Young Children's Explorations of Average in an Inquiry Classroom

Katie Makar
The University of Queensland
<k.makar@uq.edu.au>

Debra McPhee
Jindalee State School
<dmcph8@eq.edu.au>

This study situates early notions of average within an inquiry classroom to investigate the rich statistical concepts young children draw on to make sense of the questions: *Is there a typical height for a student in Year 3? If so, what is it?* Based on their deliberations over several lessons, students' ideas about average evolved through four conceptions - typical as: *reasonable*, *most common* (value or interval), *not atypical*, and *representative of the population*. Implications for teaching are discussed.

The concept of average is a key idea in statistics. Research has suggested, however, that while most older students are able to calculate an arithmetic mean, nearly all struggle to make appropriate use of an average in meaningful situations (Konold & Pollatsek, 2002; Watson, 2006;). Curricular statements in Australia and abroad have encouraged teachers to build students' early conceptual experiences with data handling in authentic contexts by, for example, asking students to collect and analyse data based on their own questions of interest (QSA, 2007; Curriculum Corporation, 2006; NCTM, 2000). The hope is that by strengthening connections between statistical measures and concrete experiences, children will already have a deep conception of data when formal ideas are introduced.

Even young children have intuitive ideas of average through their everyday experiences. By giving them opportunities to participate in statistical ideas embedded in contextualised situations, they can develop rich informal understandings (Watson, 2006). The study here reports on a Year 3 (age 8) class as they wrestle with the question, *Is there a typical height for a person in Year 3?* By considering students' conceptual development of 'typical' in a classroom culture of inquiry, the report here aims to provide insights into ways of deepening student understanding of complex statistical ideas from an early age.

Background

According to Watson (2006), the big ideas for developing statistical literacy include the concepts of variation, sampling, average, chance, and inference. While average may be considered the least difficult of these ideas, work over the past two decades suggests that students continue to struggle with it (Watson, 2007; Konold & Pollatsek, 2002). In their classic study, Mokros and Russell (1995) found that children (aged 9-14 years old) typically held five conceptions of average: mode, algorithm, reasonable, midpoint, and as a mathematical balance point. Of those who relied on the algorithm, none used it effectively; furthermore, for these students the algorithm was given more credence even when reasonableness and self-monitoring indicated a different answer. The children mistrusted reasonableness as a reliable tool for thinking about average. The mode, or single most frequent value, was another common conception of average. Children with this approach often held to it tenaciously even when prompted to think about other meanings. This can be problematic when moving from a focus on single values of data towards conceptualising data as an aggregate. The authors further suggest that children need to be given multiple opportunities to work with informal notions of average within authentic contexts "to develop their own ideas of typicality or representativeness as they describe and compare data sets" (p. 38). They promote the need to embed average in meaningful contexts,

In R. Hunter, B. Bicknell, & T. Burgess (Eds.), *Crossing divides: Proceedings of the 32nd annual conference of the Mathematics Education Research Group of Australasia* (Vol. 1). Palmerston North, NZ: MERGA.

growing out of everyday experiences to elicit students' notions of reasonableness. They contend that by age 9,

> students have developed powerful, situation-based ways of thinking about average. … Students' notions of representativeness or typicality grow out of everyday experiences and have a strong flavor of reasonableness and practicality. … Children's informal ideas about outliers helped them home in on what was typical. Reasonableness in evaluating a data set appeared as an essential strand in understanding, a strand that plays a significant role in the development of more complex notions of average (p. 21).

Their study suggests several key ideas to focus on when supporting student learning about average: (1) Young children's informal notions of average are *situation-based*, grounded in their everyday experiences with *typicality* and *representativeness*; (2) experiences with considering *outliers*, or atypical data, can help children to focus on what is typical; and (3) essential to their understanding of more complex notions of average is the idea of *reasonableness*.

In a recent and more comprehensive study, Watson (2007) interviewed students in Years 4, 6, and 8 (ages 9-13) and asked them to consider average in various contexts to elicit their informal, algorithmic, and conceptual understandings. By showing video-taped responses from other students that differed from their own response, she used cognitive conflict to prompt students to reconsider and then re-articulate their understandings. She found that in doing so, a large proportion of students were able to discuss concepts with more depth and complexity. This study suggests that it is important not just to encourage a single notion of average, but to give students experiences that require them to negotiate and debate multiple concepts of average in a complex context. In an effort to help us re-evaluate the role of imitation in learning, Ben-Zvi and Sfard (2007) propose that by supporting students through a process in which they collectively deliberate ideas, children are able to turn authoritative discourse into discourse that is persuasive for self and others.

Conceptual Framework

The outcomes of the literature on children's conceptions of average reported above were the results of clinical interviews rather than children's learning in the social culture of a classroom. We were interested in observing students' learning as they debated and collectively wrestled with notions of average using ill-structured problems. Through considering aspects of students' negotiations and representations, we were looking for evidence of student learning at a more complex level to support our hypothesis that data studied in an authentic context would produce deeper levels of children's understandings.

Two frameworks provide the conceptual basis for our study. Importantly, we situate our work within an epistemological framework of inquiry-based learning. In this study, we consider inquiry to be a pedagogical process in which a teacher supports her students as they wrestle with ill-structured problems (Makar, 2008). An ill-structured problem is one in which the problem statement or solution pathways have a number of ambiguities which require negotiation in order to mathematise or structure the solution process (Reitman, 1965). Inquiry requires an epistemological foundation that presumes knowledge is socially constructed through collaborative and iterative cycles of investigation and debate.

Drawing on relevant aspects of the literature on students' conceptions of average and inquiry, we see five key ideas underpinning our understanding of young children's explorations of average in the context of understanding typical heights.

- *Reasonableness.* In order to conceptualise an average, students need a sense of what is a 'reasonable' height (Mokros & Russell, 1995). By considering reasonableness, our aim is for students to envision a range of values that are meaningful in this context.
- *Outliers.* In deliberating what is average or typical heights, students may find it useful to make sense of outliers, or values that are 'atypical' (Mokros & Russell, 1995). We assumed that by asking students to consider atypical values, it would make the consideration of typical more apparent.
- *Typical as most common.* Mode is a fairly conventional application of average at the primary level. Researchers (e.g., Mokros & Russell, 1995; Watson, 2006) note that students often consider values with the highest frequency as being 'most popular' or 'typical'. We assumed that 'typical' would be an accessible concept of average for students at this age and that students would consider (among other possible notions of average) 'most common' as one definition of typical.
- *Comparing groups.* By comparing groups, there is opportunity to consider what is the same and what is different about two data sets. We hypothesised that one such comparison may focus students on middle values (although comparing maximum and minimum values is common in those less experienced with data, see Konold et al, 2004). Additionally, if the 'typical' value is not the same for each group, it will provide an opportunity to consider informal notions of sampling variability.
- *Informal Inference.* Watson (2006) contends that average is closely linked to concepts of inference. Makar and Rubin (2007) argue that informal statistical inference is accessible and appropriate for young children in supported inquiry-based contexts. We sought to encourage students to consider questions such as, 'What would you expect the typical height to be in the class next door?' or 'What would you predict the height might be for a new student who entered our class?' These questions were used to prompt students to go beyond description of their data towards making generalisations evidenced by the data they had collected.

Context and Method

The research reported here is situated within a longitudinal study (2006 – 2009) into teachers' emerging practices in teaching inquiry in mathematics and statistics (Makar, 2007; 2008). In this paper, the research question we are addressing is: *How do young children conceptualise 'average' in a classroom of inquiry?* To address this question, we analysed a series of one-hour lessons from a Year 3 classroom (age 8, approximately 26 children) in a large suburban middle class school in Queensland in which children had been immersed in an inquiry culture over the course of the year.

The unit took place in late October to mid-November 2007 and was the final (Term 4) inquiry unit of the year. Students in the class had already completed three inquiry units in the previous terms, each lasting 2-4 weeks: *Investigating Our Hand Spans* (Term 1), *Do We Eat a Healthy Lunch?* (Term 2), and *What Kinds of Appliances at Do We Have at Home?* (Term 3). In the *Hand Span* unit, students worked with their class's hand span data to learn to describe data, to use their data to make comparisons with data from a neighbouring class, and finally, to make predictions about other Year 3 classes (see Makar & Rubin, 2007). As part of that unit, students discussed the most common hand span first as a single value, then, as they used their data to make inferences to unknown classes, began express the uncertainty of their predictions by stating estimates of a range of values based on their own data. For example, in their own class, 16cm was the most common hand span, but for a neighbouring class it was 15 cm. When making a prediction about what hand span might be most common in a third class, students eventually expressed their predictions as a predicted range of values, like '*about* 15 – 17 cm'. In the *Healthy Lunch* and *Appliances* units, students collected, organised, and analysed increasingly complex sets of categorical data. Importantly in all of these units, the teacher focused on developing a classroom culture in which students repeatedly shared and discussed their ideas as they emerged. They were encouraged to debate ideas and articulate their thinking as it evolved.

In the unit described in this paper, the students investigated the driving questions: Is there a typical height for Year 3 students? If so, what is it? Seven lessons were videotaped about twice a week during the unit (not all lessons were videotaped, particularly those in which students were focused on measuring heights, or collecting and organising their data for most of the lesson). The video data were analysed using methods adapted from Flick (2006). Video logs were created for each lesson and the researchers annotated the lessons with particular attention to the framework discussed above. Observations that emerged were compared and debated with specific episodes identified to be transcribed for more detailed analysis. Selected transcribed segments were then analysed for meanings, focusing on progressions of student thinking over the series of lessons.

Results

Four key concepts of 'typical' (average) emerged as students debated the inquiry question: (1) Typical as a reasonable range of values; (2) Typical as the most common value or interval of data in the class; (3) Typical as contrasted with atypical heights; and (4) Debating what's 'typical' beyond their own class. Below we describe the progression of student thinking over the course of the unit, illustrating these ideas with excerpts from student discussion.

Typical as Reasonable

As students were measuring one another, their initial ideas of typical were values that seemed *reasonable*. Because students at this age have little experience with measurement beyond 100 cm, there were often errors in the values that students recorded as they measured their heights. Heather, for example, noted that the 32 cm that a peer had written didn't make sense, "because 32 is a bit too short" [25 Oct, 25:42].

As she compiles the data in her notebook, Barbara uses the word *sensible* to describe values that seem reasonable to her to be the typical height.

Teacher: So are you saying here, that you don't expect a lot of people to be 155?

Barbara: I'll show you, this. I am extra super, duper, duper sure that no one is 155 except for Charles.

Teacher: Because?

Barbara: Charles is the tallest in the class!

Teacher: Well what about one of the other numbers, then, 132?

Barbara: Well, 132 is a sensible one. I think lots of, most people will go on 132. Unless I find a more sensible one, like 130 or something. [29 Oct, 26:55]

Typical as Most Common

The greatest amount of time in the class was debating whether *typical* was the same as *most common*, and whether 'typical' referred to a range of values or to a single value. Students had explored 'most common' in their *Hand Spans* unit earlier in the year, but the current unit raised new issues to consider. For one, the data were now more complex because there were a much greater number of possible values for heights than had come up during the *Hand Spans* unit. As they organised their data, some groups clustered the data into tens and tallied heights that were in the 120s, 130s, 140s, and 150s, thereby reducing the complexity of their data. From early on in the unit, some students began deliberating 'typical' as a range of values, but struggled to articulate their reasoning.

> *Teacher*: [To the class] Is there a typical height for a person in Year 3? … Have you found that every person in the class is the same height? What are you finding out? Yes, Brett?
>
> *Brett*: Well, I would think that. Um. The height of. A typical height for a person in Year 3 would be like around 128 to 135.
>
> *Teacher*: … So why are you saying those two numbers?
>
> *Brett*: Because it's. Um. 128 is going. Keeps on going until. [long pause] [29 Oct, 36:39]

Here, Brett hypothesises a range of possible values for the typical height of a Year 3 student, but despite further encouragement, had difficulty in grounding his tentative assertion with evidence. His classmate Melanie also speculated that the typical height would be a range of values, but appeared to focus on grouped intervals after her group decided to begin tallying the data in intervals of 120s, 130s, 140s, and 150s centimetres:

> *Melanie*: Um, well, I think is that, um, that, um, most of the, um, Year 3s in the class, um, end up at, yeah, 130-something.
>
> *Teacher*: … Did you have some people that were 130-something, Melanie?
>
> *Melanie*: Yes. But um. [long pause]
>
> *Teacher*: Did you have lots of people? Or all of the people?
>
> *Melanie*: Well, not all of the people. But, yes, I think most of the people would be about 130-something. [29 Oct, 39:22]

Another idea that was raised by students is the idea of typical as 'most popular'.

> *Amy*: Well, I think typical means, like, the most popular.

Barbara disagreed.

> *Barbara*: Well, according to what I wrote …I'd actually say [the typical height is] 132. Because that's the only height, according to mine, that's [three people]…the others only have one or two.
>
> *Teacher*: Right, so what Barbara is telling us is that she's vying for 132 cm as typical because … three people according to Barbara have 132 cm so that's leading Barbara to say that that's typical. [5 Nov, 24:30]

At this stage students were putting forth two conceptual possibilities for the meaning of 'typical': (1) typical as a range of *reasonable* values based on data collected (Barbara, Brett); and (2) typical as the *most common* single value or interval of grouped data (Melanie, Amy, Barbara). Here, students are looking at their own class data and pondering whether 'most common' is a single value or an interval.

Typical in the Population – Making Inferences Beyond Their Data

When the teacher asks what would happen if they collected data from next door, the students debate what might happen. Sophie proposes taking a student from their own class that is 132 cm (the most common height) but Elaine realises they have a problem. She worries about what would happen if the 'most common' height next door were different.

Sophie: I wanted to go to that class and that class (pointing). We could actually find out if there *is* a typical height in year three. … I'll need Elaine to go with me because she is 132 cm.

Teacher: So are you going next door to try and find some people who are 132 cm? [Yes.] So you have decided that's your [typical] height measurement? [9 Nov, 1:42]

The teacher asks if going in with a set idea of what is typical will influence her findings.

Elaine: What you are saying is kind of true, [but] 32 (sic) might not be the typical height. But, um, possibly [interruption from the loudspeaker]. For example if there were five people that were 132 cm [in our class] and there were five, six people in [the class next door] which were 134 cm, like, that, yeah, so, yeah. [Makes a face] [3:34]

Elaine's suggestion prompts Sophie to rethink her plan and propose that they may actually need to look beyond their two classes for what is the typical height:

Sophie: Yes if, um, my idea is to say that if there was a typical height in grade 3, now let's just say, um, there are more people in our class that were 137 and more people in that class that were 136! Well, you should say my idea of typical is not what is now, not what's exactly, not what's [our data] *now*, I mean, not what has the *most*, but I'm saying what, what, um the *supposedly* measurement would be. Because not many people regularly are as tall as Charles. Or Dan. I mean there's a couple. [4:48]

Sophie's idea appears to encompass a third and fourth concept for what could be a typical height. She seems to be saying that in fact the two classes may have different (single) 'most common' heights. She proposes that a typical height for a Year 3 student would not just be the most common for their two classes ('what is *now*'), but seems to suggest that there is a typical beyond this (the 'supposedly' measurement) to what is common in the *population*, and contrasting these with heights of people that are unusually tall (outliers Charles and Dan). The teacher here takes the opportunity to introduce the word *atypical* to students and, building on Sophie's proposal, asks the class to consider whether there are typical and atypical heights for Year 3 students.

After collecting height data from the class next door and making comparisons, students observe that, in fact, the data from two classes do have differences (minimum, maximum, most common, and frequencies of interval heights in the 120s, 130s, 140s, and 150s cm). Students suggest that they put the data from the two classes together to find which value is most common (137 cm, different than either class). Considering the combined data propels students into deeper discussions of typical height. They deliberate several ideas—Is the typical height a single value or a range of values? Is it the greatest value (or range) for their two combined classes or is more data needed?—their proposals becoming more inferential. That is, students used their data to talk about typical heights more broadly, beyond the data they had collected. Additional issues began to arise as well as they combined the data sets. For one, they observed that middle range frequencies were higher than for the outer intervals, making informal observations about the shape of the data:

Melanie: I've made tall, medium, and small columns, and most, and the most, and the column that's got the most amount of people is in the medium. [21 Nov, 6:24]

Students continued to build on one another's ideas over the course of the final lesson:

- Sophie suggests that typical heights seem to be somewhere between 125-140cm and that those outside that range have only a few students, so are atypical.
- Melanie further proposes that if you were to combine three or four classes together, that the number of heights in the middle would get bigger and bigger.
- Emily argues that "if you went around the world" [17:45], you'd find that typical heights would be different in different places. She uses her Vietnamese background as a representative example of an Asian population, because Vietnamese children tend to be shorter than other Australian children.

These ideas prompt heated discussions when students return to their collaborative groups to come up with a conclusion for the inquiry question. They argue whether you can say there is a typical at all – as it may vary from country to country:

Elaine: If you go around the world, and you visit each Year 3 class, there's people that are like Charles and taller (untrans)

Charles: That would take a long time!

Elaine: … Some people are in the middle like us. Some people, like in Vietnam, they're really small.

Sue: And people in Japan are really tall!

Elaine: Yeah, so there is actually no typical height.

Charles: … I reckon there's only a typical height for every country. [26:31]

The class wraps up the unit with a whole class discussion of whether there is a typical height. Students decide to propose that the typical height for Year 3 students in Australia to be around 130-138 cm, but caution that as "more people from different countries are coming into Australia" [Melanie, 53:30], it may change.

Discussion

This study asked young children to investigate the question, *Is there a typical height for Year 3 children?* The results suggest that young children are more than capable of considering informal concepts of range, outliers, group comparisons, samples, populations, and informal inference in an inquiry-based environment. Their first conceptions of average focused on reasonable values for the context. Initial language about 'typical heights' were colloquial and descriptive, with students relying on context and previous understandings to structure new ideas. Students were able to converse mathematically about 'typical heights' because they were able to utilise the context to link statistical concepts with their meanings. These ideas were made accessible to students who might normally struggle in more conventional classrooms. As their language became more mathematically precise, their perception of the meanings of typical deepened and extended beyond their own data to populations beyond the classroom. By sharing ideas within a community of learners, students generated, reconsidered and built concepts through peer interaction as they thought aloud, modified ideas, argued, and shaped their thinking about mathematics. Both the teacher and students provided language models to articulate and clarify emerging ideas.

An important aspect of the study was that students were using their own data and representations to make sense of the inquiry question. This allowed them to shape the familiar into increasingly more sophisticated and abstract conceptual tools. By drawing on their own context, students made emotional and personalised connections that scaffolded them through the messiness and uncertainty that often accompanies inquiry learning. These experiences worked to break students' mathematical preconceptions of questions in mathematics having single, irrevocably true results.

The authenticity of the data had two important effects. For one, discussing heights of children was not an abstract notion, nor was it just about describing data. The familiarity of having Charles (as an outlier) and Emily (as representing wider populations) in the class generated discussion about specific data tools. Secondly, complexity of authentic situations provided opportunities for the children to dabble in challenging statistical concepts well beyond what might be expected from this year level. However, the authenticity of the problem required a culture of inquiry in order to encourage these notions to emerge through debate and deliberation.

Although this paper describes only a single classroom, it provides insight into possibilities when children are allowed to grapple with ill-structured problems. Using this classroom as an illustration, this study can help teachers undertake and design messy problems to get beyond uncritical conceptions of average—a different pedagogy to teaching statistics. A classroom culture designed to explore and deliberate key ideas in a meaningful context supports this approach.

Acknowledgements

This study was funded by the Australian Research Council, Education Queensland, and The University of Queensland. The authors are grateful to the teachers and schools that are vital partners in this research.

References

Ben-Zvi, D. & Sfard, A. (2007). Ariadne's thread, Daedalus' wings, and the learner's autonomy. *Education and Didactique, 1*(3), 123-142.

Curriculum Corporation (2006). *National Statements of Learning in Mathematics*. Melbourne: Author.

Flick, U. (2006). *Introduction to qualitative research* (3rd Edition). Thousand Oaks: Sage Publications.

Konold, C., Higgins, T., Russell, S.J., & Khalil, K. (2004). *Data Seen through Different Lenses*. Unpublished manuscript. http://srri.umass.edu/files/konold-2004dst.pdf, Accessed 9 March 2009.

Konold, C. & Pollatsek, A. (2002). Data analysis as the search for signals in noisy processes. *Journal for Research in Mathematics Education, 33*(4), 259-289.

Makar, K. (2007). Connection levers: Supports for building teachers' confidence and commitment to teach mathematics and statistics through inquiry. *Mathematics Teacher Education and Development, 8,* 48-73.

Makar, K. (2008). *A model of learning to teach statistical inquiry*. Paper presented at the Joint ICMI/IASE Study, Statistics Education in School Mathematics: Challenges for Teaching and Teacher Education. ITESM, Monterrey Mexico, 30 June – 4 July 2008.

Makar, K. & Rubin, A. (2007). *Beyond the bar graph: Teaching statistical inference in primary school*. Paper presented at the Fifth International Forum for Research on Statistical Reasoning, Thinking and Literacy, The University of Warwick, UK. 11-17 August 2007.

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

National Council of Teachers of Mathematics (2000). *Principles and standards for school mathematics*. Reston, VA: Author.

Queensland Studies Authority (2006). *Essential Learnings.* Brisbane: Author.

Reitman, W. (1965). *Cognition and thought: An information-processing approach*. New York: Wiley.

Watson, J. (2006). *Statistical literacy at school: Growth and goals.* Mahwah NJ: Lawrence Erlbaum.

Watson, J. (2007). The role of cognitive conflict in developing students' understanding of average. Educational Studies in Mathematics, 65, 21-47.