Cortina, J., Saldanha, L., & Thompson, P. W. (in press). Multiplicative conceptions of arithmetic mean. In F. Hitt (Ed.). *Proceedings of the Annual Meeting of the International Group for the Psychology of Mathematics Education*. Cuernavaca, Mexico: Centro de Investigación y de Estudios Avanzados.

MULTIPLICATIVE CONCEPTIONS OF ARITHMETIC MEAN

José Luis Cortina
Vanderbilt University
jose.luis.cortina@vanderbilt.edu

Luis Saldanha
Vanderbilt University
luis.a.saldanha@vanderbilt.edu

Pat Thompson
Vanderbilt University
pat.thompson@vanderbilt .edu

We elaborate a multiplicative conception of the arithmetic mean that is grounded in quantitative reasoning. This elaboration serves as a framework for the design and analysis of teaching experiments intended to support students' building this conception of the mean. We discuss insights into students' conceptions and their instructional implications.

Background

Mokros and Russell (1995) distinguish two basic models that have been used in the literature when defining the mean: *fair share* and *balance*. These authors note that works that have used these models seem to omit the notion of *representativeness* as an important characteristic of the mean. Mokros and Russell (1995) view the mean as an indicator of the center of a distribution, one that is useful for summarizing, describing, and comparing different data sets. For them, a powerful understanding of mean is that of a "mathematical point of balance." From this perspective, the different contributions of a data set are balanced with each other, and their point of equilibrium becomes a distinctive characteristic of the collection. Other researchers (Pollatsek, Lima, & Well, 1981; Strauss & Bichler, 1988) have characterized the mean of a set of scores as "typical" or "representative" of individual scores.

Perspective: Mean as a multiplicative measure

Thompson (1998a) claims that students are misled when the mean of a set of scores is depicted as somehow being "typical" or "representative" of either the individual scores or the set of scores. From an instructional point of view characterizing the mean as typical of individual scores is vague and may direct students to think of the mean as indicating something about individual scores. Characterizing the mean as a summary of a set of scores seems sensible only when it includes a discussion of variance as a summary of the variability of scores relative to a reference value — the mean. Moreover, we cannot infer anything about how a set of scores is distributed even when we

The research reported here was supported by the National Science Foundation under grant No. REC9814898 and grant No. The opinions expressed here do not necessarily reflect the views of the Foundation.

know its mean *and* variance. It could be distributed in any number of ways, many of which have no resemblance to each other.

Thompson (1998a) conceptualizes the mean as a multiplicative notion. His approach grounds instruction in quantitative reasoning and is intended to support conceiving of statistical measures quantitatively. In his theory of quantitative reasoning (Thompson, 1994) the notion *quantity* is a conceptual entity that a person constructs when conceptualizing situations: one thinks of a quantity when conceiving an attribute of an object as measurable. A quantity is schematic, it involves an object-image, a conceptualized attribute of the object, a tacit understanding of appropriate units of measure, and a quantification process — a process by which one directly or indirectly assigns numerical values to the attribute. Thompson's conception of the mean is rooted in images of quantifying an attribute. He stresses the importance of having students first conceive of group performance as an attribute of a group that they imagine themselves measuring. Then students need to create the mean as an adjustment on the measure of group performance, an adjustment that will allow them to compare groups of unequal sizes on an attribute that is sensitive to group size. In order for this to seem sensible to students, they need to have an underlying sense of "individual contributions to the group attribute". That is, one must imagine that in a group each individual performs and that each performance contributes to the group performance. As one "runs through" the members, one accumulates a total, aggregate measure. These images and conceptions support an understanding of the mean as a ratio that measures *group performance relative to the number of contributors in the group*. If students are mindful that they are measuring a group performance when they add up scores, then "divide by the number of contributors" can be introduced as a natural normalizing adjustment to their measurement procedure to take different group sizes into account. In this sense the mean is like an average rate (Thompson, 1994); the measure of the group contributions *per* contributor is conceived to be the same as the amount contributed by each of *n* contributors if each were to contribute equal amounts. The quantification process points to an overlap between mean as multiplicative measure of group performance and the "fair share" notion mentioned by Mokros and Russell (1995). But, they are different in that the core conception of the former is measure of group performance and relative contribution per member, whereas measurement is not an issue in “fair share.”

The aspect of these conceptions that makes the mean a multiplicative idea is that in comparing measures of group attributes one must understand that aggregate measure and group size are simultaneously accounted for in any comparison (Inhelder & Piaget, 1969; Thompson, 1994; Thompson, 1998b).

Purpose and method

The aim of our study is to investigate middle-school students’ conceptions as they engaged in problem tasks intended to support their understanding the arithmetic mean multiplicatively. In this paper we document and discuss the reasoning of four students — Marissa, Nick, Joy, and Kurt — as they confronted situations requiring them to compare the performance of unequal-sized groups.

Subjects were part of a group of 7th and 8th-grade students at a middle school in Nashville, Tennessee who participated in clinical interviews and teaching experiments. These experiments typically involved one to three 50-minute sessions in which students worked in pairs on one or two problems. Two researchers were present during the sessions; their role was to encourage the students’ joint participation in reasoning about the problems, and to encourage them to publicly explain their reasoning. Each session was videotaped and field notes were taken. Data were analyzed from the perspective of mean as a multiplicative measure, as elaborated above.

Results

In the individual interviews, Nick and Marissa (8th graders) were presented with two data sets representing the distances ridden by individual members of two cycling teams in a contest. The teams had an unequal number of members (6 and 9). The problem was to give criteria for choosing the winning team. Our intent in using this problem was to raise the issue in students’ minds of the need to account for the difference in group sizes when comparing groups on similar characteristics. We wondered whether students would recognize the usefulness of the mean in this situation. Neither Marissa nor Nick mentioned the mean of scores as a criteria. Both Marissa and Nick seemed to be thinking of the performance of each bicycle team; they added up the scores in each team but excluded three of the scores in the bigger team — truncating the data was their way of accounting for the size difference.

In retrospect, we conjectured that this first task may have facilitated this additive-like normalization. We therefore designed a subsequent task for the teaching experiment in which students would have little opportunity to reason in this way. The task was also intended to focus students' attention on group performance, consequently the text of the problem contained no information about individual contributors. Additionally, we wanted to set a problem that would not evoke an automatic response to use or calculate the mean. In this task (see figure 1) students were asked to use sample data to help them devise a criteria by which to select a winning troop of Girl Scouts nation-wide.

"The National Office of the Girl Scouts of America is presenting an award to one troop of Girl Scouts for participating in the annual cookie sale. It is your job to advise the National Office. How should they select the wining troop?"

Here is an example of the kind of information that is available to the National Office:

Town	Population	Number of troop members	Amount of money earned	Number of boxes sold
Amesbury	15,00	20	480	120
Westport	37,000	200	3,060	1,020
Kendal	25,000	100	2,275	700
Ellesmere	72,000	300	3,250	650
Union City	1,000,000	500	4,500	1,000

Figure 1. The "Girl Scout Cookie" problem statement and data used by students

In this problem both Marissa and Nick sensed that considering only the number of boxes sold or only the amount of money earned would not be a fair way of determining a winner. Nick's strategy to solve this problem was to divide the number of boxes by the number of troop members. He viewed this as a way to express the number of boxes sold as a multiple of the size of the troop, and he apparently took this multiplicative comparison as an indicator of the group's performance.

Researcher: So can you explain your method again?

Nick: Yes, it's really just like dividing, I guess, cause am I was just looking at it like, Amesbury (inaudible) sold, I am trying to say this right, six times the number of troop members.

While Nick's strategy can be considered identical to calculating the mean, he seemed to interpret the product of the division as a multiplicative comparison of magnitudes. The idea of "contribution per contributor" did not seem to be part of his reasoning. When trying to help Nick explain his method, Marissa did seem to have the latter interpretation:

Researcher: What does that tell you? What are those numbers telling you?

Marissa: How many boxes each person would've had to have sold

Researcher: What do you mean each person would've had to have sold?

Marissa: About, how, from the number of members in the troop, how many boxes each person in the troop would've had to have sold to come up with that number of boxes.

But Marissa did not view this "equal distribution" as a scaled measurement of group performance that allowed her to make comparisons. During the session she made it clear that she did not consider "Nick's way" to be an adequate way of comparing the troops. Instead Marissa proposed to solve the problem by giving two awards, one for troops in towns with populations less than 500,000 and another for those in towns with populations greater than 500,000.

When the 7th-grade students, Joy and Kurt, first engaged in this task neither one saw a need to normalize the groups in order to compare them. These students thought that the winning criterion should simply be "number of boxes sold". Kurt argued that troop size should not be a consideration because the purpose of selling cookies is to collect funds, so the troop with the highest sales should win. After we pushed the students to consider whether this was a fair criterion the discussion led to their considering the comparison of boxes sold and troop size. Kurt recognized this as an average, and both students had an image of average that is consistent with Marissa's: they interpreted it as a hypothetical equal distribution of the total boxes sold, but did not see it as a normalized measure of troop performance. Kurt seemed to think that average was just a different criterion from total boxes sold, one that was unrelated to group performance.

These observations inspired us to design a task intended to have Joy and Kurt confront the problematicity of comparing the performance of groups of unequal sizes in a situation that did not explicitly involve a competition. The problem asked to compare crime statistics in the cities of San

Diego (SD) and Tijuana (TJ) in order to decide whether the smaller town (TJ) had a higher incidence of crime (see figure 2 for an abridged version of the statistics examined by students).

San Diego	
Total crimes	54,421
Population	1,214,000

Tijuana	
Total crimes	45,868
Population	813,380

Figure 2. Crime statistics for the cities of San Diego and Tijuana

At first, students focused only on the total incidents of crime. Only after being asked whether TJ could be considered a "safer" place than SD did their conceptions entail a consideration of population sizes. They seemed to have a strong sense that the difference in total crime between the cities was not proportional to their population sizes, and they took this as indicative of a difference in *safeness* of the two cities. Joy's comment illustrates their thinking: "I'd probably say that SD, it, to me it, SD looks safer to me 'cause it has so many people. And then TJ just has 800,000 people with 45,000 crimes".

Students' idea of how to deal with this problem seemed to be based on hypothetically equalizing their populations and proportionally increasing or decreasing the number of crimes.

> Kurt: Uhh, yeah if you could just like find a way to show like population evenly between the two. Like just say give them each a million people but took the crimes that are still there and, like some way figure what it would be with a million people by either adding or subtracting.
>
> Researcher: In each city, each having a million people?
>
> Kurt: Yeah , in each city having a million people, and what the crimes would be using the statistics you have now. And so each city would be even as far as population.

Conclusion

Our students demonstrated relatively sophisticated multiplicative reasoning and their understanding of "average" was more than purely procedural, it seemed consistent with the "fair share" conception mentioned by Mokros and Russell (1995). In spite of this, conceiving of a normalized measure of group performance was non-trivial for them. With the exception of Nick, even when students formed the ratio of group total to group size, for them this did not constitute a *measure* of group performance relative to group size. It was unclear to them what this ratio was an indicator of, and their strategies for dealing with the difference in group sizes suggests that they did not see that this ratio inherently accounts for the difference. Our challenge in these experiments was to design instruction that would support students thinking about *group* performance and their envisioning the need to consider group size when comparing group performances. The results of these experiments point us in a clear direction for the future: to focus on first having students come to conceive of group performance as a *measurable* attribute and then to support their conceiving of the mean as an appropriate *measure* of that attribute.

References

Inhelder, B. & Piaget, J. (1969). *The early growth of logic in the child: Classification and seriation.* New York: Norton.

Mevarech, Z. (1983). A deep structure model of students' statistical misconceptions. *Educational Studies in Mathematics,* 14, 415-429.

Mokros, J. & Russell, S. (1995). Children's concepts of average and representativeness. *Journal for Research in Mathematics Education*, 26(1), 20-39.

Pollatsek, A., Lima, S., & Well, A. D. (1981). Concept or computation: Students' understanding of the mean. *Educational Studies in Mathematics*, 12, 191-204.

Strauss, S., & Bichler, E. (1988). The development of children's concepts of the arithmetic average. *Journal for Research in Mathematics Education*, 18(1), 64-80.

Thompson, P.W. (1994). The development of the concept of speed and its relationship to concepts of speed. In G. Harel & G. Confrey (Eds.) *The development of multiplicative reasoning in the learning of mathematics,* (pp. 179-234). NY: SUNY Press.

Thompson, P.W. (1998a) Manuscript. Measures of group characteristics. Vanderbilt University.

Thompson, P.W. (1998b) Manuscript. Foundations of multiplicative reasoning. Vanderbilt University.