



Children's Reasoning about Samples and Sampling in a Project-Based Learning Environment

Efi Paparistodemou, e.paparistodemou@cytanet.com.cy
Cyprus Pedagogical Institute

Maria Meletiou-Mavrotheris, m.mavrotheris@euc.ac.cy
European University Cyprus

Abstract

Building connections between sample and population lies at the heart of informal statistical inference (Pratt, Johnston-Wilder, Ainley and Mason, 2008). As Zieffler, Garfield, delMas and Reading (2008) point out, informal reasoning about statistical inference is the way in which students build connections between observed sample data and unknown or theoretical populations, and the way they make arguments or use data-based evidence to support these connections. The present study aims to investigate children's reasoning about samples and sampling in a project-based learning environment. Children analyzed collected data using TinkerPlots® as an investigation tool, and made a presentation of their findings to the whole school. The research aimed at providing detailed information about upper elementary school children's developing knowledge and intuitions regarding key statistical concepts related to samples and sampling, the type of informal inferential reasoning and thinking possible for the specific age group, and supportive instructional activities. Findings from the study suggest that the use of dynamic statistics software has the potential to enhance statistics instruction by making inferential reasoning accessible to young learners.

Keywords

Sampling, Sample, Informal Inferential Reasoning, TinkerPlots

Literature Review

Most of the research on children's reasoning about samples and sampling conducted in the past had primarily focused on understanding current conceptions rather than on developing them over time (Makar et al., 2011). Recent research, however, has shown that young children can demonstrate quite sophisticated levels of informal reasoning about samples and sampling if provided with an interesting and motivating learning context. Gil & Ben-Zvi (2010), for example, studied reasoning about sample and sampling among Grade 6 students (age 12), in the context of a collaborative, project- and inquiry-based learning environment designed to develop their informal inferential reasoning. They witnessed significant development of children's reasoning about key conceptions related to sample and sampling.

Although recent research on informal inferential reasoning has shown some promise in helping children develop deeper understanding of samples and sampling, research in this area is still at an infant stage. The current study contributes to the existing literature by investigating ways to support the development of primary school children's sampling conceptions in the context of making informal statistical inferences.

Advances of technology provide new tools and opportunities for the opening windows on



mathematical thinking (Noss & Hoyles, 1996). Having such a set of tools widely available to young learners has the potential to give children access to advanced statistical topics including inferential statistics and the broader process of statistical investigation (Makar & Rubin, 2007), by removing computational barriers to inquiry. This leads to a shift in the focus of statistics instruction at the school level from learning statistical tools and procedures (e.g., graphical representations, numerical measures) towards more holistic, process-oriented approaches that go beyond data analysis techniques (Makar & Rubin, 2007). Statistics can be presented as an investigative process that involves formulating questions, collecting data, analyzing data, and drawing data-based conclusions and inferences (Guidelines for Assessment and Instruction in Statistical Education (GAISE) Report, 2005).

Sampling reasoning is at the core of statistical investigations. Sample size and sampling method are the main determinants of the validity of statistical inferences. Despite, however, the central role of sampling in statistics, there has been relatively little research into the development of students' sampling cognitions. Among the few conducted studies on children's early conceptions of samples and sampling are the studies of Jacobs (1999) and Schwartz et al. (1998) who investigated Grades 4 and 5 children's informal understanding of sampling issues in the context of interpreting and evaluating survey results. They found that children's distrust of simple random samples and preference for stratification of the sample or for self-selection was attributed to their pre-occupation with issues of fairness and the wish to ensure representation of the diversity in the population in the sample. Also, children were more likely to identify potential bias issues in restricted sampling methods than in self-selected methods. Watson and Moritz (2000) investigated the characteristics of children's constructions of the concept of sample, and identified two key ideas for developing the sampling concept: appreciation for variation in the population and sensitivity to bias. The authors have found a trend for higher level performance with increasing age. The youngest children in their study (Grade 3, age 8-9) had fairly primitive, idiosyncratic notions of samples and sampling derived from everyday experiences with sample products or medical-related contexts.

In a study of high school students, Rubin et al. (1991) found a tension existing between the ideas of sample variability and sample representativeness. On most instances, students' comments suggested that they over-relied on sample representativeness, believing that a random sample has to be representative of the population, and that not randomness but some other mechanism must have caused sampling variability. Similarly to Rubin et al. (1991), Watson and Kelly (2006) found that elementary and middle school students often express beliefs that in a sample "anything can happen". Saldanha and Thompson (2002), who designed a teaching experiment to develop senior secondary students' concept of sampling distribution, found that due to lack of a suitable sense of the variability and the repeatability of the sampling process, students tended to judge a sample's representativeness only in relation to how different they thought it was to the underlying population parameter and not on how it compared to a clustering of the statistic's values.

Concurring with Clements and Sarama (2007), we espouse hierarchic interactionism, a theoretical framework which views children's development of mathematical reasoning as resulting from an interplay between internal and external factors, including innate competencies and dispositions, maturation, experience with the physical environment, sociocultural experiences, and self-regulatory processes. According to hierarchic interactionism, most content knowledge is acquired along developmental progressions of levels of thinking. These progressions play a special role in children's cognition and learning because they are particularly consistent with children's intuitive knowledge and patterns of thinking and learning at various levels of development, with each level characterized by specific mental objects (e.g., concepts) and actions (processes). The children's



environment and culture affect the rate and depth of their learning along the developmental progressions. Instruction based on learning consistent with natural developmental progressions is more effective, efficient, and generative for the child than learning that does not follow these paths.

The present article contributes to the emerging research literature on the early development of informal inferential reasoning by focusing on children's understanding of sampling issues. Moreover, it describes the interaction between children and the dynamic environment of TinkerPlots as a trajectory for expressing the idea of sample. Building connections between sample and population lies at the heart of informal statistical inference (Pratt et al., 2008). As Zieffler et al. (2008) point out, informal reasoning about statistical inference is the way in which students build connections between observed sample data and unknown or theoretical populations, and the way they make arguments or use data-based evidence to support these connections. The present study aims to investigate children's reasoning about samples and sampling in a project-based learning environment. The research aimed at providing detailed information about upper elementary school children's developing knowledge and intuitions regarding key statistical concepts related to samples and sampling, the type of informal inferential reasoning and thinking possible for the specific age group, and supportive instructional activities.

Methodology

A teaching experiment was designed to promote understanding of sampling issues in a Grade 6 (11 year-old students) classroom. Nineteen children participated in data-centered activities, in contexts familiar to them, which provided them with opportunities to investigate real world problems of statistics using technology. They posed questions of interest to them, devised and carried out a sample data collection plan, and worked in small groups to formulate and evaluate data-based inferences using the dynamic statistics software Tinkerplots[®] as an investigation tool.

During the study, the research team collected and analyzed a wealth of data to assess students' growth in understanding and reasoning about samples and sampling. Students' learning processes were studied using written assessments, audio-recordings of class sessions, video-records of group sessions, interviews of selected students (the interviewing took place while students were working in groups for analyzing their data), and classroom observations and artifacts.

The videotapes collected during the course were first globally viewed and brief notes were made to index them. The goal of this preliminary analysis was to identify representative parts of the videotapes indicative of students' approaches and strategies when performing specific statistical problem solving tasks. The selected occasions were viewed several times and were transcribed. The transcribed data, along with other data collected in the study, were analyzed in order to investigate children's ways of thinking about samples and sampling while informally drawing inferences from data. The results section shares some of the insights gained from the study regarding patterns and mechanisms of development in children's reasoning about samples and sampling.

Results

The main data source for the activities taking place during the teaching episode was a survey developed and administered by the children, which investigated the community service and volunteerism habits of students in their school. The development of children's volunteering ethos was a priority set by the Cyprus Ministry of Education for the entire school year. Students were introduced to the importance of involved and responsible citizenship in a cross-curriculum



environment. Different subjects in the school curriculum aimed at fostering service learning, by informing children about the benefits of volunteerism and by encouraging them to get more actively involved in community service and voluntary work. A number of pro-social activities that provided volunteer opportunities for children were organized by the school.

Being sensitized to the importance of voluntary work, the sixth graders in our study decided to conduct a survey in order to investigate the status of school and community service among students in their school. Towards that purpose, they constructed a survey questionnaire. They worked in small groups, and then in a whole class setting, for finding ‘important questions’ to include in the questionnaire. The constructed questionnaire inquired students about their gender, age, whether they were familiar with each of the main volunteering organizations in Cyprus, and whether they wished to become members of such an organization. It also asked students to indicate the approximate number of times they participate each year in events organized by volunteering organizations, and to specify in which of a range of volunteer work activities planned to be take place at their school they would wish to participate.

Students participating in this research first completed the questionnaire by themselves and then decided to compare their answers with those of their classmates. Children analyzed the data and drew conclusions regarding the volunteerism habits of children in their class. Finally, they started thinking about conducting a survey of the students of the school in order to present their results to a school fair at the end of the year.

Next, students devised and carried out a data collection plan in order to obtain information about the volunteering habits of all students in the school. Given the large number of students in the school, they decided that it would be very difficult to administer the questionnaire to all students in the school. Instead, they decided to collect data from a sample of students from grades four, five, and six of the school. The sample selection process was decided after a long class discussion.

The following whole-class episode shows how students explore the need of having a representative sample from data.



Line 1 Teacher/Second author (T): So, what do you think? How can we get data for
Line 2 this survey?
Line 3 Student 1 (S1): We can get data only from 6th grade. Students at 6th grade are
Line 4 the oldest students at school and they will give better answers.
Line 5 Student 2(S2): But then, we cannot say that our survey comes from the whole
Line 6 school. What about the other classes?
Line 7 Student 3 (S3): *Why don't we ask all the students?*
Line 8 T: That's a good idea. What do others think?
Line 9 S4: Good idea. We need to split in groups of 2 and visit all the classes.
Line 10 T: But, think of having 220 questionnaires to analyze...
Line 11 S3: Is there another way to get a sample that *represents* our school?
Line 12 S2: Well, we need to have students from all the 3 grades. This is for sure.
Line 13 S4: Ok. Let's get the five students from the elected committee of each class.
Line 14 The children of the class voted for these students, so let's ask their opinion.
Line 15 S5: I don't agree with this. It is not fair. These students were selected to
Line 16 *represent* their class in the school's decisions, not to *represent* what all the
Line 17 students think. I didn't want to be in my class committee, but I would like to
Line 18 answer the questionnaire.
Line 19 S6: I agree.
Line 20 S7: Well, for having a *fair sample* of students, we also need to have the same
Line 21 number of boys and girls.
Line 22 T: Why do you think that?
Line 23 S7: We are different from boys. It is not fair to ask more boys than girls.
Line 24 T: So, what shall we do?
Line 25 S8: We need to select children by chance. Without knowing...
Line 26 T: How?
Line 27 S2: We can get the catalogue of each class. We need to select from each
Line 28 catalogue 5 boys and 5 girls.
Line 29
Line 30

The first reaction of some students (S3, Line7) was to ask all the students at school. It is interesting how the particular children came to the conclusion about the need to get a random sample. Fairness seems to be a big issue for them and this is also the reason for deciding to have a stratified sampling method (stratification by class and gender).

The teacher (Line 10) does an intervention in order students to start thinking of the idea of sample. This point was critical because in the following lines the students started to construct the idea of sample. They used phrases like 'that represents our school' (Line 11), and 'fair sample' (Line 20) to justify their decisions. It is also noteworthy that the first "fair decision" for them was to get data from all the students of the school (Line 7). It seems that the possibility of including the whole population, and the practical difficulties that this would entail, was the driving force for deciding to instead select a representative sample. The above episode shows also how the context of the survey influences the idea of having a 'fair' sample. The S5 girl was very interested in this project. This was one of the few times in which she expressed a wish to take part in a class activity. Her comments (Lines 15-18) were critical on continuing the discussion with children. In Lines 27-28, S2 suggests a way of getting a random sample. He suggested a random stratified sample. We believe that came easily to his mind, as children were very familiar to use the



catalogue of the class. Actually, the class catalogue was used for absences, grading etc. The student knew that each class had a catalogue, so it was very easy to him to refer on it.

After collecting these real data about themselves and from a sample of students from the whole school, students worked in small groups to explore the data, using the dynamic statistics software TinkerPlots® as an investigative tool. The class was divided in five groups. Each group got one questionnaire, entered the data in a TinkerPlots datacard, got another questionnaire etc. Because of that, the final number of questionnaires in each group was different. In the end, Group A had access to 15 cases in their datacards, Group B had 26 cases, Group C had 4 cases, Group D had 21 cases and Group E had 31 cases. The big difference between the Group C (4 cases) and the other groups was because of technical problems with the laptop the group used. The teacher provided technical interventions here, that it wasn't worth to mention them.

Each group analyzed the data they had at hand and discussed their findings with each other. In group discussion, students tried to draw conclusions about the data they had in front of them. Firstly, they were making some data-based argumentations (Paparistodemou & Meletiou-Mavrotheris, 2008) like 'children in our data are willing to volunteer' (Student 2, Group A). Group A constructed the following graph (Figure 1):

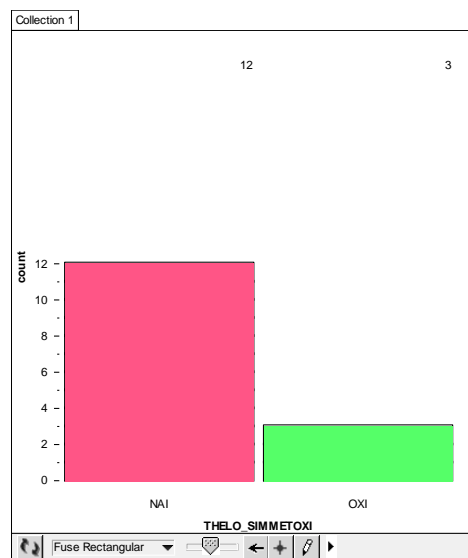


Figure 1: How many students of the data of Group A wanted to volunteer

(‘NAI’=YES, ‘OXI’=NO)

Line 31 S2: We see children in our data are willing to volunteer.

Line 32 S1: From those that they didn't want to...Shall we check if there were boys or girls?

Children constructed another graph based on their question (see Figure 2). They added to their graph the attribute of Gender and they continued their thinking.

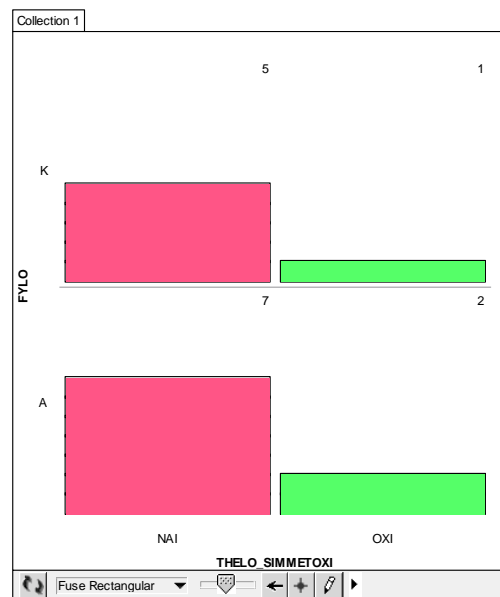


Figure 2: How many students of the data of Group A wanted to volunteer

(‘FYLO’=GENDER, ‘A’=‘BOY’, ‘K’=‘GIRL’)

- Line 33 S1: From the children they said ‘No’, most of them are boys.
 Line 34 S4: But, here we have a *sample of the sample*. We need to see the graphs from all
 Line 35 of our cases and this would be done at the end.
 Line 36 S1: Let me see from which class these boys come from [she clicks on the particular
 Line 37 part of the graph and look at the DataCard].

The above episode is an example of how children interacted with TinkerPlots in order to draw conclusions about their data. It is interesting how S4 express the meaning of the sample (Line 34). It seems to accept the results from the whole sample but need more evidence for drawing conclusions from a small number of data. Moreover, the children made some data-based argumentations and generalizations (Paparistodemou & Meletiou, 2008) like ‘The whole school is willing to help’ (Student 6, Group C), ‘For most of our students to respect means to be kind’(Student 4, Group A).

After finishing with the datacards, the teacher (first author) got the different files from groups, and at the end she got all the cases in one file for analysis (97 cases). She then initiated a whole class discussion:

- Line 38 Teacher: Let me open the cases of Group A. We have 15 cases. What do you
 Line 39 think of the analysis of the results?
 Line 40 S9: We need to say that all 15 cases were *randomly selected*.
 Line 41 S1: Yes, but this sample is not *representative*. The children are not enough.
 Line 42 S2: Ok. They were randomly selected, but they are not too many.
 Line 43 S5: We cannot say that the findings from this group are the final findings.
 Line 44 S10: *We need to have a file with all of the cases.*



- Line 45 T: How about Group B's cases? Look, here we have 26 cases.
 Line 46 S11: I think it is the same as before. *I think* that we cannot have only 26 cases,
 Line 47 for analysing 220 [the total students of the school]. 26 students is the number
 Line 48 of only one class.
 Line 49 S4: But, it is better than having 15 cases.
 Line 50 S2: Yes, it is better, but not enough. More children means more answers. That
 Line 51 means we have a better opinion of what is happening in our school. A better
 Line 52 opinion of what's happening, but not the best!
 Line 53 T: So...I am not going to ask you about the 4 cases of Group C...
 Line 54 S12: *Definitely* we cannot draw any conclusions. The number is too small.
 Line 55 S2: Too small of a number. *That's for sure*. If we think that we have 116 more
 Line 56 opinions...you can imagine...
 Line 57 T: But our sample was randomly selected.
 Line 58 S12: Is there a possibility of having 1 child in each class? If we have one child
 Line 59 from each class, then with 3 cases we can say something.
 Line 60 T: Hmm...
 Line 61 S2: Do you think that one child from grade 4 can represent all grade 4
 Line 62 students? Do you think that the questionnaire you filled for yourself can
 Line 63 represent all grade 6 students?
 Line 64 S12: No, but this is the least we can do. If we had 15 cases, but there is not
 Line 65 even one grade 4 student in these cases, it is worse...
 Line 66 S4: *Why don't we analyse all of our cases?*

The above snapshot shows how children come to realize the disadvantages of drawing generalizations about the population from a small sample. We recognize phrases like 'representative' (Line 41), 'randomly selected' (Line 42), but we also recognize that the number of the cases in a sample influence their opinion. In Line 46, the student uses the phrase 'I think', thus not making a strong statement about the 26 cases. In addition, in Line 54 and Line 55, students are convinced that they cannot draw conclusions from 4 cases. This is the reason that they use strong statements like 'definitely', and 'for sure'. Moreover, it is interesting how stratified sampling seems fair to them. The dynamic software helped children to construct multivariate graphs reflect on stratified sample. In Lines 58-65, S12 argues about having a random sample of 3 cases, but selecting it from all three grades. It is interesting that he is claiming that having a sample of size 3 is better than having a sample of size 15, if the sample of size 3 includes one child from each grade but the sample of size 15 does not.

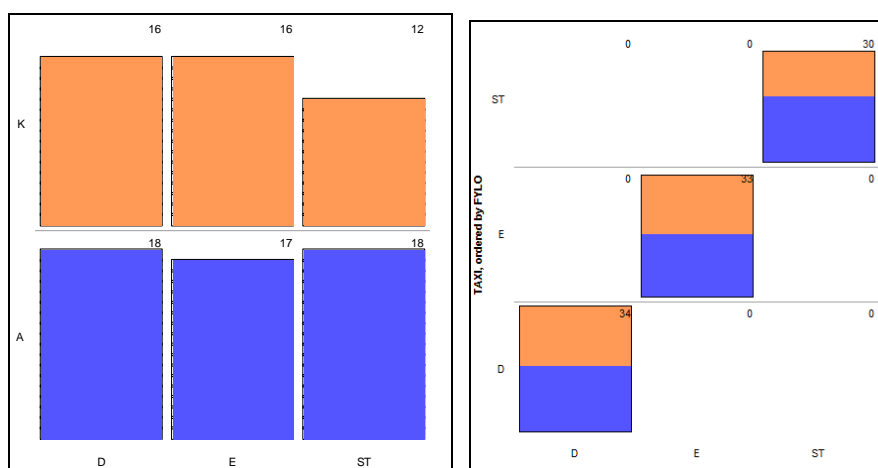


Figure 3: Number of boys and girls ('A'=Boy, 'K'=Girl) in each grade ('TAXI', where 'D'=Grade 4, 'E'=Grade 5, 'St'=Grade 6)



The following episode comes from the analysis of all 97 cases of their randomly selected sample. The children looked at graphs such as the ones in Figure 3 and drew some conclusions regarding the validity of the sampling scheme they chose.

- Line 67 T: What do you think about our data?
Line 68 S7: Our numbers are good. You see, in grade 4 we have almost the same
Line 69 number of students as in grade 5 and in grade 6. The boys are a bit more...
Line 70 S2: Our method is *totally random*. So, having a bit more boys is ok.
Line 71 T: Do you think we can draw 'fair' results about our school now?
Line 72 S4: Yes....You see, now we have 97 cases out of 220 children. And you see,
Line 73 in in each grade we have *almost the same number of students* [in each class].
Line 74 T: So, is this a good sample?
Line 75 S2: Yes. Because it would be very difficult to collect data from all 220
Line 76 children. We chose these children randomly. I think they can represent our
Line 77 school.

In the above case, we see that children are satisfied with the sample selection process they had employed. They use phrases like 'totally random' (Line 70), 'almost the same number [in each class]' (Line 73) to express their satisfaction. Another issue that bothers them again is to ensure that their sample is representative of the whole population of students in their school. In Lines 75-77, S2 provides a justification for not collecting data from the entire population. Besides this justification, it shows that total fairness in results comes from asking the whole population. It is a strong statement that teachers can build on when introducing informal statistical inference in early mathematics instruction.

Conclusions

The 11-year-old students in the study experienced statistics as an investigative, problem-solving process. They formulated questions of interest to them and designed a survey instrument to use for data collection purposes. After a long discussion, they decided that it was more appropriate to use a sample of the population rather than a census to collect their data. They devised and carried out a sample collection plan to answer their research questions. This opportunity to experience the statistical problem-solving process through genuine collection and analysis of sample data, encouraged children to build, refine, and reorganize their intuitive understandings about samples and sampling. Their informal reasoning regarding the effects of sampling method and sample size progressed from rudimentary forms to more sophisticated ones. They began to appreciate the principles underlying sampling, and particularly the need for an adequately large sample size and a random-based sampling procedure.

The students in our study used the dynamic statistics software TinkerPlots® as an investigation tool. The presence of the dynamic software facilitated students' interest in the statistical investigation; it gave them the opportunity to explore data and draw data-based arguments and inferences in ways that would not have been possible for them without the software (Hammerman & Rubin, 2003) like interacting with a constructed graph (see Figure 1 and Figure 2) and drawing conclusions for two attributes at the same time (e.g. Figure 2). Attributes of TinkerPlots® like the ability to operate quickly and accurately, to dynamically link multiple representations, to provide immediate feedback, and to transform an entire representation into a manipulable object enhanced students' flexibility in using representations and provided the means for them to focus on statistical conceptual understanding. This study is an example of an approach to improving



students' use of statistical reasoning and thinking by embedding statistical concepts within a purposeful statistical investigation that brings the context to the forefront. For young children like those participating in our study, personal experience and interest play a key role in learners' interactions with data. Our findings illustrate how young learners can begin to reason about sampling issues and other key inferential ideas when their interest in the task is high. Children's focus during their statistical investigations was on understanding the situation at hand, rather than on examining decontextualized data. Their engagement in an authentic, real world context encouraged students to seek ways to collect sample data that would enable them to draw valid inferences extending beyond their class to the whole school. The children were very much involved with their school project and the conclusions drawn from the data were important for them in order to understand what was happening at their school.

We focus our efforts on building sound foundations of inferential reasoning at a young age. As pointed out by Clements and Sarama (2007), young children possess an informal knowledge of mathematics that is surprising broad and complex. The current study and several other studies (e.g. Paparistodemou & Meletiou-Mavrotheris, 2008) have illustrated that when given the chance to participate in appropriate instructional settings that support active knowledge construction, even very young children can exhibit well-established intuitions for fundamental statistical concepts related to statistical inference. Through genuine data exploration, they can investigate and begin to comprehend abstract statistical concepts, developing a strong conceptual base on which to later build a more formal study of inferential statistics during high school and at the university level.

References

- Clements, D. H., & Sarama, J. (2007). Early childhood mathematics learning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 461-555). Greenwich, CT: Information Age Publishing.
- GAISE Report (2005). *Guidelines for assessment and instruction in statistics education: A Pre-K-12 Curriculum Framework*. Alexandria, VA: The American Statistical Association. [Online: <http://www.amstat.org/education/gaise>]
- Gil, E. & Ben-Zvi, D. (2010). Emergence of reasoning about sampling among young students in the context of informal inferential reasoning. 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.
- Hammerman, J., & Rubin, A. (2003). Reasoning in the presence of variability. In C. Lee (Ed.), *Reasoning about Variability: A Collection of Current Research Studies* [CD-ROM]. Dordrecht, The Netherlands: Kluwer Academic Publisher.
- Makar, K., & Rubin, A. (2007, August). Beyond the bar graph: Teaching informal statistical inference in primary school. Paper presented at the *Fifth International Research Forum on Statistical Reasoning, Thinking, and Literacy (SRTL-5)*, Warwick, UK.
- Makar, K., Wells, J., & Allmond, S. (2011). Is This Game 1 or Game 2? Primary Children's Reasoning About Samples during Inquiry. Paper presented at the *Seventh International Research Forum on Statistical Reasoning, Thinking and Literacy*, Freudenthal Institute, Utrecht, the Netherlands.



- Noss, R., and Hoyles, C. (1996). *Windows on Mathematical Meanings: Learning Cultures and Computers*. Dordrecht: Kluwer
- Paparistodemou, E., & Meletiou-Mavrotheris, M. (2008). Developing young children's informal inference skills in data analysis. *Statistics Education Research Journal*, 7(2), 83-106.
- Pratt, D., Johnston-Wilder, P., Ainley, J., & Mason, J. (2008). Local and global thinking in statistical inference. *Statistics Education Research Journal*, 7, 107-129
- 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, The Netherlands: International Statistical Institute.
- Saldanha, L., & Thompson, P. (2002). Conceptions of sample and their relationship to statistical inference. *Educational Studies in Mathematics*, 51, 257-270.
- Schwartz, D. L., Goldman, S. R., Vye, N. J., Barron, B. J., & The Cognition and Technology Group at Vanderbilt. (1998). Aligning everyday and mathematical reasoning: The case of sampling assumptions. In S. P. Lajoie (Ed.), *Reflections on statistics: Learning, teaching and assessment in grades K-12* (pp. 233-273). Mahwah, NJ: Erlbaum.
- Watson, J.M., & Kelly, B.A. (2006). Expectation versus variation: Students' decision making in a sampling environment. *Canadian Journal of Science, Mathematics and Technology Education*, 6, 145-166.
- Watson, J., and Moritz, J. (2000). Developing Concepts of Sampling. *Journal for Research in Mathematics Education*, 31 (1), 44-70.
- Zieffler, A., Garfield, J., delMas, R., & Reading, C. (2008). A framework to support research on informal inferential reasoning. *Statistics Education Research Journal*, 7(2), 40-58. [Online: <http://www.stat.auckland.ac.nz/serj>].