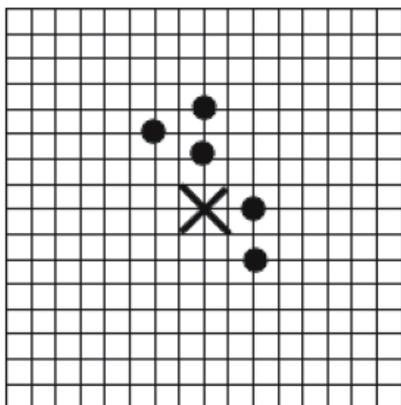
- National Council of Teachers of Mathematics (NCTM). *Principles and Standards for School Mathematics*. Reston, VA: NCTM, 2000.
- Schoenfeld, A. H. *Mathematical Problem Solving*. New York: Harcourt Brace Jovanovich, 1995.
- Schwartz, D. L., and T. Martin. “Inventing to Prepare for Future Learning: The Hidden Efficiency of Encouraging Original Student Production in Statistics Instruction.” *Cognition and Instruction* 22, no. 2 (2004): 129–84. Appendix available at AAALab.Stanford.Edu.∞

KATHY HIMMELBERGER, khimmelberger@pausd.org, teaches mathematics at Palo Alto High School, Palo Alto, CA 94301. She is interested in increasing the mathematical confidence of the average mathematics student. DANIEL SCHWARTZ, danls@stanford.edu, is a professor of education at Stanford University, Stanford, CA 94305. He is interested in making classroom learning satisfying and effective and preparing students to continue learning once they leave the classroom.

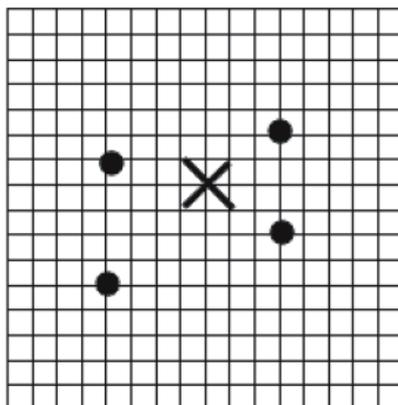
Batter Up!

Name _____

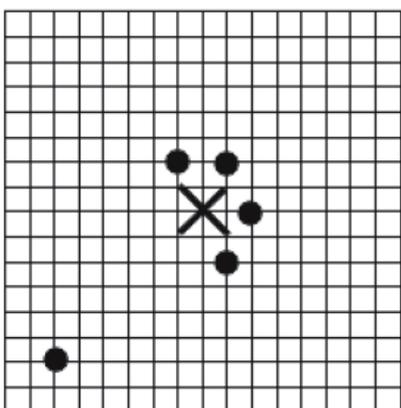
The four grids below show pitching results from four different pitching machines. The X represents the target, and the black dots represent where different pitches landed. Your task is to invent a procedure for computing a reliability index for each pitching machine. There is no single way to do this, but, to ensure a fair comparison among the machines, you must use the same procedure for each. Using the grids below, compute a reliability index for each pitching machine. Write your procedure and the index value you compute for each.



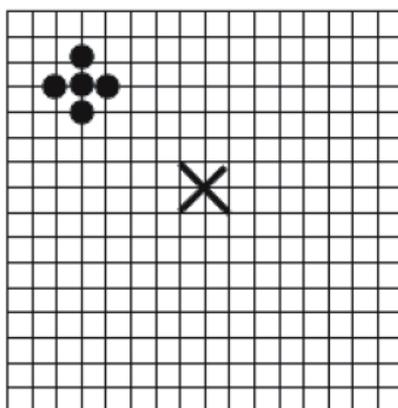
Ronco Pitching Machine



Big Bruiser Pitchomatic



Fireball Pitchers



Smyth's Finest