STUDENTS' REASONING ABOUT UNCERTAINTY WHILE MAKING INFORMAL STATISTICAL INFERENCES IN AN "INTEGRATED PEDAGOGIC APPROACH"

Hana Manor, Dani Ben-Zvi, and Keren Aridor
The University of Haifa, Israel
hana.manor@gmail.com

Reasoning about uncertainty is a key and challenging element in informal statistical inferential reasoning. We designed and implemented an "Integrated Pedagogic Approach" to help students understand the relationship between sample and population in making informal statistical inferences. In this case study we analyze two sixth grade students' reasoning about uncertainty during their first encounters with making informal statistical inferences based on random samples taken from a hidden TinkerPlots2 Sampler. We identified four main stages in the students' reasoning about uncertainty: Account for, examine, control, and quantify uncertainty. In addition, two types of uncertainties–contextual and a statistical–shaped the students' reasoning about uncertainty and played a major role in their transitions from stage to stage. Implications for research and practice are also discussed.

INTRODUCTION

The recognition that judgments based on sample data are inherently uncertain is a key idea in statistical inference. This implies that developing an understanding of statistical inference requires fostering probabilistic considerations. To support students' reasoning about uncertainty in the context of making informal statistical inferences (ISIs), we developed an *Integrated Pedagogic Approach (IPA)* comprised of data and model worlds. In this case study we analyze students' reasoning about uncertainty during their first stages in making ISIs in an inquiry-based learning environment usingTinkerPlots2 (Konold & Miller, 2011).

SCIENTIFIC BACKGROUND

Formal Statistical Inference

The rationale behind collecting data in statistics is learning about real world situations. "Statistical inference moves beyond the data in hand to draw conclusions about some wider universe, taking into account that variation is everywhere and that conclusions are uncertain" (Moore, 2007, p. xxviii). In its simplest form the question of statistical inference deals with the manner of reaching general conclusions about what the true, long run situation is actually like, based on outcomes of a sample that can be collected only once. Given only the sample evidence, the statistician is always unsure of any assertion he makes about the true state of the situation. The theory of statistical inference provides ways to assess this uncertainty and to calculate the probability of error in a particular decision.

Uncertainty and Statistical Inference

The ability to deal intelligently with uncertainty is one of the challenging goals of instruction about statistical inference. Furthermore, our intuitive perception of chance profoundly contradicts the laws of probability that describe actual random behavior (Tversky & Kahneman, 1974). People tend to base predictions of uncertain outcomes on causal explanations instead of information obtained from repeating an experiment (Konold, 1989).This orientation might be problematic in learning statistical inference because students have to consider the relative unusualness of a sampling process outcome.

Informal Statistical Inference

In order to give students a sense of the power of drawing reliable inferences from samples, and given that statistical inference is challenging for most students (Garfield & Ben-Zvi, 2008), Informal Statistical Inference (ISI) and Informal Inferential Reasoning (IIR) have recently became a focus of research (Pratt & Ainley, 2008; Makar, Bakker, & Ben-Zvi, 2011). ISI is a data-based generalization made without formal statistical procedures that includes an articulated component of

In K. Makar, B. de Sousa, & R. Gould (Eds.), *Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9, July, 2014), Flagstaff, Arizona, USA.* Voorburg, The Netherlands: International Statistical Institute. iase-web.org

uncertainty (Makar & Rubin, 2009). IIR - the reasoning processes that lead to the formulation of an ISI- include “the cognitive activities involved in informally drawing conclusions or making predictions about ‘some wider universe’ from patterns, representations, statistical measures and statistical models of random samples, while attending to the strength and limitations of the sampling and the drawn inferences” (Ben-Zvi, Gil, & Apel, 2007, p. 2).

METHOD

This study is part of an extended design research (Cobb et al., 2003) of the *Connections* Project (2005-2015) that aims to study children’s statistical reasoning in an inquiry-based and technology-enhanced statistics learning environment in grades 4 to 9 (Ben-Zvi et al., 2007). This paper focuses on the research question: *How can students’ reasoning about uncertainty emerge while making ISIs in an IPA learning environment*? To address this question, we analyze a pair of students’ reasoning while drawing ISIs using *TinkerPlots2*.

The IPA was developed to guide the design and analysis of the experimental tasks (as part of Manor’s PhD study). It is comprised of *data* and *model* worlds to help students learn about the relationship between sample and population. In the *data world*, students collect real sample, frequently through a random sampling process, in order to study a particular phenomenon in the population. In this world, students may be aware of the variability between samples, but might not necessarily account for probabilistic considerations, e.g., the chance variability that stems from the random sampling process. In the *model world*, students build a model (a probability distribution) to a known (hypothetical) population and produce data of random samples from this model. Hence their attention is paid to a model and to a random process, which produces the outcomes of samples from this model. In this world, due to randomness, the details vary from sample to sample, but the variability is controlled. That is, given a certain distribution of the population, the likelihood of certain results can be estimated. Students in the IPA experiment transitions and build connections iteratively between the two worlds by working on the same problem context in both worlds. They examine sample results in relation to a hypothetical model of the population. Our hypothesis is that the IPA can support students’ development of reasoning about uncertainty when making ISIs. In this paper we focus on one task (“The Hidden Model of Social Networks” - HMSN) in the IPA learning trajectory that served as a scaffold for bringing the two worlds closer.

Participants

Yam and Shon, a pair of students (grade 6, aged 12), in an Israeli private school, were selected by their high communication skills to provide a window to their statistical reasoning. They have participated in the *Connections* experiment in fifth grade, collected and investigated data about their peers using *TinkerPlots1*. Following the growing samples heuristic (Ben-Zvi, Aridor, Makar, & Bakker, 2012), they were gradually introduced to samples of increasing size to support their reasoning about ISI and sampling.

The Learning Trajectory

The IPA learning trajectory encompassed six activities (total of about 20 hours) that first introduced the two worlds separately. In the data world, the students planned a statistical investigation: Chose a research theme, posed a question, formulated a conjecture, and decided about the sampling method and sample size (Act. 1-3). In the model world, they built a hypothetical model for the population distribution using the “sampler” (a TP2 object that creates models of probabilistic processes and generates random samples) based on their research conjecture (Act. 4).In order to encourage the students to examine the connections between the worlds, they were asked “what if” questions on optional real data results while exploring samples produced by model simulations. To refine students’ understanding of the connections between the two worlds, they were given a fifth Activity – HMSN, which is the focus of the current study. The students were asked to study a hidden TP2 sampler with unknown data distributions (built by two other students) by exploring random samples drawn from this sampler. In the last activity, the students returned to their own investigation (from Act. 1-4), explored data and models in the two worlds simultaneously by examining the real sample results in relation to their hypothetical model of the population.

The Episode

We focus here on Shon and Yam's discussions during one episode in the HMSN task, exploring samples drawn from the hiddenTP2samplerin order to make ISIs. In the beginning, the researcher presented the three interconnected attributes of the hidden sampler: Average time length that a teen spends on social networks (minutes per day), friends' number of a teen in social networks (#FSN), and grade. They were then asked to draw ISIs based on growing sample sizes, but each time they wanted to increase the sample size, they had to explain why. They began their exploration with sample size 10.

Data Collection and Analysis

The activities were entirely videotaped and computer screens were simultaneously captured using Camtasia. Interpretive micro-analysis (e.g., Meira 1998) was used to analyze the data. It is a microgenetic qualitative detailed analysis of the transcripts, which takes into account verbal, gestural, and symbolic actions within the situations in which they occurred. The goal of the analysis was to infer students' reasoning of uncertainty as they made ISIs based on random samples taken from the hidden model.

RESULTS

We identified four main stages in Shon and Yam's expressions of uncertainty: Account for, examine, control, and quantify uncertainty.

Stage I: Account for Uncertainty

Shon and Yam's initial understanding was that small samples are not good enough to draw valid conclusions: "We can't take [draw] more than ten [cases in a sample]. It's pretty bad" [31]. At the beginning, they were uncertain about the first sample data that were in contradiction to their previous knowledge. For example, Shon doubted that a fourth grade student has 600 FSN: "something doesn't make sense" [45]. Furthermore, they used a hesitant tone in describing the #FSN distributions of each grade ("it seems to me that…" [80]) and asked several times to increase the sample size.

They handled this uncertainty in two ways: a) added the #FSN mean to find a signal in the data, and b) drew a mean trend line (MTL) using theTP2's pen to recognize the pattern (sample 1 in Fig. 1). Restricted by sample size ten, Shon and Yam decided to draw more random samples from the hidden model, which only increased their uncertainty level: "It [sample 2] is completely different [than sample 1]" [147].Hence, they tried to characterize the variability between the samples.

Stage II. Examine Uncertainty

To examine their uncertainty about the variability between samples, Shon and Yam developed the "capture MTLs" method [155] to compare between samples. They kept the MTL they drew for each sample on the graph (Fig. 1) and noticed the large variability between the MTLs: "It is very different each time" [158], and concluded that "a sample size ten is too small" [160]. They consequently handled uncertainty by requesting to increase the sample size to 20.To examine their uncertainty about samples of 20, the boys used again the "capture MTLs" method and compared the MTLs for three samples (1-3in Fig. 2).

185	S	*So they [the MTLs 1-3, Fig. 2] get closer, as if they look very similar. One [MTL in sample size 10] was previously here [points to the left edge of Fig. 2] and one was there [points to the right edge of Fig.2].*
192	S	*They are more alike. Not very…*
193	Y	*[The MTLs are similar] not in terms of differences, but in terms of shape, yes. The shape [of the MTLs 1-3] increases till here [grades 4-5 or grades 6-7] and then it decreases.*
200	H	*But what does it mean about sample size 20in comparison to sample size 10?*
201	S	*It is better [to infer from]. It is a little better.*
202	Y	*There is still a problem: The peak [the largest mean] here [in MTL1, Fig. 2] is here [grades 6-7], but the peaks in these two [MTL2 and MTL3] are here [in grades 4-5].*

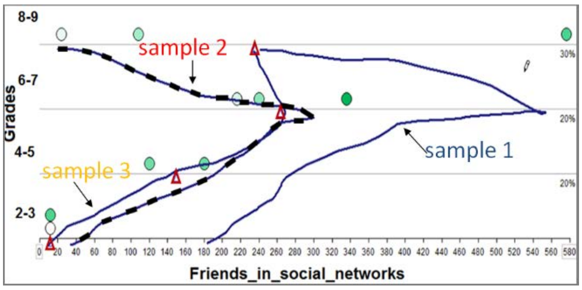

Fig.1: Comparison of the MTLs of #FSN over three samples of size 10.

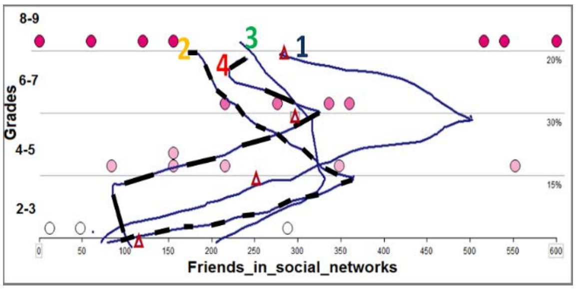

Fig. 2: Comparison of the MTLs of #FSN over four samples of size 20.

In their attempt to account for the variability between the samples, the boys referred to the similarity and differences in location, shape and "peak" between the MTLs (Fig. 2). These considerations resulted in reduced uncertainty level, although the boys were unsatisfied with sample size 20. A fourth sample (MTL4, Fig. 2) surprised them and destabilized their confidence regarding the trend. They decided therefore to disqualify the ability to infer from a sample size 20. The researcher's attempt to refine the boys' level of uncertainty about their informal inferences led them to notice different types of uncertainty by examining the variability in #FSN means within the grades over several samples.

Stage III. Control Uncertainty

The boys developed a new graphical method –"capture means"– to capture the variability between the means to control the uncertainty. They drew more samples size 20 and evaluated the variability of means within the grades by drawing circles to capture the means' signals (Fig. 3).

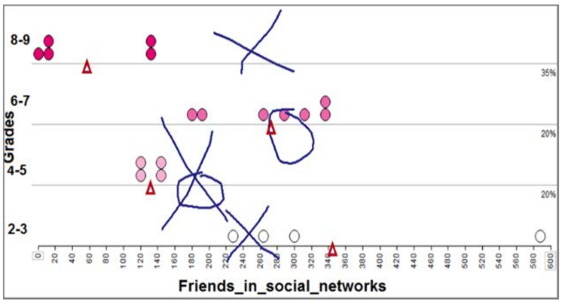

Fig.3: "Stable" and "constantly varying" mean signals over several samples size 20.

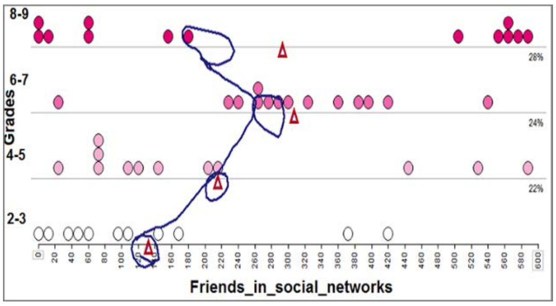

Fig.4: The hypothetical MTL of #FSN over several samples size 50.

236 Y *Grade 9 students – no [ability to capture the mean]. It [the mean of grade 9] constantly changes…Okay, it [the mean of grade 6] stays around here [draws a circle around grade 6means, Fig. 3].*

239 S *What can we do? Grades 6 and 7 are the most stable classes.*

The boys indicated that there was a difference in the variability of means between the grades over the repeated samples of size 20. They drew a circle to capture the mean's signal – allowing for a noise around it– in grades that are "stable" [239], and a 'X' in grades that their mean was "constantly varying" [236] (Fig.3). This method allowed them to decide whether the sample variability is too large within the grades. They used it for several samples of size 20 for each distribution to confirm and refine their findings [256].

246 S *In grade 6, I think it [the mean variability] is relatively stable because it [the mean] is usually in the area of the circle [Fig. 3]. That's why I say they [the sample variability] are relatively stable.*

251 Y *The mean is usually close to this circle [in grades 4 and 6]. And those marked [by] X [grades 2 and 9].Okay here [grade 4] also not really… [draw an X on the circle of grade 4, Fig. 3].*

257 H *Is your confidence level in the sample is connected to whether it is close or far?*

263 Y *We were able to score only this [the mean of grade 6] [nodding].*

According to their innovative "capture means" graphical method, when the means of a particular grade could be captured inside a drawn circle, they concluded that the variability within that grade was small [248]. Otherwise, the variability was too large, and they expressed higher uncertainty about conclusions that could be drawn [263] and about the sample size [265].

Stage IV. Quantify Uncertainty

Shon and Yam applied their "capture means" method on larger samples of size 50, drew a circle for each grade that captured well the means of that grade over many samples: "Grade 9 stays in this area. It really jumps around this spot [draws a circle around grade 4 mean]" [284]. Encouraged by their results, the boys expressed higher confidence level and were satisfied with the sample size: "[sample of] 50 in my opinion will be enough" [286]. Their certainty about the MTL increased and they drew it (Fig. 4) saying they were "absolutely" [290] certain.

During the next meeting, their sense of confidence encouraged them to refine their hypothetical MTL over samples of size 50. They drew a few random samples, but were surprised that several of them showed a significantly different trend ("type 1" in Fig.6) than the hypothetical trend ("type 0" in Fig. 5): "Here [in Fig. 5], it was an increase [from grade 2 to 6] and now [in the new sample, Fig.6] it is a decrease [from grade 2 to 4], [then an] increase [from grade 4 to 6]" [392]. Disappointed by the contradicting results, the boys became highly uncertain: "We can't draw an inference, because it is different all the time" [381]. They tried to handle the growing uncertainty about the trend by drawing bigger random samples of size 65 and noticed that there were more samples with "type 0" trend than "type 1". To quantify their uncertainty about the trend, the boys calculated the difference between the number of samples with each trend, and related to this difference as a "breakpoint". Namely, when this breakpoint equals a certain number, decided in advance, it would point at the more likely trend. Setting the breakpoint to three, the boys strengthened their previous assumption and rejected "type 1" trend over "type 0" with a subjective confidence level of 80%. They explained their high confidence level and even found a way to increase it:

463 Y *Because we had three times more [cases of "type 0" than "type 1"]. There are still times it's like this "type 1"], but most of the time it's like this ["type 0"].*

464 S *We will wait until it [the breakpoint] will be more than five. Here again… we'll wait until it will arrive at five…If there's one more time [a sample with "type 0" trend], then I believe in 90%.*

At this point they generalized the meaning of the breakpoint as an estimate of the confidence level: The bigger the breakpoint– the higher the confidence level is.

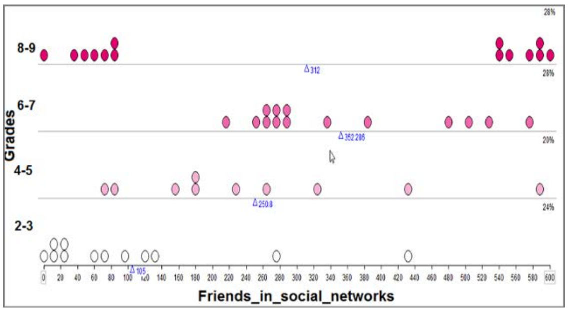

Fig. 5: The #FSN means within the grades over samples size 50 with trend type 0.

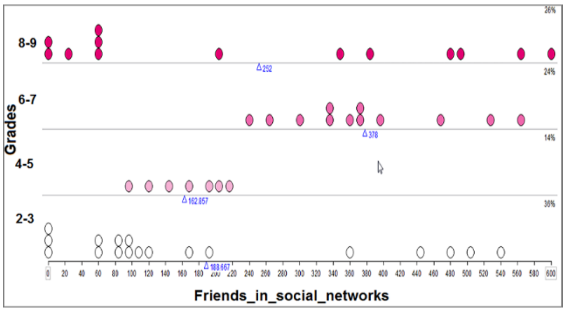

Fig. 6: The #FSN means within the grades over samples size 50 with trend type 1.

DISCUSSION

The question of this study was how can students' reasoning about uncertainty emerge *while making ISIs in an IPA learning environment?* In the above illustration we carefully examined Shon and Yam's work in the HMSN task, and found out four stages in their reasoning about uncertainty while learning to integrate the data and the model worlds. We describe in this section two types of uncertainty derived from the boys' expressions, the role they played in the transitions from stage to stage, and the study's implications.

We identified two types of uncertainty in the boys' expressions that shaped their reasoning about uncertainty in moving between stages: *contextual* and *statistical* uncertainty. The contextual uncertainty stemmed from a conflict between the boys' context knowledge and the data. For example, when the boys explored a sample size 10, Shon doubted that a fourth grade student had the biggest #FSN and said that "it is strange". Such a conflict increased the boys' uncertainty about the ability to infer from a sample and their context knowledge. The *statistical uncertainty* stemmed from the variability in the data and the sample variability. Disturbed by small sample sizes and restricted by the task design, the boys examined by self-invented graphical methods the variability between means and MTLS over many samples and the large variability in sample data. These situations raised the boys' uncertainty about inferring from a single sample of a certain size.

Salient conflicts between data and contextual knowledge that were expressed by contextual uncertainty drove the boys to refine their methods of examining, controlling and quantifying the statistical uncertainty in their transitions from stage to stage. In detail, exploring data that contradicted their previous knowledge in the first stage drove the boys to draw repeated samples and examine the statistical uncertainty in sampling variability in the second stage. In their transition from the second to the third stage, a surprising sample showing MTL that made no sense pushed the boys to control the statistical uncertainty by inventing the "capture means" method. A quantification of the statistical uncertainty in the fourth stage was a result of their contextual uncertainty regarding their hypothetical MTL.

Although this short description is far from exhaustive, this study strengthens our hypothesis that the IPA can support students' development of reasoning about uncertainty when making ISIs by experimenting with iterative transitions and building connections between the data and the model worlds. Authentic inquiry activities in the data world with clear purpose can raise the contextual uncertainty in students' reasoning. Probabilistic considerations in the model world can raise the statistical uncertainty in their reasoning. We believe that the boys' ability to handle both contextual and statistical uncertainties stemmed inter alia from their engagement in a learning trajectory based on the IPA. We hope to contribute to the important discussion in the statistics education community on new ways to combine data and chance, EDA and probability, in order to support students' informal inferential reasoning.

REFERENCES

Ben-Zvi, D., Aridor, K., Makar, K., & Bakker, A. (2012). Students' emergent articulations of uncertainty while making informal statistical inferences. *ZDM – The International Journal on Mathematics Education*, *44*(7), 913-925.

Ben-Zvi, D., Gil, E., & Apel, N. (2007, August). *What is hidden beyond the data? Helping young students to reason and argue about some wider universe*. Paper presented at SRTL-5, University of Warwick, UK.

Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher*, *32*(1), 9–13.

Garfield, J., & Ben-Zvi, D. (2008). *Developing students' statistical reasoning: connecting research and teaching practice*. Emeryville, CA: Springer.

Konold, C. (1989). Informal conceptions of probability. *Cognition and Instruction*, *6*(1), 59-98.

Konold, C., & Miller, C. (2011). *TinkerPlots* (Version 2.0) [Computer software]. Key Curriculum Press.

Makar, K., & Rubin, A. (2009). A framework for thinking about informal statistical inference. *Statistics Education Research Journal*, *8*(1), 82-105.

Makar, K., Bakker, A., & Ben-Zvi, D. (2011). The reasoning behind informal statistical inference. *Mathematical Thinking and Learning*, *13*(1&2), 152-173.

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.

Moore, D. (2007). *The basic practice of statistics* (fourth ed.). New York: W. H. Freeman and Company.

Pratt, D., & Ainley, J. (2008). Introducing the special issue on informal inferential reasoning. *Statistics Education Research Journal*, *7*(2), 3-4.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124-1131.