EMERGENCE OF REASONING ABOUT SAMPLING AMONG YOUNG STUDENTS IN THE CONTEXT OF INFORMAL INFERENTIAL REASONING

Einat Gil and Dani Ben-Zvi
University of Haifa, Israel
egil@construct.haifa.ac.il

This paper discusses students' evolving statistical reasoning about randomness and sampling in the context of inquiry-based activities designed to develop their informal inferential reasoning (IIR). The knowledge of sampling and randomness are key concepts to understanding statistical inference (Garfield & Ben-Zvi, 2008). In the 'Connections' project (Ben-Zvi, Gil & Apel, 2007), sixth grade students were engaged in an inquiry-based learning environment using TinkerPlots (Konold & Miller, 2005) that was designed to develop their IIR. In this design experiment (Brown, 1992; Collins, 1992), the students' intuitive concepts of sampling and randomness were used to design instructional activities that nurture the emergence of ideas of random vs. biased sample and inference. This knowledge was later applied by the students to investigate authentic data and draw informal statistical inferences from a random sample to a population.

THEORETICAL BACKGROUND

In this paper we study the emergence of students' reasoning about sample and sampling in the context of inquiry-based activities designed to develop their Informal Inferential Reasoning (IIR). IIR refers to "the cognitive activities involved in informally drawing conclusions (generalizations) from data (samples) about a 'wider universe' (the population), while attending to the strength and limitations of the sampling and the drawn inferences" (Ben-Zvi et al., 2007), and "articulating the uncertainty embedded in an inference" (Makar & Rubin, 2009). IIR involves a consideration of multiple dimensions: properties of data aggregates, the idea of signal and noise, various forms of variability, ideas about sample size and the sampling procedure, representativeness, controlling for bias, and tendency (Rubin, Hammerman & Konold, 2006).

Ideas of sampling and using samples for statistical inference are at the heart of statistical investigations (Garfield & Ben-Zvi, 2008). Two central ideas of sampling–sampling representativeness and sampling variability–are important and related foundations for understanding statistical inference. Overreliance on sampling representativeness leads students to think that a sample tells us everything about a population, while overreliance on sampling variability leads students to think that a sample tells us nothing useful about a population (Rubin, Bruce, & Tenney, 1991). According to Tversky and Kahneman (1971), people tend to rely too much on small random samples and perceive them as representative. This has been suggested to be true also for school students (Shaughnessy, Garfield & Greer, 1996).

The term *randomness* is related in everyday use to incidental events (contrary to intentional acts), while in statistics it is related to the principle of equal probability (Batanero & Serrano, 1999). To overcome bias in sampling that might be caused by personal choice, a statistician uses random sampling. Simple Random Sample (SRS) consists of n individuals from the population chosen in such a way that every individual has an equal chance to be in the sample, and every sample in the size n has an equal chance to be selected (Eisenbach, 2005). A larger sample is more likely to predict the desired parameter and thus to produce a smaller sampling variability (Moore & McCabe, 2006).

Difficulties in understanding and using the concept of sample, random sample and sampling biases are described in the literature. For example, Metz (1999) found that many elementary school students thought that they cannot draw inferences from a sample to a population due to the need to ask everyone in the population. Watson (2004) classified children's reasoning about samples to six hierarchical categories of understanding that takes into account reference to sampling method (random/biased), sample size and other sample/sampling characteristics. Her findings showed that elementary school students improved their understanding of a random sample and sampling bias during the four years between one interview to another.

Jacobs (1999) investigated students' reasoning about sampling before they learned the subject formally, and found that about a third of the students (grades 4 -5) estimated correctly the quality of the survey distinguishing between a sample taken randomly that results in a non-biased

In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8, July, 2010), Ljubljana, Slovenia*. Voorburg, The Netherlands: International Statistical Institute. www.stat.auckland.ac.nz/~iase/publications.php

sample and a sample taken by a deliberate choice that results with a biased sample. Other students (12%) preferred the principle of fairness and tended not to trust random samples. Lack of representation of the diversity in the population has been suggested as a source for students' preference of biased samples and distrust of random samples (Jacobs, 1999; Watson, 2004).

Samples and sampling are included in current school statistics curricula. For example, students in grades 5 to 8 in the USA are expected to view sampling as a critical issue in data collection by considering, "whether every student must be questioned and if not, how can randomness in the sampling be assured … random samples, bias in sampling procedure, and limited samples all are important considerations" (National Council of Teachers of Mathematics, 1989, p. 106). In Australia, the National Statement on Mathematics for Australian Schools (AEC, 1991) states that "the dual notions of sampling and of making inferences about populations, based on samples, are fundamental to prediction and decision making in many aspects of life. Students will need a great many experiences to enable them to understand principles underlying sampling and statistical inference…" (AEC, 1991, p. 164).

Several learning trajectories are suggested for teaching concepts related to sample and sampling (e.g., Garfield & Ben-Zvi, 2008). Rubin et al. (1991), based on a senior high school study, suggest the following sequence of ideas: a) the center value of the sample is not necessarily identical to that of the population but can give good estimates about its center value; b) there is a difference between different samples, sampled from the same population (more variability will be between smaller samples); c) the difference between one sample to another is not due to a mistake but occurs because of the sampling process. It happens even if the sampling process is not biased; d) although a statistic is different than the parameter value, they are different in a way that one can anticipate/predict it. Most sample statistics are close to the parameter and a few are significantly larger or smaller than it; e) a sufficiently large random sample can give good prediction of the parameter; and f) this prediction relates to the sample size, larger samples can give statistics with smaller variability from the parameter than smaller samples.

METHOD

The research question

The following research question was formulated to further study some of the challenges that have been identified in students' reasoning about samples and sampling. It focuses on how sixth-grade students' reasoning about sample and sampling emerges in the context of an inquiry-based learning environment aiming at developing IIR: *What perceptions of samples and sampling emerge among young students in this context of studying IIR?*

The Setting

In the *Connections* Project (grades 4–6, 2005–2007), the investigators, statistics education researchers from the University of Haifa, worked with primary school teachers and students to study students' evolving ideas of statistical reasoning within a statistical inquiry cycle in a computerized learning environment. Students actively experienced some of the processes involved in experts' practice of data-based inquiry by working on data scenarios, investigated by peer collaboration and classroom discussions. A central feature of the learning environment was the use of *TinkerPlots* (Konold & Miller, 2005), a statistical dynamic visualization tool that is designed to help students develop statistical reasoning by providing new interactive and dynamic ways of representing data.

The sixth grade learning trajectory (Gil & Ben-Zvi, 2007) provides ample opportunities for students throughout the five-week intervention to account for, describe, explain and argue as they draw random samples and make informal inferences. In addition, it provided a window for the researchers to uncover students' intuitive and emerging understanding of these concepts while they experienced guided and open-ended hands-on activities to support the understanding of random sampling (such as the "Stringing Students Along" activity, Shaughnessy & Chance, 2005, p. 43). A gradual transition from concrete to abstract methods of drawing random samples was offered, for example, by drawing at first folded notes from a hat, followed by drawing random numbers in *TinkerPlots*, to using a "slider" and "hot keys" in *TinkerPlots* to resample and increase sample size.

In the second half of the curriculum, students autonomously posed the questions they wished to investigate about the population, formulated hypotheses, analyzed additional samples of data, interpreted the results and drew informal inferences about the population. At this stage, students investigated data that was collected from all students in grades 6 and 7 in the school using a 17-item questionnaire about gender and age, issues related to transfer from primary school to middle school (e.g., homework load) and sportsmanship (e.g., long jump results, favorite sport). The population data (n=206) was never exposed to the students, who were only allowed to randomly sample from this file in order to informally infer about the parameters of this population.

Participants and episodes

The *Connections* intervention took place at a science-focused primary school in Haifa, Israel. Most of the students in this school came from affluent backgrounds and participated in the *Connections* lessons in fourth and fifth grades. These previous encounters made them fluent with the software and basic informal statistical ideas, language, skills and perspectives. In the current study we follow a group of three sixth grade students - Odi, Eli and Asi (pseudonyms, males, age 12), very good math students. Asi is a gifted student who occasionally joins the group. The participants were high academic and verbal students chosen to provide a rich source of information about their learning and reasoning as they worked independently (resembling the “intensive sampling method”, Patton, 2002).

Three episodes from the participants’ work will be reported in the ICOTS-8 presentation, of which two segments are briefly discussed below to provide a flavor of the data and context. In episode 1, Odi and Eli worked on their first data investigation (of the second half of the curriculum). They investigated a self-chosen question, comparing the homework load of sixth and seventh graders, using a random sample (n=20). The second segment took place a few days later while Odi, Eli and Asi studied the same topic using another random sample (n=30).

The third segment is taken from a different investigation that occurred a few days later. The trio felt unsatisfied with the triviality of the homework load investigation, and decided to look for a more challenging and interesting research topic: a comparison of long jump results between sixth and seventh graders in relation to favorite sports. They started by comparing long jump results from a random sample (n=20) and the result seemed surprising and unreasonable to them: sixth grade mean was greater than seventh grade mean. This finding dramatically affected the course of their investigation.

Analysis

To examine the emergence of reasoning about sample and sampling in IIR, these inquiry episodes were fully videotaped and transcribed. The analysis of the videotapes was based on interpretive microanalysis (see, for example, Meira, 1998): A qualitative detailed analysis of the transcripts, taking into account verbal, gestural and symbolic actions within the situations in which they occurred. The goal of such an analysis was to infer and trace the development of cognitive structures and the sociocultural processes of understanding and learning. Utterances and reasoning processes about samples and sampling were carefully identified by two and sometimes three experienced researchers. The researchers discussed, presented, and advanced and/or rejected hypotheses, interpretations, and inferences about the students’ cognitive structures. Advancing or rejecting an interpretation required: a) providing as many pieces of evidence as possible (including past and/or future episodes and all sources of data as described earlier), and b) attempting to produce equally strong alternative interpretations based on the available evidence.

RESULTS

In this section we present two segments of students’ reasoning about sample and sampling as they informally infer from a sample.

Segment 1: Initial ambiguity regarding the value of random sampling

In the following transcript, taken from the first investigation with *TinkerPlots*, the students infer that seventh graders have larger homework load than sixth graders. The researcher asks them how sure they are in their inference.

17	R	… How sure are you about the sample, in these results?
18	Odi	For now ah…
19	Eli	We are not sure since we chose [the sample] randomly, so maybe… when we compare it [this sample to other samples] - we might really see that the results are different.
20	R	Different than what?
21	Eli	[Different] than other samples… so the results may be different since we in fact chose randomly…
22	R	So is it good or bad that you have chosen randomly?
23	Odi	Quite good
24	R	Why does it look quite good to you?
25	Odi	[It looks quite good] since if we chose [only] the excellent students, it [the results] would not be so interesting.
26	R	I didn't hear you.
27	Odi	If we chose only the excellent students, it would have not been so interesting, since everyone [in this sample] could complete it [the homework] in an hour or even less.
28	R	[To Eli] And what do you think? You said that random sampling is good. Does it help or do you find it problematic?
29	Eli	It [random sampling] is very good, since everyone has a chance to be in the sample, unlike [a situation in which] the weaker students, the better students, or students that have more free time [will be deliberately chosen].

In response to the researcher's question about their level of confidence in their inference, Eli answers that they are not sure (19). The reasons he provides for his doubt are related to random sampling and sampling variability (19, 21). The students seem ambiguous at this stage about the role and value of randomness in sampling. This ambiguity drives on the one hand their doubt in the strength of their inferences, but on the other hand they start to appreciate the benefits of random sampling. They are aware that a sample taken by a deliberate choice that includes only one type of students, such as academically outstanding students, is less interesting because it has less variability (Odi, 26); and that a random sample gives each individual an equal chance to be included in the sample (Eli, 28). Their explanations reflect a disharmony of their perception of random sampling when they consider their level of confidence in their inferences. They thus find it hard to decide between their emerging recognition of the advantage of random sampling over a sample taken by a deliberate choice that results with a less representative sample, and at the same time their doubt in random sampling due to expected relatively large sampling variability.

Segment 2: Sample representativeness and randomness

The second segment is taken from the trio's investigation of a second sample (n=30) in which the homework load of seventh grade is larger on average than sixth grade. The researcher asks Eli, Odi and Asi whether this sample is representative of the population.

40	Asi	I think it [the sample] represents [the population]… it's also reasonable to assume that seventh graders will have greater [homework] load… heavier than sixth graders… [this inference is true] unrelated to the [results of this] sample.
41	R	Ok, you are basing [your conclusion] also on…
42	Asi	Both on the sample and on [unheard]
43	R	Both on the sample and on… how reasonable it is … a reasonable inference?
44	Asi	Yes
45	R	If the conclusions were unreasonable in your opinion, what would you think?
46	Asi	I would think that the sample is probably not reliable or…
47	R	And what would you do?
48	Eli	Just a minute… maybe we did not choose [the sample] so randomly… [if the conclusions were unreasonable] we would investigate ourselves to see if we really chose [the sample] in a random way.

Asi claims that the second sample represents the population well and he considers the results reasonable context-wise (40). The reasonableness of the results seems to support his acknowledging of the representativeness of the sample. At this point, the investigator posed a hypothetical experiment to the students, in order to confront them with a conflicting situation, in which the sample results seem unreasonable. In response, Asi replies that if the sample results were not reasonable he would consider the sample unreliable (46). Eli spontaneously adds that he would suspect in this case that the drawing of the sample was not done randomly (48). Thus, another ambiguity rises, whether unreasonable results mean unreliable sample or reflect biased sampling methods.

In sum, in these two segments we observe students' rather sophisticated ideas of sample and sampling, as they evolve, while trying to infer from a random sample about the population. In the first segment we observed a duality in students' considering a random sample both as a "good" and useful tool in making informal inference while concurrently reflecting a lower level of confidence in the inference since it is based on random sampling. The second segment focused on the representativeness of the random sample and the reasonability of the inference. The students provided two explanations to the sources of an unreasonable hypothetical data-based inference: an unreliable sample or a biased sampling method. In the ICOTS-8 presentation a third segment will be presented, in which the three students realize that the sample results are unreasonable, and struggle with explaining this conflicting situation in terms of the problem context, confounding variable and their understanding of sample and sampling.

DISCUSSION

In this paper we presented one possible snapshot of the emergence of reasoning about sample and sampling of six grade students in an inquiry-based learning environment, designed to promote IIR. Through investigations of an authentic topic and an attempt to informally infer about the school's 6th and 7th grade population from random samples, the students struggled with the meaning of random sampling and were ambiguous regarding its contribution to their inference. The six graders' ideas of a random sample in the context of IIR included recognition of the importance of giving an equal opportunity to every student to be included in the random sample. At the same time, they acknowledge the existence of sampling variability and therefore random sampling is perceived by them as reducing level of confidence in the inference. They are concerned that a different random sample will show different, sometimes contradicting, results. Moreover, unreasonable sample results are linked to rechecking sampling methods and the idea of randomness.

These results are aligned with a development in the emergence of ideas associated with sample and sampling that was described by Gil (2008) in the *Connections* project. In particular, the key ideas of randomness and random sampling were partially understood and used by students at this age who were able to consider the implications of sampling representativeness and sampling variability, but not to resolve the relations among them. It seems that a learning environment designed to promote IIR such as the one described in this paper might help to promote the emergence of reasoning about random sampling. In another study, elementary school students gradually improved their reasoning about sampling toward recognizing the value of random sampling (Watson, 2004). This improvement was verified in an interview with the same students four years later.

Students' difficulties to rely upon a random sample due to overreliance on sampling variability were described also by Rubin, Bruce and Tenney (1991). We suggest that special attention ought to be given in curricula design to building understanding and authentic experiences of random sampling (cf., Garfield & Ben-Zvi (2008). In light of the complexity involved in the concepts of sample and random sampling, further pedagogical efforts should be carried out to address the multiple challenges. As seen in the *Connections* study, understanding and using the concept of random samples entails complex perceptions for the young learners that are far beyond the formal definition of a simple random sample.

REFERENCES

Australian Education Council. (1991). *A national statement on mathematics for Australian schools.* Carlton, Vic.: AEC.

Batanero, C., & Serrano, L. (1999). The meaning of randomness for secondary students. *Journal for Research in Mathematics Education*, *30*, 558-567.

Ben-Zvi, D., Gil, E., & Apel, N. (2007). What is hidden beyond the data? Helping young students to reason and argue about some wider universe. In D. Pratt & J. Ainley (Eds.), *Reasoning about Informal Inferential Statistical Reasoning: A collection of current research studies. Proceedings of the Fifth International Research Forum on Statistical Reasoning, Thinking, and Literacy (SRTL-5)*, University of Warwick, UK, August, 2007.

Eisenbach, R. (2005). *Statistics for Non-statisticians*. Jerusalem: The Hebrew University, Academon (Hebrew).

Garfield, J., & Ben-Zvi, D. (2008). *Developing students' statistical reasoning: Connecting Research and Teaching Practice*. Springer.

Gil, E. (2008). *Development of informal inferential statistical reasoning among young children in research-based learning environment*. Unpublished MA Thesis, University of Haifa, Haifa (Hebrew).

Gil, E., & Ben-Zvi, D. (2007). *Informal Statistical Inference: Data Analysis for Sixth Grade Using TinkerPlots – Student's Workbook.* Haifa, Israel: University of Haifa (Hebrew).

Jacobs, V. R. (1999). How do students think about statistical sampling before instruction? *Mathematics Teaching in the Middle School*, *5*(4), 240-246,263.

Konold, C., & Miller, C. (2005). *TinkerPlots: Dynamic data exploration, statistics software for middle school curricula.* Emeryville, CA: Key Curriculum Press.

Makar, K., & Rubin, A. (2009). A Framework for thinking about Informal Inference. *Statistics Education Research Journal*, *8*(1), 82-105.

Meira, L. (1998). Making sense of instructional devices: The emergence of transparency in mathematical activity. *Journal for Research in Mathematics Education*, *29*(2), 121-142.

Metz, K. E. (1999). Why sampling works or why it can't: Ideas of young children engaged in research of their own design. In F. Hitt & M. Santos (Eds.), *Proceedings of the 21st annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*, Vol. 2 (pp. 492-498). Cuernavaca, Mexico: PME.

Moore, D., & McCabe, G. P. (2006). *Introduction to the practice of statistics*, (5th Ed.). New York: W. H. Freeman and Company.

National Council of Teachers of Mathematics (1989). *Principles and standards for school mathematics*. Reston, VA: NCTM.

Patton, H. J. (2002). *Qualitative research & evaluation methods*. Thousand Oaks: Sage Publication.

Rubin, A., Bruce, B., & Tenney, Y. (1991). Learning about sampling: Trouble at the core of statistics. In D. Vere-Jones (Ed.), *Proceedings of the Third International Conference on Teaching Statistics*. Vol. 1 (pp. 314-319). Voorburg: International Statistical Institute.

Rubin, A., Hammerman, J. K. L., & Konold, C. (2006). Exploring informal inference with interactive visualization software. In A. Rossman & B. Chance (Eds.), *Proceedings of the Seventh International Conference on Teaching Statistics*, Salvador, Brazil. Voorburg: The Netherland: International Statistical Institute.

Shaughnessy, J. H. & Chance, B. (2005). *Statistical questions from the classroom*. Reston,
VA: National Council of Teachers of Mathematics.

Shaughnessy, J. M., Garfield, J., & Greer, B. (1996). Data handling. In A. J. Bishop, K. Clements, C. Keitel, J. Kilpatrick & C. Laborde (Eds.), *International handbook of mathematics education, Part 1* (pp. 205-237). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin*, *76*(2), 105-110.

Watson, J. M. (2004). Developing reasoning about samples. In D. Ben-Zvi & J. B. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking*. Springer.