# It's a Home Run!

## Using Mathematical Discourse to Support the Learning of Statistics

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, and 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, and 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 and they can talk about the wrong things; or they may not have a driving question to keep their discussions lively and moving forward. Exceptional mathematical exercises and knowledgeable teachers are needed to elicit effective forms of discourse – discourse that naturally moves towards precision, clarity, generality, and gives students the foundation to more easily learn future, related concepts. Here, we present an example of 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 the students to actually discover the canonical solutions for measuring variability (though that would have been a fine outcome). Instead, the goal was for the students to have productive discussions that prepared them to deeply understand the standard solutions, which were eventually presented by the teachers. 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.

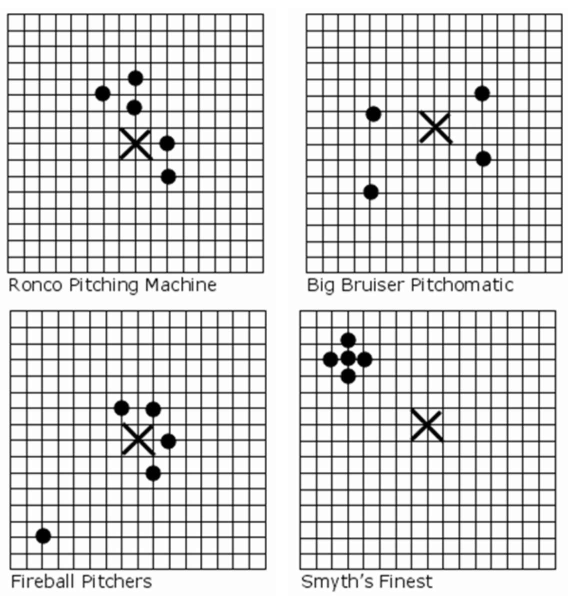



Figure 1. Pitching Machine Task

To clarify how the inventing tasks promote student discussion and learning, we use the example of an activity called “Pitching Machines.” (We have included the Pitching Machine activity in the back as a worksheet.) Students received 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 to compute a *reliability index* for each machine; the method should produce a score that ranks each machine on a reliability continuum so people can decide which one they would like to buy. When students asked what *reliability* means, the teacher encouraged them to create a definition based upon characteristics a baseball coach would look for in purchasing a pitching machine.

The grids used an instructional technique called *contrasting cases*. Contrasting cases, like tasting glasses of wine 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 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 with a tight cluster of pitches that are far away from the target. When contrasted with the other machines, this case helped students notice that variability is not the same thing as inaccuracy, which is a common early confusion for students.

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

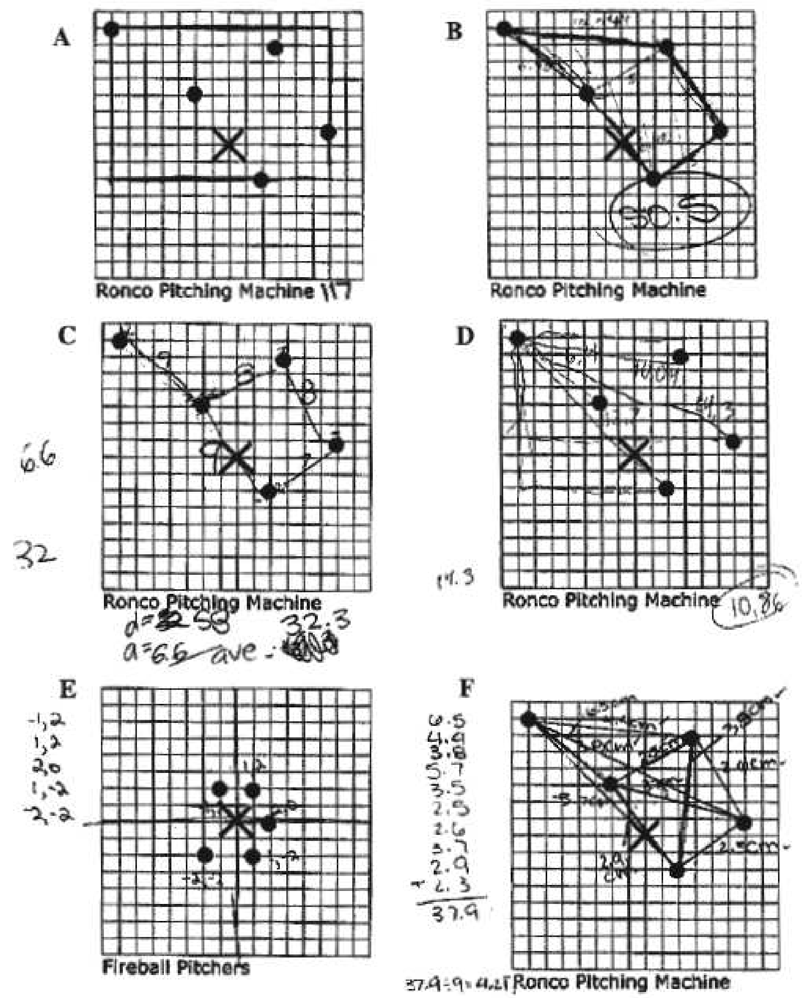



Figure 2. A few sample student solutions.

The students worked in small groups (2 to 5 students) for approximately 45 minutes to invent their reliability index. Figure 2 shows some of the creative ways they tried to solve the problem: (A) Found the area covered by the pitches; this is equivalent to a range formula because only the far points affect the answer, (B) Found the perimeter using the grid marks and used the Pythagorean Theorem to compute the length of each line segment; this is similar to summing the distances between consecutive numbers in a data list, except it often ignores interior points, (C) Found the average distances between pairs of points, although pairings are selected haphazardly; this method uses the average instead of summing, (D) Found the average distance from a haphazardly chosen point to all other points; if they had chosen the distance from the mean of the points instead of an arbitrary starting point, this would be equivalent to the mean deviation, (E) Found the frequency of balls in each of the four quadrants; this is a rare frequency-based solution, (F) Found the average distance between all pair-wise points using a ruler: this is 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 with other statistical concepts. We were amazed by the level of sophistication shown in their discourse, not necessarily by the words they used, but by 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 the 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 among one group of students. The students often pursued their individual ideas, but they also vigorously commented on each other's ideas and the overall state of their problem solving process. The visual nature of the problem also made it easy for the students to point to particular features and to be sure that they were all looking at the same thing (although this makes it harder for the reader to follow along in the transcript).

Early on, the 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 a method that uses a range-like 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 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 of the advantages of groups, when they work well, is that there will be more ideas to work with compared to doing a task alone. Although the conversation may seem fractured, this is a natural feature of the early stages of moving through many different ideas. The following exchange provides an example of how students productively moved 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 led to an important discussion where 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 [Smythe] 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…
JUAN: … the problem is… Then you'll say this [Big Bruiser] is very reliable because the distances [shortest and longest] are the same. I was trying…
MICHAEL: Although this one [Ronco] would be very reliable because all of them are closer to the target. Like for this one [Smyth], we can always move the target this way, so that you know every single ball….

At this point, the teacher arrived at their table and helped the students understand the task. In particular, the teacher emphasized that their solution method had to yield values that ranked the machines in accordance with their 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 do it; 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.
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 pre-defined 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 bigoted --- 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 got it.
JUAN: I figured how this class works already.

We use block periods (90 minutes) twice a week at our high school; therefore, we were able to fit group discussion and whole class presentations of the solutions into a single block. After all groups had devised their methods, each group recreated their solution for one of the grids on an overhead. It was important that their solutions be transparent and precise, because a student volunteer from a different group was selected to explain the method to the class 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 concerning the rationale and generality of their methods. The teacher remained an interested bystander, only helping to clarify what students said, if necessary.

During the regular 50-minute period the following day, the teacher gave a second, but briefer task, to make sure the students could connect what they 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} v. {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} v. {1 3 5 7 9} and {1 3 5} v. {1 1 3 3 5 5}.

Then the teacher gave a brief 5-10 minute lecture about the mean deviation, one solution method that mathematicians invented to solve these kinds of problems. The teacher applied it 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, the students spent 10-15 minutes practicing with a new small data set. Our hope was that the students' use of discourse and their experiences with the Pitching Machine and trampoline activity better prepared them to understand variability and to 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), the students took a written test. One question asked the students to compute a measure of variability. Eighty-six percent accurately computed the mean deviation. A year later we retested a random subset of 30 students; fifty-seven percent remembered how to compute the mean deviation, even though they had only practiced for 10 minutes the year before with no intervening practice. We compared these students with a random sample of college students from a top-20 public university (according to *US 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 $\frac{\sum|x-\overline{X}|}{n}$ divides by *n*. Sixty-four percent of students indicated that it addressed the 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% 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 for 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, groups of two or more students, and a teacher comfortable with and proficient at facilitating classroom discourse. We discuss each component in turn so teachers may be able to find or design their own lessons.

The activity must be approachable for the students. The students should be able to try out things 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 machine problem, plus its comprehensible context and contrasting cases, allowed the students to start sketching answers quickly, before they started to compute specific answers. Problems with several solution paths are also desirable because they make it possible for many students in the group to bring different 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 http://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. Alan Schoenfeld (1995) noted in his studies of mathematical problem solving that students working alone tended to quickly search for familiar, formal mathematics they could apply to novel problems, with little regard to the accuracy and feasibility of their response. In the pitching machine task, we tried to solve this problem by having students work in groups on a novel inventing task where, instead of causing a rushed solution, their shared *ignorance* of how to begin 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 be willing to let the groups hash out their ideas by not jumping in immediately to point out either erroneous **or** correct solutions. 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 the groups. For example, with the pitching machine task, many students used an area solution (Figure 2A). Several teachers who saw this solution asked 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 sum, for these types of activities, the teacher assumes the role of manager who must decide whether to let the ball players just 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 extra class time compared to more traditional *tell-and-practice* lessons, and they may not elicit correct responses during the invention session. An additional class period may be required to present the correct solution. Most math teachers allow one day per topic or section in the book, 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, which can potentially 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 the students were tested directly on the computational aspects of the lessons. The big difference showed up later. The students in the discourse condition learned over twice as much from later, related lessons.

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

## Works Cited

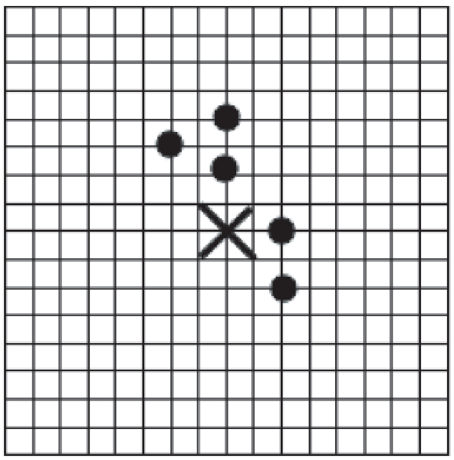
National Council of Teachers of Mathematics. *Principals and Standards for School Mathematics*, 2000.

Schoenfeld, A. H. *Mathematical Problem Solving*. New York: Harcourt Brace Jovanovich, 1995.

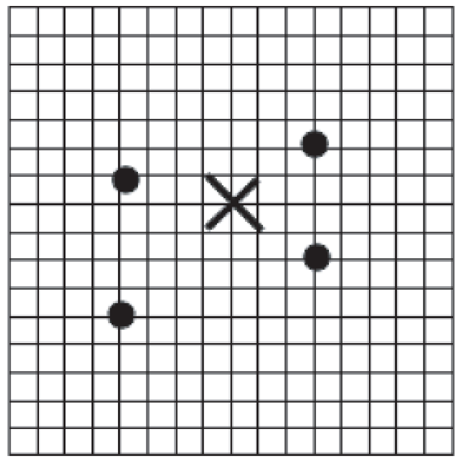
Schwartz, D. L., & Martin, T. "Inventing to Prepare for Future Learning: The Hidden Efficiency of Encouraging Original Student Production in Statistics Instruction." *Cognition and Instruction*, 2004, 22(2), 129-184.

## Batter Up!

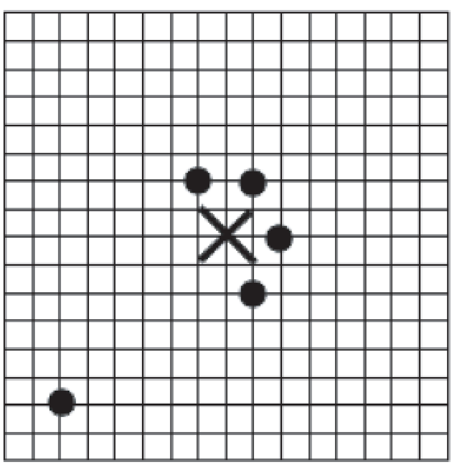
Here are four grids showing the results from four different pitching machines. The X represents the target and the black dots represent where different pitches landed. Your task is to invent a procedure for computing a reliability index for each of the pitching machines. There is no single way to do this, but you have to use the same procedure for each machine, so it is a fair comparison between the machines. <u>Write your procedure and the index value you compute for each pitching machine using the grids below.</u>



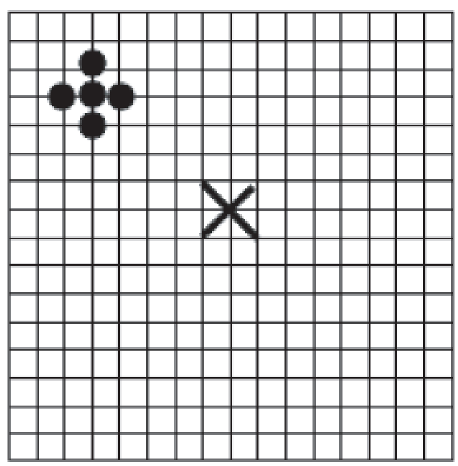
Ronco Pitching Machine



Big Bruiser Pitchomatic



Fireball Pitchers



Smyth's Finest