International Journal of Statistics and Probability; Vol. 2, No. 1; 2013
ISSN 1927-7032 E-ISSN 1927-7040
Published by Canadian Center of Science and Education

Estimating Parameters from Samples: Shuttling between Spheres

Theodosia Prodromou[1]

[1] School of Education, University of New England, Australia

Correspondence: Theodosia Prodromou, School of Education, University of New England, Australia. Tel: 61-2-6773-3237. E-mail: theodosia.prodromou@une.edu.au

Received: December 31, 2012 Accepted: January 24, 2013 Online Published: January 27, 2013

doi:10.5539/ijsp.v2n1p113 URL: http://dx.doi.org/10.5539/ijsp.v2n1p113

Abstract

In order to better understand the thinking of students' learning to make informal statistical inferences, this research examined the thinking of senior secondary school students (age 17) engaged in the task of using observed data to make point estimates of a population parameter within a computer-based simulation. Following the "Growing Samples" instructional model, the point estimation activity involved sampling and estimating across three tasks with different sample sizes. This research study aimed to trace the evolution of the students' thinking, with particular attention to use of the statistical concepts in making informal inferences from sampling. The students in this study were observed to rely primarily on mathematical thinking, which, perhaps, inhibited their ability to construct meanings about the basic statistical concepts underpinning sampling when performing point estimates. At times in the process students were seen to shift between mathematical thinking, statistical thinking, and thinking about the context, but the mathematical thinking seemed to dominate their attempts to create estimates. These research findings are useful for informing the teaching of point estimation of a population parameter to school-aged students. The research findings also stress the need for teachers to rethink the relationship between statistical thinking and mathematical thinking in order to promote statistical thinking in relevant learning situations for their students.

Keywords: informal statistical inference, point estimates, population, samples, statistical thinking, mathematical thinking

1. Introduction

A productive and authentic way of teaching the statistical reasoning necessary when working with samples is to provide opportunities for school students to engage in activities that involve informal inferential reasoning (Makar, Wells, & Allmond, 2011). Such activities also provide an important way for students to progress from working with descriptive statistics to working with inferential statistics because they offer the opportunity to reason informally. The word "informally" is used to "emphasize that we are not expecting students to rely on formal statistical measures and procedures to formulate their inferences" (Makar, Bakker, & Ben-Zvi, 2011, p. 153). Such Informal Inferential Reasoning (IIR) has been defined as the process of drawing generalised conclusions from data, involving four critical principles: generalising beyond data (parameter estimates, conclusions, and predictions); using data as evidence of the generalisation; articulating the degree of certainty (due to variability) embedded in the generalisation (these three principles were articulated by Makar and Rubin, 2009); and comparing datasets with a model such as ideal (targeted) distributions (proposed by Bakker, Kent, Derry, Noss, and Hoyles, 2008).

The research reported in this paper focuses on the first of these four principles, generalising beyond data, in particular estimating parameters. Estimating parameters is a process by which one makes inferences about a population based on information gained from one (or more) sample(s). A sample is a representative part of a population selected when sampling for the purpose of drawing inferences about unknown populations (e.g., estimating parameters or predicting).

2. Growing Samples

Recent research has sought to understand how better to approach the topic of making informal inferences about a population based on information gained from one or more samples (Ben-Zvi et al., 2011, 2012; Prodromou, 2011).

The instructional idea of "Growing Samples", suggested by Konold and Pollatsek (2002) and then developed by Bakker (2004), plays a predominant role in providing a useful perspective of the role in shedding some light in the development of students' informal inferential reasoning, and reasoning about sampling, samples, and variation.

Bakker helped eighth grade students who engaged with a sequence of "Growing Samples" activities to see stable patterns generated by larger samples, thus students understood that larger samples are less variable and better represent population. Bakker suggested that asking students to make conjectures about the growing samples builds students' reasoning about sampling in the context of variability and distribution. Research literature provides evidence that the growing samples approach is helpful in supporting coherent reasoning, based exclusively on the integration of key statistical concepts such as sampling, data, distribution, variability, and tendency (Ben-Zvi et al., 2011, 2012; Prodromou, 2011).

Ben-Zvi (2006) found that the growing samples processes enhanced students' sensitivity to uncertainty and variation in data, enabling students to know something about the population. Research studies by Ben-Zvi et al. (2011) and Prodromou (2011, 2012), which were in line with the Growing Samples literature, showed that students developed inferential reasoning about sampling while working with *TinkerPlots*. The students in those studies not only experienced the limitations of small samples when making inferences about a larger population, but also experienced an emerging quantification of confidence in making such inferences, interconnections of concepts of sampling, and informal statistical inference with key concepts such as spread, distribution, likelihood, randomness, average, and graph interpretation.

When the students were encouraged to express their confidence about how certain they were about their inferences, they tended to either express extreme confidence in knowing that something can be inferred from samples or express that nothing could be concluded (i.e., complete certainty vs. extreme doubt; Ben-Zvi et al., 2012). The growing samples task design provided students with opportunities to witness increasing evidence for (or against) particular conjectures, thus develop a language to talk about "grey areas of this middle ground" (p. 923).The research reported in this paper is based on the growing samples design of the investigations that support students' informal inferential reasoning when estimating population parameters from samples.

3. Using Samples to Estimate Parameters as Part of Informal Inferential Reasoning

Parameters can be estimated by providing either a point estimate or an interval estimate. A point estimate involves the use of sample data to calculate a single value (best known as a statistic) that can be used as a "best guess" or "best estimate" of an unknown (fixed or random) population parameter. For example, a sample mean is a point estimate used to estimate the population mean. An "interval estimate" involves the use of sample data to calculate an interval of possible (or probable) values of an unknown population parameter within which a population parameter lies. For example, 1< sample mean < 4 is an interval estimate within which the population mean lies.

The student reasoning process that leads to informal statistical inferences (ISI) when estimating parameters can help teachers to gain insights in student thinking and identify critical elements that support and nurture student ISI such as estimating parameters. Statistical thinking is needed when students engage in informal inferential reasoning and teachers need to be familiar with this type of thinking and to nurture it for students to be supported in their informal inferential reasoning.

4. Tension between Statistical Thinking and Mathematical Thinking

Mathematics teachers need to be aware that the thinking a student requires to solve a statistical problem will differ from the thinking required to solve most mathematical problems. If students are not equipped with sufficient statistical thinking capabilities they may approach statistical problems using mathematical thinking. So how might statistical thinking and mathematical thinking differ?

What is 'thinking'? Generally speaking, thinking can be defined as "the process of considering or reasoning about something" (Oxford University Press, 2012). While this definition provides a basic guide to the concept of thinking, more detailed explanations have been provided that are discipline specific, of these the statistical and mathematical thinking are of most interest. Before considering these in more detail, what of the distinction between thinking and reasoning?

In statistics education some have described reasoning as a form of thinking with both reasoning and other forms of thinking needed to be able to work on a task, while others have attempted to make a clear distinction reasoning and thinking. A useful approach to distinguishing between reasoning and thinking is to consider the task being undertaken and conceptualise thinking as *knowing* "when and how to apply knowledge and procedures", and reasoning as *explaining* "why results were produced or why a conclusion was justified" (delMas, 2004, p. 85).

Thus examples of reasoning can be found in particular stages of a person's thinking, such as where the person is expected to imply, justify, or infer. Now back to the two types of thinking, mathematical and statistical.

What is mathematical thinking? Mason, Burton, and Stacey (2010) described four fundamental processes involved in mathematical thinking: (MT1) specialising-considering special cases or examples; (MT2) generalising-looking for patterns and relationships; (MT3) conjecturing-predicting relationships and results; and (MT4) convincing-finding and communicating reasons why something is true. From the previous discussion it might be concluded that convincing (MT4) is mathematical thinking that involves "reasoning".

What is statistical thinking? In attempting to answer this question, a statistician and a mathematics educator (Wild & Pfannkuch, 1999) worked together to build up four dimensions which contribute to the "rich complexity" of statistical thinking: (ST1) the investigative cycle-continuously through the stages problem, plan, data, analysis and conclusion; (ST2) types of thinking-recognition of need for data, transnumeration, consideration of variation, reasoning with distinctive set of statistical models, integrating the statistical and contextual information, knowledge, and conceptions; (ST3) the interrogative cycle-continuously through the stages generate, seek, interpret, criticise and judge; and (ST4) dispositions-including scepticism, imagination, curiosity and awareness, openness to ideas that challenge preconceptions, a propensity to seek deeper meaning, being logical, engagement and perseverance.

Amongst the types of thinking skills (ST2), Wild and Pfannkuch recognised, in particular, the importance of the raw materials on which statistical thinking works. These raw materials are statistical knowledge, context knowledge, and the information in data. However, the thinking itself occurs by the synthesis of these elements. In particular, one has to bring to bear all appropriate knowledge regarding the undertaken task, and then to build connections amongst existing context-knowledge and the outcomes of statistical analyses. Wild and Pfannkuch (1999) described the synthesis of context-knowledge and statistical knowledge as one that "traces the (usual) evolution of an idea from the earliest inkling through to the formulation of a statistical question precise enough to be answered by the collection of data, and then on to a plan of action" (p. 228). They also emphasize the continual shuttling backwards and forwards between thinking in the context sphere and the statistical sphere. The interplay between context and statistics is continuous until the questions in hand are satisfactorily answered. For example, Wild and Pfannkuch (1999) explain how, in the analysis stage, context knowledge leads to questions that require consultation of the observation data, which pushes learners into the statistical sphere of thinking, but then characteristics of the data push learners back to the context sphere to answer basic questions like, "Why is this happening?", and "What does this mean?" (p. 228).

While there may be similarities between the mathematical thinking and statistical thinking, these two types of thinking are dissimilar in two important elements: variation and context. All statistical thinking must be grounded within a context (delMas, 2004), while mathematical thinking may or may not make use of contexts. All statistical thinking involves some form of consideration of variation (Pfannkuch & Wild, 2004), which is very different from the concept of variables dealt with in mathematical thinking. The fact that variation is an observable phenomena and that it is always present (Wild & Pfannkuch, 1999) is of relevance to all aspects of statistical thinking.

To appreciate the tension between thinking mathematically or statistically, consider research reported by Lane-Getaz (2006) where students engaged in simulation activities were good at "mathematically" calculating statistics but once they were exposed to activities that allowed them to explore variation within distributions they were able to produce explanations of their projects that demonstrated better statistical thinking.

Rather than promoting the differences between the two types of thinking, teachers who rethink the relationship between statistical thinking and mathematical thinking can help their students to learn how to synthesize the two types of thinking. This would help teachers to promote statistical thinking in relevant learning situations for their students. To inform teachers in developing such support, the researcher became interested in what type of thinking students will engage in when completing an activity that requires statistical thinking.

5. Aim

This exploratory research study examined how senior secondary school students construct meanings about basic statistical concepts underpinning sampling when making informal inferences from data. The focus was on observing the development of students' thinking as they construct meaning about the key statistical concepts of 'sample' and 'sampling,' while the students engaged in an informal statistical inference task that involved making point estimates of a population parameter within a computer-based simulation. It was expected that some insights might be gained into the conceptual struggle that takes place when 17-year-olds engage in inferential reasoning when making point estimates of a population parameter.

In this research a constructivist stance is used to search for nave conceptions that might serve as resources in developing more sophisticated strategies. In addition, this might shed some light on the tension that a student may experience when opting for thinking mathematically or statistically and how this tension can be resolved.

6. Point Estimation Activity

The point estimation activity, a computer simulation titled Murphy's Dam was presented in a spreadsheet and introduced the context of a dam containing three fish species (Bass, Perch, Trout; Figure 1). The spreadsheet allowed students to simulate drawing a sample of fish from the dam and displayed the number and percentage of each species of fish within the sample (Figure 2). The students were asked to provide the owner of the dam, Brian Murphy, with advice about the percentage of each species of fish in his dam. To inform their advice, the students were engaged in drawing "catch and release" samples of 20, 50, and then 100 fish from the dam (Tasks 1, 2, and 3, respectively). In each task students were requested to make estimates of the percentage of bass, perch and trout in the dam after each sample was drawn.

Brian Murphy has a dam, on his farm, which contains many fish of three different species: Bass, Perch and Trout. Since introducing each of the three species the number of fish has grown considerably. Brian would like to estimate the percentage of each species he now has in the dam.
You have been consulted to provide this advice to Brian. Your estimate of the percentage of each of the three species will be based on a sample of fish that you draw (catch and release) from the dam.
Your sampling of fish will be based on the assumption that you are drawing the fish in such a way that each fish (no matter what species) is equally likely to be caught.

Figure 1. Murphy's dam scenario

Draw a sample of 20 fish			
species	Bass	Perch	Trout
number	2	15	3
percentage (%)	10	75	15

Figure 2. Simulated sample of 20 fish

For each task students were required to draw ten separate samples from the dam and for each sample drawn record the observed percentage for each species of fish caught and a point estimate of percentage of bass, perch and trout (Figure 3).

Fish caught from the dam						Estimates of fish in the dam		
Bass (%)		Perch (%)		Trout (%)		Bass (%)	Perch (%)	Trout (%)
7	35	10	50	3	15	30	40	30
5	25	9	45	6	30	20	40	40
4	20	13	65	3	15	15	70	15
2	10	9	45	9	45	5	50	45
3	15	15	75	2	10	20	70	10
9	45	11	55	0	0	50	50	0
5	25	11	55	4	20	20	60	20
3	15	14	70	3	15	20	65	15
9	45	8	40	3	15	50	35	15
4	20	14	70	2	10	15	75	10

Figure 3. Example of completed recording sheet (second iteration of Task 1)

When the three tasks were completed, the students were asked to reflect on the point estimation activities across the three tasks, reason about their estimates by comparing the estimates, and attempt to estimate the actual percentage of each species of fish in the dam (In the simulation, the percentage of fish in the dam was set as 30% bass, 50% perch and 20% trout).

The design of the point estimation activity evolved around the idea of growing samples, starting from a sample of size 20, moving to about 50, then 100, and finally the entire population. Using a sequence of "growing sample" activities was a pedagogical design conjecture to help students progressively develop their inferential reasoning about samples, and their ability to make point estimates of parameters of the population.

7. Methodology

7.1 Participants

The point estimation activity was undertaken by three pairs of average-ability female students studying Mathematics General (Year 11-age 17 years) in an Australian secondary school. Participation was voluntary; students self-selected a partner; and the tasks were undertaken out of class time. The teacher made the final choice of which students participated by recommending those who were able to better articulate their thinking. The choice of senior secondary students was a curriculum-based decision because, in accordance with Australian curriculum guidelines (Australian Curriculum, Assessment and Reporting Authority [ACARA], 2011), Year 11 students in this school had been taught sampling prior to involvement in the research. The choice of average ability students was based on the assumption that above-average ability students would construct the knowledge too quickly and possibly not take the time to verbalise their thinking, while below-average ability students may not be able to articulate their thinking.

The researcher was a participant observer, working with each pair of students as they completed the three tasks. The researcher interacted with the students to probe for reasons that might help to explain their actions and therefore provide some insight in their thinking.

7.2 Data Analysis

Student work with the simulation was recorded using Camtasia (2000) software. The data collected included audio recordings of the student voices, video recordings of the screen activity, and worksheets completed by the students.

All the audio recordings were transcribed and screenshots of the simulation spreadsheet and student recording sheets were included as needed to make sense of the transcription. The researchers discussed the data and chose those sections of the transcript that most clearly demonstrated student thinking as they reasoned.

8. Findings

The findings are only reported for one pair of girls, Cathy and Liz (pseudonyms), because their articulation during the three tasks gives the most informative illustration of their thinking. When the two girls first engaged with the three tasks they simplified the required approach by estimating after all ten samples had been drawn rather than after each sample. A second iteration of the three tasks was undertaken to achieve what was originally planned for the point estimation activity, having the girls estimate after each sample was drawn. The findings are reported for each of the two iterations separately.

8.1 First Iteration of the Three Tasks

When Cathy and Liz were engaged in Task 1 (drawing samples of size 20), they began working on separate recording sheets (as instructed) but insisted on drawing all ten samples (contrary to instructions) from the dam before estimating the percentage for each species of fish. They made two requests: (i) to share one recording sheet rather than work on two separate sheets; and (ii) to record the number of fish as well as the percentage of fish for each species. Both requests were allowed. When making their estimate (after the ten samples were drawn) they insisted on trying to develop an algorithm to calculate the estimate. They took the average number (over the 10 samples) of bass caught (5) and divided this by the total number of fish caught (20) and converted it to a percentage (25%). They repeated this process to calculate the percentage of Perch (50%) and the percentage of Trout (25%). The working was done using a calculator. When they were asked to explain the algorithm they developed to make the estimate they wrote "number of fish caught sample of fish x 100" (Algorithm A). Using an algorithm like this is an example of mathematical thinking. They did not explain why they averaged the number of fish caught over all ten samples and then calculated the percentage, rather than just averaging the percentages.

When Cathy and Liz engaged in Task 2 (drawing samples of size 50), they tried to apply algorithm A to calculate the estimate of the percentage of bass. They drew the ten samples and performed their calculations, this time using a spreadsheet. The average of the ten samples came out to 14.8 but they chose to use 14 instead of 14.8 in their algorithm. In addition, they said they did not "like" the answer they were getting to estimate the percentage of bass because it was too small and they chose 20 instead. It is not clear how or why the girls chose 20% as their estimate. No percentages were calculated for perch and trout and the girls did not explain why they did not estimate the percentages of perch and trout.

When Cathy and Liz engaged in Task 3 (drawing a sample of size 100), they realised that in the observed data for each sample the percentage of each species of fish was equal to the number of fish caught. They again applied their algorithm to calculate the estimates. They wrote out the estimate for bass, and then argued that the estimates for the other two species of fish would simply equal the percentages they had drawn in their samples. As explained above the girls developed an algorithm to form their estimates and their explanations were computational rather than statistical in nature. The researchers decided to do a second iteration of the three tasks, this time insisting that the girls estimate the percentage of each species of fish after each sample drawn as was planned in the original task. The second iteration of the three tasks was performed two weeks after the first iteration.

8.2 Second Iteration of the Three Tasks

When Cathy and Liz engaged in Task 1 they experimented with a mathematical algorithm: "number of a species of fish divided by percentage of a species of fish multiplied by 100" (Algorithm B). For example, after drawing the first sample, they calculated 5 divided by 25 multiplied by 100 to give an estimate of 20 for the percentage of bass. Similarly for perch, 9 divided by 45 multiplied by 100 gave 20, and even for trout they calculated 20 as the estimate. Although they did not express surprise that all three estimates were the same, they did realise that the three percentages should sum to 100 (conservation principle) and as the sum was only 60, they concluded that that their algorithm did not work.

They then calculated the estimated percentage of each species using an alternate algorithm: "percentage of a species of fish divided by the number of species of fish multiply by 100" (Algorithm C). This calculation resulted in 50% as the estimate for each of bass, perch, and trout. This time they concluded that their algorithm did not work because they calculated the same estimate, 50, for each species of fish.

Not satisfied with either attempt at estimation, they suggested drawing another sample of 20 fish so that they could compare the new sample and the previous sample to observe any possible change. Cathy and Liz used the new and previous observed percentage caught to produce a new estimate, which was not related to the previous estimate. If the percentage of fish caught had gone up (down), then the estimated percentage of fish was set at 5% more (less) than the caught amount (Algorithm D). This algorithm was used for both bass and perch. Then the estimated percentage of trout was always calculated by adding the bass and perch estimates together and subtracting from 100. Algorithm D was the first attempt by the girls to provide estimates that in some way were linked to the fluctuations (variations) between the samples drawn but made no use of the previous estimate to calculate a new estimate. Obviously this algorithm only worked for the second or subsequent estimates, otherwise the change direct, up or down, could not be determined. The increments were always only 5%, up or down, irrespective of the size of the percentage of fish for the new and previous sample.

An example of the application of Algorithm D from the 8th estimates and 9th estimates of Task 2 (Figure 4) follows. To find the 9th estimate for bass Cathy and Liz noticed that the 9th catch (observed percentage) for bass, 45, was larger than the 8th, 15, and so the 9th estimate for bass was the observed value (45) plus 5, giving 50 as the estimated percentage of bass. Similarly for the 9th estimate for perch, 40 (new observed percentage) was less than 70 (previous observed percentage) and so 5 was subtracted from the observed percentage (40) to give 35 as the estimated percentage of perch. Finally, the 9th estimated percentage of trout was 15 was calculated as 100-(50+35). It should be noted that the 8th estimate (20 65 15) was not used at all by the students in producing the 9th estimate Although this application of Algorithm D has been explained using 8th and 9th catches, this algorithm was applied for calculating the second and subsequent estimates.

When Cathy and Liz were engaged in drawing "catch and release" samples of size 50 (Task 2), from the dam and estimating the percentage of bass, perch, and trout in the dam (see student recording sheet in Figure 5), they needed to form an estimate for the first sample drawn because Algorithm D could not be applied. Cathy tried Algorithm B, which resulted in 50 for each of the three percentage estimates, bass, perch, and trout. The girls realised that the estimate cannot be 50 every time, concluding that the algorithm did not work.

They tried to find a number that went into 28, 50, and 22 (the percentages caught in the first sample drawn). They concluded that 2 goes into 22, 50 and 28 and thus 2 was the number they could increase or decrease the observed percentage by the work out the percentage estimate.

Fish caught from the dam		
Bass (%)	Perch (%)	Trout (%)
15	70	15
45	40	15

Estimates of fish in the dam		
Bass (%)	Perch (%)	Trout (%)
20	65	15
50	35	15

Figure 4. Example of the application of Algorithm D to estimate the percentages of each species of fish

Fish caught from the dam

Bass (%)		Perch (%)		Trout (%)	
14	28	25	50	11	22
17	34	22	44	11	22
14	28	27	54	9	18
11	22	31	62	8	16
23	46	22	44	5	10
16	32	21	42	13	26
15	30	23	46	12	24
12	24	27	54	11	22
20	40	21	42	9	18
16	52	24	48	10	20

Estimates of fish in the dam

Bass (%)	Perch (%)	Trout (%)
24	54	22
38	40	22
24	58	18
18	66	16
50	40	10
28	46	26
26	50	24
20	58	22
44	38	18
28	52	20

Figure 5. Student recording sheet for Task 2 in the second iteration. Each row represents a trial

The following comes from a point where they tried to make an estimate after the first sample was drawn.

1) Cathy: Maybe if we would work out like the next fish (draw another sample) because then we could see if there is a pattern which might help with the formula.

2) Researcher: Remember that you need to make estimates after each sample was drawn.

3) Liz: maybe if we could work out how many fish equalled what percentages like we did for Task 1 (sample size 20). 1 fish equalled 5% and we will work things out.

4) Cathy: Well it's always half. 1 fish equals 2%, 2 fish equals 4%, 4 fish equals 8%.

They had decided confidently that 2 was the number they could increase or decrease the observed percentage by the work out the percentage estimate.

The girls used then a new algorithm (Algorithm E), without discussing this algorithm. They went down by 2 fish for the Bass and so took away 4% from the "caught" percentage (28%) of Bass to give the "estimate" percentage (24%) of Bass. They then added 4% to the percentage of Perch. Then they found the percentage of trout by adding the percentage of bass and perch and subtracting the sum from 100%. They estimated [24, 54, 22] (See student recording sheet in Figure 5).

The girls tried to use Algorithm D developed for Task 1 (sample size 20) to make the second and subsequent estimates. However, when applying the algorithm this time the relevant change, either up or down, was only by 2% and not 5%. As before, the estimated percentage of trout was calculated using the conservation principle.

When Cathy and Liz were engaged in drawing "catch and release" samples of size 100 (Task 3), they applied Algorithm E, for the first estimate (see student recording sheet in Figure 6). However, they went down by 1 fish for the Bass and so took away 1% from the "caught" percentage (26%) of Bass to give the "estimate" percentage (25%) of Bass. They then added 1% to the percentage (58%) of Perch to give the estimated percentage (59%). Then they found the percentage of trout by adding the percentage of bass and perch and subtracting the sum from 100%.

Fish caught from the dam

Bass (%)		Perch (%)		Trout (%)	
26	26	58	58	16	16
29	29	46	46	25	25
33	35	51	"	16	"
30		44		26	

Estimates of fish in the dam

Bass (%)	Perch (%)	Trout (%)
25	59	16
30	45	25
34	50	16
29	45	26

Figure 6. Student recording sheet for Task 3 in the second iteration

They drew three samples and made three estimates based on Algorithm D, to make the second and subsequent estimates. They relevant change was either up or down, by 1%. Cathy pointed out that the algorithm was working "even" when the number of fish caught was 100 because the number of fish equals the percentage and so 1 fish equalled 1%. As before, the estimated percentage of trout was calculated using the conservation principle. They then concluded that they did not need to draw any more samples because it was "easy" to estimate.

When the girls were asked to reflect on all the activities when they "caught and released" samples of sizes 20, 50 and 100 and attempt estimate the actual percentage of each species of fish in the dam. They averaged the percentage of the fish caught when 100 fish caught from the dam. The researcher asked them to estimate the percentage without using a formula. Then they looked at the range of the percentage caught for a particular species across the three activities (see student recording sheets in Figure 3, Figure 5, and Figure 6), e.g., for bass 26 was the least and 33 was the most and then they found the median "middle value" (33-26=7) and then halved it so that the middle value was 3.5. They concluded that it could be 29 or 30 but then stayed with 30.

9. Summary

The conclusions drawn indicate when making point estimates of the percentage of bass, perch, and trout in the dam, the two secondary students, Cathy and Liz, generally focused on whether the percentage of fish caught had gone up (or down), then the estimated percentage of fish was set at 5% (samples of size 20), or 2% (samples of size 50), or 1% (samples of size 100) more (less) than the caught amount.

A number of algorithms emerged and as part of this process the students tended to experiment with choosing a "relevant" algorithm each time a new sample was drawn.

Despite the apparently arbitrary choice of the three following algorithms to forms estimates of the percentage of bass, perch and trout in the dam:

(1) "number of fish caught divided by sample of fish multiplied by 100" (Algorithm A).

(2) "number of a species of fish divided by percentage of a species of fish multiplied by 100" (Algorithm B)

(3) "percentage of a species of fish divided by the number of species of fish multiply by 100" (Algorithm C).

When Cathy and Liz engaged in Task 2 (drawing sample of size 50) and Task 3 (drawing sample of size 100), they tried to apply algorithm E to make the first estimate:

(4) "go down by 2 fish when drawing a sample of size 50 (or 1 fish drawing a sample of size 100) for the Bass and so take away $2 \cdot 2\% = 4\%$ (or 1%) from the "caught" percentage of Bass to give the "estimate" percentage of Bass. They then added 4% to the percentage of Perch. Then they found the percentage of trout by adding the percentage of bass and perch and subtracting the sum from 100%" (Algorithm E).

The girls tried to use Algorithm D to make the second and subsequent estimates:

(5) "the relevant change, either up or down, was by 5% (sample of size 20) or 2% (sample of size 50) or 1% (sample of size 100). The estimated percentage of trout was calculated using the conservation principle" (Algorithm D).

Eventually, students appeared to settle on Algorithm E when making the first point estimates of a population parameter and on algorithm D when making the second and subsequent estimates.

Whilst making the point estimates, the students demonstrated very little evidence of the notion of stabilizing the values used for the points estimate as more information becomes available with each successive sample. So, despite looking back to previous estimates to form the new estimate, there was no sense of refining previous estimates, just a sense of using them as a basis for the new estimate.

In this paper, students' engagement in the point estimation activity has been presented in which mathematical thinking dominates over statistical thinking.

Students did not construct any meanings about basic statistical concepts underpinning sampling in the context of making informal inferences from data when performing point estimates.

10. Discussion

For the contribution to the topic of estimating parameters from samples when engaged in an informal statistical inference task that involved making point estimates of a population parameter within a computer-based simulation, the research findings are useful for informing the teaching of point estimation of a population parameter to school-aged students. The type of thinking students engaged in when completing the point estimation activity that requires statistical thinking became the focus of the research findings. One has to bring to bear all relevant knowledge on the tasks in hand, and then to draw connections between existing context-knowledge, mathematical knowledge and the previous estimates of parameters from samples. Wild and Pfannkuch (1999) developed a theoretical framework that illustrates the Interplay between context and statistics and also emphasizes the continual shuttling backwards and forwards between thinking in the context sphere and the statistical sphere.

In this study, the findings show that the students persistence to think mathematically rather than statistically prevent them from constructing any meanings about basic statistical concepts underpinning sampling in the context of making informal inferences from data when performing point estimates. The research findings stress the need to rethink the relationship between statistical thinking and mathematical thinking.

The researcher considers that statistical knowledge, mathematical knowledge, and context knowledge are the raw materials on which thinking works. The thinking required for estimating the percentage of the population parameters is in fact a synthesis of these elements to produce implications, insights and conjectures. One cannot indulge in statistical thinking without having some context knowledge and mathematical knowledge.

The researcher based on Wild's and Pfannkuch' (1999) constructed a new framework that illustrates the construction of students' knowledge when engaged in an informal statistical inference task that involved making point estimates of population parameters from samples within a computer-based simulation. Figure 7 illustrates the continual shuttling backwards and forwards between thinking in the context sphere, the statistical sphere, and the mathematical sphere.

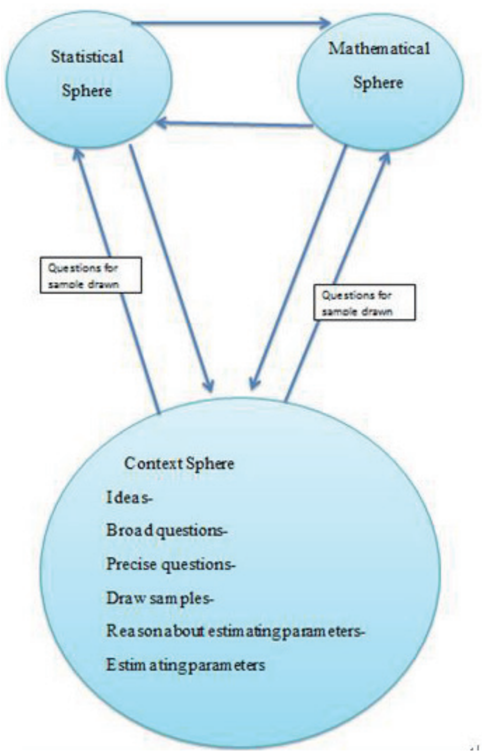

Figure 7. Shuttling between spheres

Figure 7 traces the evolution of an idea from the earliest inkling through to the formulation of a precise statistical question that is to be answered following the different stages of the "growing samples" instructional idea. The role of the context is crucial at the earliest stages of the statistical task because learners' constructions of informal statistical inference understandings are driven almost entirely by context knowledge.

For example, at the parameter estimating stage questions are suggested by context knowledge that require consulting the samples drawn, which temporarily pushes learners into the mathematical sphere because learners are more familiar with constructing and working using mathematical formulae. Features seen in samples propel learner back to the context sphere to answer scaffolding questions: "Why is this happening?" and "What does this mean?" Such scaffolding questions are meant to push learners into the statistical sphere to help them construct statistical concepts underpinning their activities. Statistical knowledge contributes more as the thinking crystallises.

In this study, students were not able to construct meanings about basic statistical concepts underpinning sampling in the context of making informal inferences from data when performing point estimates because the scaffolding questions did not push leaners into the statistical sphere.

It might be argued that students' pursuit of mathematical thinking at the earliest stages of learning statistics has been the root of all the problems that have been encountered by students being taught elementary statistical concepts. At these earliest stages, the reduction of emphasis on mathematical thinking in statistical teaching is crucial because the earliest stages are driven almost entirely by context knowledge and learners' thinking is moving backwards and forwards between thinking in the context sphere and the statistical sphere. The statistical concepts learners construct and reason about are usually "informal" and embrace elements of "naivety".

A fundamental change in teaching elementary statistics at school is required in order to allow and facilitate the simplification of core statistical ideas. In particular, the key ideas of statistical inference can be developed without relying on any mathematical models used by formal probability theory. Statistical thinking at elementary levels can be developed relying on processes and not mathematics.

On the contrary, for constructing statistical knowledge beyond the elementary level, the use of mathematical formulae of probability theory becomes increasingly important, as mathematics is critical for the development of statistical methods.

The reader must consider the limitations of this research to elaborate the research questions. One limitation is that the researcher only analysed the reasoning of one pair of students, and focused on the more interesting illustrations of the emerging ideas. While the overall analysis of the other pairs of students followed the same broad strokes, some interesting variations used by the students in their reasoning might justify further discussion, but are outside the scope of this paper.

A second limitation is the interview technique used to ask students to explain their responses was not explicit enough. Despite students being asked "why" they had formed the estimates the way they did and "why" they made any changes to their estimates, they mostly gave superficial responses in their reasoning. More probing questions were needed to direct the students to explain their reasoning and propel students back to the context sphere to reflect on context knowledge.

10.1 Future Research

Despite the limitations of the study, it reveals some aspects of students' reasoning while making point estimates. Although some interesting point estimation algorithms emerged, little evidence of using the core concepts, sampling, sample, and variation.

The results of this research provide a strong basis to help those teaching point estimation to secondary school students better understand their students' reasoning. As far as further research is concerned, in raising the notion that there may be a better way to investigate statistical reasoning, especially as involved in the sampling process, it is acknowledged that the focus should not be on mathematical approaches to estimation.

There is still more research that needs to be done in exploring students' statistical reasoning when sampling and making point estimates and many important research questions exist that the above research has not addressed. There are two areas that should be the focus for future researchers. The first area is a need for research that also studies students' quantification of the level of confidence (Ben-Zvi et al., 2011; Prodromou, 2011) when engaged in an informal statistical inference task that involved making point estimates of a population parameter within a computer-based simulation.

The estimation tasks used in the reported study could be expanded to include student expression of their level of confidence in relation to their informal inferential reasoning while sampling. This could be achieved, for example, by letting the students sample until they are confident that they have a “good” estimate (i.e., the students decide when to stop sampling), rather than instructing them to do a specified number of trials.

The second area is a need to investigate the relationship between statistical thinking and mathematical thinking in order to promote statistical thinking in relevant learning situations for students being the outcome of a balanced synthesis of ideas and information from the context sphere, statistics sphere and mathematics sphere.

References

Australian Curriculum, Assessment and Reporting Authority. (2011). *Australian Curriculum: Mathematics*. Version 1.2. Retrieved March 15, 2011, from http://www.acara.edu.au

Bakker, A. (2004). Reasoning about shape as a pattern in variability. *Statistics Education Research Journal, 3*(2), 64-83.

Bakker, A., Kent, P., Derry, J., Noss, R., & Hoyles, C. (2008). Statistical inference at work: Statistical process control as an example. *Statistics Education Research Journal, 7*(2), 131-146.

Ben-Zvi, D. (2006). Scaffolding students’ informal inference and argumentation. In A. Rossman & B. Chance (Eds.), *Proceedings of the Seventh International Conference on Teaching of Statistics*, Salvador, Bahia, Brazil, 2-7 July, 2006. Voorburg, The Netherlands: International Statistical Institute. Retrieved from http://www.stat.auckland.ac.nz/*iase/publications/17/2D1_BENZ.pdf (Accessed 10 December, 2012)

Ben-Zvi, D., Aridor, K., Makar, K., & Bakker, A. (2012). Children’s emergent articulations of uncertainty while making informal statistical inferences. *ZDM Mathematics Education, 44*, 913-925. http://dx.doi.org/10.1007/s11858-012-0420-3

Ben-Zvi, D., Makar, K., Bakker, A., & Aridor, K. (2011). *Children’s emergent inferential reasoning about samples in an inquiry-based environment*. Paper presented to the 7th Congress in Mathematics Education, Rzeszow, Poland, 9-13 February. Retrieved from http://www.cerme7.univ.rzeszow.pl/WG/5/CERME_BenZvi-Makar-Bakker-Aridor.pdf (Accessed 20 December, 2012)

Camtasia: Tec Smith Corporation. (2000). *Catania studio* (Version 6.0) [Computer software]. Okemos, MI: Tec Smith Corporation. Retrieved October 20, 2009, from http://www.techsmith.com/camtasia.asp

delMas, B. (2004). A comparison of mathematical and statistical reasoning. In D. Ben-Zvi & J. Garfield (Eds.), *The Challenges of developing statistical literacy, reasoning and thinking* (pp. 79-95). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Kobold, C., & Pollatsek, A. (2002). Data analysis as a search for signals in noisy processes. *Journal for Research in Mathematics Education, 33*(4), 259-289. http://dx.doi.org/10.2307/749741

Lane-Getaz, S. J. (2006). What is statistical thinking, and how does it develop? In G. F. Burrill & P. C. Elliott (Eds.), *Thinking and reasoning with chance and data* (pp. 273-290). Reston, VA: The National Council of Teachers of Mathematics.

Makar, K., Bakker, A., & Ben-Zvi, D. (2011). The Reasoning Behind Informal Statistical Inference. *Mathematical Thinking and Learning, 13*, 152-173. http://dx.doi.org/10.1080/10986065.2011.538301

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

Makar, K., Wells, J., & Allmond, S. (2011). Is this game 1or game 2? Primary children’s reasoning about samples during inquiry. In *Presentation Papers from the International Collaboration for Research on Statistical Reasoning, Thinking and Literacy (SRTL7)* (pp.17-39). Utrecht, the Netherlands: Utrecht University.

Mason, J., Burton, L., & Stacey K. (2010). *Thinking Mathematically* (2nd ed.). United Kingdom: Pearson Education.

Oxford University Press. (2012). *Oxford Dictionaries*. Retrieved 24 October, 2012, from http://oxforddictionaries.com/definition/english/thinking

Pfannkuch, M., & Wild, C. (2004) Towards an understanding of statistical thinking. In D. Ben-Zvi & J. Garfield

(Eds.), *The Challenges of developing statistical literacy, reasoning and thinking* (pp. 17-46). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Prodromou, T. (2011). Students' emerging inferential reasoning about samples and sampling: *In Proceedings of the Thirty Fourth Annual conference of the Mathematics. Education Research Group of Australasia* (pp. 640-648). Alice Springs, Australia. Retrieved from http://www.merga.net.au/documents/RP_PRODROMOU_MERGA34-AAMT.pdf (Accessed 10 December, 2012)

Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review, 67*(3), 223-265. http://dx.doi.org/10.1111/j.1751-5823.1999.tb00442.x