LINKING PROBLEMS, CONCLUSIONS AND EVIDENCE: PRIMARY STUDENTS' EARLY EXPERIENCES OF PLANNING STATISTICAL INVESTIGATIONS

Jill Fielding-Wells
School of Education, University of Queensland, Australia
j.wells2@uq.edu.au

An overview of many primary programs demonstrates the passivity of statistical learning in the junior years. Students are usually provided clean, orderly, simplistic data, or data representations, with which to work. When students are encouraged to collect their own data, it is limited to that which could be expected to cause little difficulty. The focus on contrived and unsophisticated data collection and analysis denies younger students the opportunity to design their own statistical investigations. The research reported here derives from the introduction of the statistical investigative cycle (Wild and Pfannkuch, 1999) to a classroom of 9-10 year old students. The students initially experienced difficulty envisioning the investigation process, despite both explicit instruction and multiple prior experiences with investigative learning. A focus on connecting problems and conclusions to evidence enabled students to plan investigations more efficiently.

INTRODUCTION

In an increasingly data rich society, there is a necessity for the development of statistical knowledge and understanding amongst citizens. This need has been acknowledged by the existence of statistics in the curriculum or curricular standards from an early age in many countries (e.g., Australian Curriculum, Assessment and Reporting Authority, 2009; Ministry of Education [New Zealand], 2007; National Council of Teachers of Mathematics, 2000). However, research into how young children develop statistical understandings is sparse with a significant need for further work.

The development of conceptual knowledge in statistics is to be valued, although it is insufficient to develop the critical thinking that is needed in statistical investigations. As Wild and Pfannkuch (1999) argue, undertaking 'projects' may broaden students' statistical experiences but without explicit instructional support and guidance, it is insufficient to develop statistical thinking and understanding of key theoretical structures. The importance of students' experience with the entire investigative process highlights the purposeful nature of statistics and its ability to explain, describe and/or quantify phenomena. To remove the problematised context that is inherent with statistical investigation is often to provide contrived, decontextualised, or artificial data with which to work; the detail required to make links visible is removed, thus obscuring essential connections. Despite this, much of the work that is done with children in the school setting focuses on data representation and low-order analysis/interpretation. Statistics problems encountered rely predominantly on neat, carefully derived data sets and well-defined problems in which many of the difficult decisions have been removed. In his review of research in statistics education, Shaughnessy (2007) acknowledged that statistics education in the United States had a tendency to emphasize data collection, analysis and conclusion while neglecting the nature of the problem and planning phases.

> If students are given only prepackaged statistics problems, in which the tough decisions of problem formulation, design and data production have already been made for them, they will encounter an impoverished, three-phase investigative cycle and will be ill-equipped to deal with statistics problems in their early formation stages. (p.963)

As a profession, educators have little detailed understanding of how best to scaffold student learning of the statistical cycle, challenges they may face and developmentally appropriate sequencing. The purpose of this paper is to provide insights that emerged from observation and analysis of a class of fifth grade students (aged 9-10) engaged in the early phases of learning to manage their own data investigations. Details of the framework chosen and the situational context are also provided.

FRAMEWORK AND LITERATURE

The statistical investigative cycle acts as a framework to build and develop statistical problem solving. While several models have been proposed (see, for example, Franklin et al, 2007;

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

Wild & Pfannkuch, 1999) with some variation noted, they are more similar than different. Wild and Pfannkuch's (1999) statistical investigative cycle (known as the PPDAC cycle) was adopted as a framework to introduce and scaffold students' learning in statistics. The PPDAC cycle describes five steps of a statistical investigation: Problem, Plan, Data, Analysis, and Conclusion. At the primary level, this could be described as:

Problem - The deconstruction, negotiation and refining of the problem in conjunction with context familiarisation.

Plan - The identification of the data needed to address the problem and consideration of effective collection, recording and analysis of that data.

Data - Data collection, recording and cleaning.

Analysis - Organising, manipulating, and interpreting data to identify trends or patterns which provide evidence with which to address the problem.

Conclusion - Reflecting upon the evidence identified in the analysis stage and linking it back to the initial problem in order to provide a response to that problem.

There is little research on students' experiences in planning statistical investigations, although its importance is alluded to in several documents. For example, Makar and her colleagues talk about the importance of allowing students "to envision the inquiry process, looking forward and backward to link the question-evidence-conclusion connections meaningfully as they work" (Makar, Bakker, & Ben-Zvi, under review, p. 21) and that "[making] this connection more explicit may help teachers to better support these links [for their students]" (Makar & Rubin, 2009, p. 86). In describing how experienced statisticians work, the GAISE Report (Franklin et al, 2007) notes the importance, and the difficulty, of linking questions to data to conclusion:

> At the point of question formulation, the statistician anticipates the data collection, the nature of the analysis, and possible interpretations. … In the end, the mature practitioner reflects upon all aspects of data collection and analysis as well as the question, itself, when interpreting results. Likewise, he or she links data collection and analysis to each other and the other two components [question and conclusion]. Beginning students cannot be expected to make all of these linkages. They require years of experience and training (p. 12-13).

Students cannot be expected to fully determine the nuances of variation in a data set, or to be able to make decisions about data collection methodology, unless they have been exposed to planning for the collection and analysis of data themselves. While "the main purpose in statistics is to give an accounting of the variability in the data" (Franklin et al, 2007, p. 12), it is difficult, if not impossible, to adequately describe the variability of a data set without understanding its collection, design and context. Shaughnessy (2007) suggests "a need for several iterative cycles just between the Problem ↔ Plan phases of the Investigative cycle" (p. 963) to support students in making these links.

METHOD

Participants

The focus group of students was a single, co-educational class of twenty-four, 9-10 year old students in a suburban primary school in Australia. The students had been exposed to a range of teaching methods in learning mathematics and statistics prior to the current year with some having experienced a predominantly traditional, transmission-based model, some having had multiple experiences with inquiry-based learning, and others having experienced both. At the time this research commenced, the class had been together for over five months and the classroom teacher had used both a traditional and inquiry-based approach to teaching mathematics and statistics. During this time the students had undertaken two sizeable inquiries. In both instances, students had been provided with research questions and, using questioning and discussion rather than explicit direction, had planned and implemented mathematical inquiries to determine a possible outcome. The classroom teacher/researcher had no specific training in mathematics or statistics teaching; however, she had completed several research methodology subjects at the post-graduate level.

Tasks

Students were engaged with two tasks designed to give students experience with multiple iterations of planning a statistical investigation. The first task presented to the students was to determine whether shower timers that had been issued by the State water authority were useful. The local area had been experiencing long-term drought and shower timers had been provided to residents to encourage them to limit showers to 4 minutes. They were similar to egg-timers in appearance and designed to be affixed to the shower wall. The researcher's purpose in selecting this topic was twofold: first to engage the students with something topical, authentic and of high community importance; and second to provide the students with a problem that required the generation of data and that would trigger a statistical investigation cycle. For the second task, a member of the school staff needed help choosing the music for the upcoming Year 4 and 5 school dance. The students were asked to investigate which songs would be the most popular to play at the school disco. These tasks allowed for multiple cycles of planning to be undertaken by students with numerous attempts at a shower timer task (Cycles 1-3), followed by a second inquiry into selecting songs for the school disco (Cycles 4-6).

Data and analysis

At times, students worked both collaboratively and individually. Students' work samples and journals were collected over the course of both inquiries to analyse and assess development. In order to study the reasoning of students in more depth, one student from each of six collaborative groups was randomly chosen and their work tracked in detail. The planning tasks the students completed were assessed according to the planning considerations identified:

- Data Collection: Identifying the data required, identifying the target sample/population, proposing a method, and considering the need to record the data.
- Data Analysis: Identifying a method of organising/handling the data to enhance its usefulness and ability to be interpreted
- Conclusions: Recognition of who would require the 'answer', identifying a suitable format for communicating the answer.

These areas were then scored according to the level and appropriateness of detail included. Table 1 provides an overview of these levels.

Table 1. Analytic framework applied to students' attention to investigative planning

Level	Description
0	No attention or inappropriate response
1	Indirectly mentioned
2	Directly mentioned with no details OR details provided but would not answer the question
3	Relevant, workable details provided, but insufficient detail to replicate
4	Substantial and relevant details provided

RESULTS

In this section, each cycle is described in more detail, highlighting challenges students encountered as well as elements that triggered progress. Cycles 1-3 relate to the shower timer task whereas Cycles 4-6 describe students' work on the disco song task. A table summarising the progress of the sample over Cycles 4-6 is provided at the end of this *section.*

Shower Timer Task

Cycle 1. Initially, each student was provided with a shower timer, given time to experiment (play) with it, and asked what questions they had about the timers. The class collectively negotiated a question "Is there a difference between the 4 minute shower timers distributed by [name of state] Water?" They were specifically tasked with planning an investigation that would enable them to provide an answer to the question. The students' initial responses were fairly arbitrary with some form of formal or informal measurement. For example many students began racing timers against other students without any attempt to record the results. No student followed the instruction to create a plan, despite attempts by the researcher to refocus the students on the task.

Cycle 2. A class discussion ensued during which students brainstormed the general activities they undertake during an inquiry. These activities were clustered to formally introduce them to the PPDAC cycle (Wild & Pfannkuch, 1999); their brainstorming had identified 4 of the 5 steps of the PPDAC cycle, with only 'conclusion' being omitted. The stages of the PPDAC cycle were deconstructed with each stage of their last inquiry mapped against the PPDAC framework, and then used to generate the students' own version of a 'planning checklist' to further enhance their understanding. Students worked in groups to attempt to construct an improved plan. The completed plans revealed that the majority of groups had focussed heavily on methodology; how to start, how/when to turn the timers, how to record the data and so forth, without considering whether the resultant data was useful. One group of students focussed almost purely on what could be called administrative issues; for example, how many students would be required to carry out the timings. Another focussed on designing a histogram to record the data, with no reference as to how to collect data; and a further group created an attractive, carefully drawn poster of the PPDAC cycle with no content relevant to the investigation.

Cycle 3. In order for students to recognise shortcomings, they were asked to implement their plans, noting any changes needed. In doing so, they quickly realized that significant modifications were required as their plans did not include sufficient detail. Students continued to have difficulty envisioning the steps that were required and frequently lost sight of the purpose of the investigation. A class discussion allowed students to suggest a number of difficulties and omissions in their plans, for example their attention to insignificant issues, collection of data that did not answer the question, insufficient sample size, and preoccupation with graphing (rather than recording and interpreting) the data. A concerted focus was placed on the need to make a clear link between the question, the evidence and the conclusion.

Disco Song Task

Cycle 4. The context was purposely changed in order to assess students' ability to transfer their learnings to a new context: Choosing the most popular songs that could be played for the school disco. The students were instructed to individually write down a question and a plan to enable the question to be answered. The students' initial responses showed similar difficulties to those obtained in the shower timer unit, with few students making any reference to analysis or communication of the data once it had been collected. Some students suggested that the information might be available on the internet and that they could 'Google' it. The problems encountered previously were still present even after significant explicit discussion and instruction.

Cycle 5. The statistical investigation cycle was reviewed and students were asked to go through their plan systematically and identify elements they had included (or omitted) in their planning checklist (Cycle 2). The students quickly realized that their plans did not adequately consider the specifics of what data was required or how it would be recorded. As the students had more success with the front end of the investigation, the researcher focused on a reverse process, identifying an anticipated conclusion in response to the question asked, for example a list of the 15 songs most desired by students at the school dance. The students were asked to envisage what evidence would be required to argue or convince others of their findings. A scaffolded planning sheet was provided to focus the students on the need for a link between question, evidence and conclusion. The students found that the specific structure of the individual sections helped to focus on the necessary links (see Figure 1). Students' responses were much more developed when the scaffolded planning guide was offered.

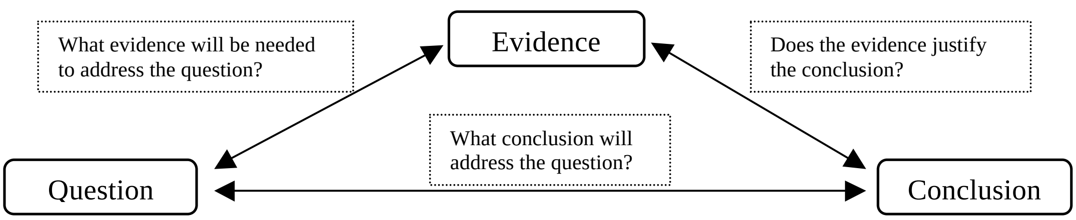

Figure 1. Connections between question-evidence-conclusion in a statistical investigation

Cycle 6. Notable gains were observed after the students were directed to consider the question↔evidence↔conclusion links; however, students displayed varied strengths and weaknesses in envisioning different aspects of the cycle. To respond to this, the researcher formed the students into small working groups of three or four, and asked them to discuss and defend their plans until they had formed the best plan they could as a group. This plan would then be presented formally to the class and the teacher for feedback. The collaborated versions of the students' work typically showed significant gains towards a workable investigation. It was notable that the few instances where gains were not as great could be attributed to groups where a dominant personality had done much to force their own plan on the group. After presenting to the class, students were offered the opportunity to adjust their plan based on the feedback received. The one group that had demonstrated individual dominance then adjusted their work collaboratively to achieve a more workable investigative plan.

Summary

The six cycles with which the students engaged showed increasingly sophisticated plans being developed as students gained experience, were provided with explicit structures, collaborated, and then presented their findings to the class. Table 2 summarises the outcome levels of the sample of students (n = 6) chosen from each group for further analysis in Cycles 4-6. An explanation of the assignment of values can be found under the section 'Data and analysis'. These results suggest that considerable gains in Cycle 5, when students were provided with scaffolded planning sheets which focused the students on the conclusion prior to considering data analysis and collection. Cycle 6 also resulted in further improvement with the implementation of collaborative work.

Table 2. Summary scores of randomly selected student samples across Cycle 4, 5 and 6

Cycle	Average 'Data Collection' Score	Average 'Data Analysis' Score	Average 'Conclusion' Score	Average Total Score
4	0.63	0.50	0.20	0.43
5	1.79	1.58	2.08	1.82
6	2.88	2.33	3.25	2.82

DISCUSSION

Shaughnessy (2007) and Franklin et al. (2007) both observe that it is not easy for children to develop the understandings necessary for statistical thinking, reasoning and literacy, and their arguments were clearly supported by the research findings. The students involved in this study, despite having had multiple experiences with inquiry learning, continued to have difficulty envisioning the statistical investigative process. Clearly the need to teach the investigative cycle is critical as even students who were familiar with inquiry did not appear to make those connections – despite being able to easily identify most of stages in the PPDAC cycle. Wild and Pfannkuch (1999) similarly argued that letting students 'do' projects was insufficient to develop knowledge of the statistical cycle, that explicit teaching was necessary to develop student understanding.

The majority of the difficulties noted were in making linkages between question, evidence and conclusion. These can be broken into three categories:

1. Question ↔ Evidence: Difficulty with generating a question and then being able to envisage what evidence might be needed in order to address that question, the type of data to be collected and the method of collection.
2. Evidence ↔ Conclusion: To examine the evidence (data), understand its source, and draw upon that knowledge to make conclusions, draw inferences and so forth, within the context of the investigation.
3. Conclusion ↔ Question: Considering the conclusion and linking this back to the initial question asked, determining whether the question asked has been answered.

Focusing students specifically on a connections model (Figure 1) had the most noticeable impact on students developing understandings. Discussing and displaying this model made it possible for the researcher to continuously focus the students on the need to make essential connections.

The other area of significant gain occurred as students progressed from individual work to group work. Possible explanations for this trend may be the effects of argumentation and accountability. There is evidence to support the increased understandings that develop when students work collaboratively to develop shared understandings about a topic through evidence and persuasion (Berland & Reiser, 2008). In addition, as the students realised that their work would be presented to a peer and teacher audience, the increased accountability may also have motivated them to further improve their plans, attending to clarity and addressing possible gaps in reasoning.

On a final note, the first attempt to create an investigative plan for the School Disco task showed minimal ability to envisage the investigative cycle at all, despite students having recently completed significant work on the Shower Timer plan. This would indicate that direct transfer was not immediate across contexts. Several researchers have identified the necessity for multiple iterations of activities in order to develop linkages between aspects of statistical investigations (Franklin et al, 2007), across a range of contexts in order to adequtely develop deep learning. The results from the tasks reported here would add some weight to those arguments.

CONCLUSION

Shaughnessy's (2007) assertion has merit, focusing on data collection, analysis and conclusion alone is not sufficient to develop deep understanding of statistical process; but at what age can students begin to develop these understandings? There is significant need for work to be undertaken in this field as relatively little is known about the process by which children develop the ability to plan their own statistical investigations.The purpose of this research was to initially observe young students as they tried to envision the statistical investigation process and the results are encouraging in terms of the possibilities.

REFERENCES

Australian Curriculum, Assessment and Reporting Authority (2009). *The shape of the Australian curriculum*. Access date 29/10/09 from http://www.acara.edu.au/verve/_resources/Australian_Curriculum_-_Maths.pdf

Berland, L. K., & Reiser, B. J. (2008). Making sense of argumentation and explanation. *Science Education, 93* (1), pp.26-55.

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, R., & Scheaffer, R. (2007). *Guidelines for assessment and instruction in statistics education (GAISE) report: A preK–12 curriculum framework*. Alexandria, VA: American Statistical Association.

Makar, K., Bakker, A., & Ben-Zvi, D. (under review). Informal statistical inference: Examining the emergence of students' statistical reasoning. Paper under review for *Mathematical Thinking and Learning*, Special Issue on the Role of Context in Informal Statistical Inference (Makar & Ben-Zvi, Guest Editors).

Makar, K., & Rubin, A. (2009). A framework for thinking about informal statistical inference. *Statistical Education Research Journal, 8* (1). Access date 29/10/09 from *http://www.stat.auckland.ac.nz/~iase/serj/SERJ8(1)_Makar_Rubin.pdf*

Ministry of Education. (2007). *The New Zealand curriculum*. Wellington: Learning Media Ltd.

National Council of Teachers of Mathematics (2000). *Principles and standards for school mathematics*. Reston, VA: NCTM. http://www.nctm.org/standards

Shaughnessy, J. M. (2007).Research on statistics learning and reasoning. In F. K. Lester (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning*. Charlotte, NC: NCTM.

Wild, C. J. & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review 67*(3), 223-265.