WHAT ABOUT THE P IN THE PPDAC CYCLE? AN INITIAL LOOK AT POSING QUESTIONS FOR STATISTICAL INVESTIGATION.

Pip Arnold
The University of Auckland, New Zealand

Posing questions is at the heart of any statistical investigation cycle. This paper describes a pilot study exploring the questions 14-15 year old students pose about given multivariate datasets. An overview of the method used including the tasks given to the students, a possible categorisation framework for the different types of questions, and initial findings about the different categories of questions (summary and comparison) posed by students are discussed. Suggested further research in the area is given.

OVERVIEW

The new statistics curriculum in New Zealand has consolidated the use of the statistical investigation cycle in high school classrooms. The need for the PPDAC (Problem, Plan, Data, Analysis, Conclusion) cycle (Wild & Pfannkuch, 1999) to be modelled in teaching programmes is important and resources have been developed to support teachers. Students need to become familiar with all aspects of the cycle, from understanding and defining the problem, to:

> determining appropriate variables and measures; considering sources of variation; gathering and cleaning data; using multiple displays, and re-categorising data to find patterns, variations, relationships, and trends in multivariate data sets; comparing sample distributions visually, using measures of centre, spread, and proportions; and presenting a report of findings (Ministry of Education, 2007, p. 55).

The pilot study, which was focused on exploring specialised statistical content knowledge that teachers need, was based in one Year 10 (14-15 year olds) classroom. A sequence of lessons based on Census at Schools data (www.censusatschool.org.nz) was used to introduce the students to the statistical investigation cycle. The aim of the lessons was to develop the complete PPDAC cycle with the students from posing questions through to writing the conclusions. Previously the focus in teaching had been on drawing graphs and calculating statistics. During the teaching of the lessons, the posing of questions from multivariate datasets was identified as a potential obstacle for students. This obstacle is problematic, since in Year 11 posing questions is an integral part of a statistics assessment for national qualifications. The research question for this paper addresses the different categories of questions (summary and comparison) students are posing at the beginning of a statistical investigation based on a given multivariate dataset.

LITERATURE REVIEW

The statistical investigative cycle is a framework that supports statistical problem solving (Wild & Pfannkuch, 1999; Franklin & Garfield, 2006; Konold & Higgins, 2003; Graham, 2006; Zayac, 1991; Bishop & Talbot, 2001). The different authors present the statistical investigative cycle with either four or five phases. The commonalities amongst them are greater than the differences. The investigative cycle while dynamic in its nature has at its heart a starting point based on a problem.

Research on the questions students pose appears to be limited and this has been noted and identified as a potential problem area for students and teachers by Burgess (2007) and Rubick (2000). There is research that mentions posing questions where data collection is required (Konold & Higgins, 2003; Lehrer & Romberg, 1996; Russell, 2006). Konold and Higgins, (2003) report on classroom episodes where teachers and students together transform general questions into statistical questions. Lehrer and Romberg (1996) describe a study on elementary students' reasoning where they had students in their study pose questions about their own lives and then select the questions to be asked. In addition to the students posing their own questions for others to answer, the students were asked to come up with the question that they wanted to investigate from the data. Lehrer and Romberg explain that the students were given seed questions to help them and reported on the questions posed, but there was no classification or discussion about the students' questions. Russell, (2006, p. 18) describes the "complexity of creating data that

represents events in the world" and the need for teachers and students to turn real-world questions into statistical questions that can be answered using data. Other research reports on teacher and researcher generated questions which students answer from given data (Bakker & Gravemeijer, 2004; Friel, Curcio & Bright, 2001).

An element to consider when posing summary questions is the individual view versus the aggregate view. Franklin and Garfield (2006, p.350) consider this element when formulating statistical questions. They state, "the formulation of a statistics question requires an understanding of the difference between a question that anticipates a deterministic answer and a question that anticipates an answer based on data that vary". Other research around the individual and aggregate view is in relation to student reasoning. Konold and Higgins (2003, p. 203) report "the answers to many of the questions that interested students – for instance, Who is the tallest? Who has the most? Who else is like me? - require locating individuals, especially themselves, within the group". They suggest that it is not until students have to answer a question such as whether two groups may differ that students will start to think about the group characteristics rather than the individual.

There are various frameworks in the literature that focus on or include statistical questions. Some describe the different types of questions students will ask depending on the nature of the data (Pfannkuch & Horring, 2005; Graham, 2006). Questions can be classified as summary, comparison or relationship questions. Summary questions are posed when a description of the data is needed and are usually about a single data set. Comparison questions are posed for comparing two (or more) subsets of data across a common variable, for example, male and female, young and old. Relationship questions are posed for looking at the interrelationship between two paired variables. Franklin and Garfield's (2006, p. 353) conceptual framework model describes posing questions and student movement towards the "statistics question distinction" as one of the statistical problem solving processes. Friel et al. (2001, p. 130), however, when defining graph comprehension make a distinction between question posing and question asking. They discuss three levels of graph comprehension from the beginning level of "extracting data from a graph" through an intermediate level "interpolating and finding relationships in the data as shown on a graph" to an advanced level "extrapolating from the data and analysing the relationships implicit in a graph". Friel et al. propose a three level framework for question asking "read the data, read between the data, and read beyond the data".

When posing questions in the statistical investigative cycle context it is important to separate out two distinctive question purposes. The *investigative question*, the question being asked of the data, and the *survey question*, the question being asked to get the data. Lehrer and Romberg (1996, p. 80) describe the questioning processes they used in their study. First there is the survey question in relationship to the constructing of the data, where students generate questions for a survey. Second there is questions for analysing the data, where students come "up with questions that they may ask about the data". Russell (2006, p. 21) appears to refer to two different question purposes when making suggestions for what teachers can do to help their students in the problem and planning phases. She says, "help students to keep in mind their original questions and interests and to consider whether their data collection questions and methods are resulting in data that yield information about those original questions". In these research examples the students are posing questions with two different purposes, which provides students with the opportunity to connect survey and investigative questions. However, when students are given a multivariate dataset often the survey questions do not exist or are not given and this connection is not made. The connection between the survey and investigative questions is important as it forms part of the context of the investigative questions.

All the research either describes events that occurred in the classroom in relation to a situation where students were posing questions to collect their own data, or presents frameworks for posing questions for instruction purposes. However, there appears to be no research documenting the different categories of questions that students actually pose from given multivariate datasets generating an ideal platform from which to launch an initial exploration about the different categories of questions that students actually pose. This is not an idea to be taken lightly as posing questions and interrogation of the data occur throughout the whole enquiry process. Therefore in this paper the research question that will be addressed is: What different

categories of investigative questions (summary and comparison) are students posing at the beginning of a statistical investigation from given multivariate datasets?

METHOD

This paper will report on the analysis of the pre- and post-test responses of a group of 15 average to below average students to a question on posing questions and an analysis of the summary and comparison type questions the students and the teacher were posing during the first two class sessions. In the previous year the large (2500 students) coeducational multicultural school had not covered any statistics with these students. These 15 students from a class of 24 gave permission for their test results to be analysed. Only 12 of these students completed the post-test. A ten-lesson teaching unit was developed collaboratively by the researcher and the teacher to follow the PPDAC cycle. The lessons were videotaped and transcribed.

The teacher and researcher were involved in a reflection process after each lesson. This allowed adjustments to the original plan as issues came to light and it seemed sensible to make the changes immediately. This process forms part of a teaching experiment (Bakker, 2004). While reflecting on student generated questions from the first lesson some of the issues around the questions were clarified, for example, questions that had an individual case view versus questions that had an aggregate view. As a result of this discussion the second lesson was modified.

Categories for summary and comparison questions were hypothesised and these were refined during the analysis process. Classroom based questions from the first two lessons were included with the pre- and post-test questions to develop the different categories of questions.

TASKS

In the pre-test students were asked to pose a question in relation to a given multivariate dataset. The dataset had 14 students randomly selected from the 2005 Census at School New Zealand database. The variables given for each student were: gender, age, region they live in, years they have lived in New Zealand, ethnicity, height (in cm), right foot length (in cm), time it usually takes to travel to school (in mins), main activity they did at lunchtime and year level. The students were then asked to answer their question by drawing a graph to display the data and then to make three statements about what their graph showed.

In the first two lessons the teacher focused the work on posing questions from a given multivariate dataset. The data were given to the students in the form of data cards (Watson, 2006). An example of a data card is shown in Figure 1. The dataset contained 180 students, 30 from each year level, 5 to 10 (age 9-15) randomly selected from the 2005 Census at School New Zealand database. Groups of students in the class were given a set of 30 data cards from one year level.

Number:	4
Gender	Girl
Ethnicity	NZ European
Year Level	10
Height (cm)	170
Right foot (cm)	27
Armspan (cm)	173
Wrist circ. (cm)	18
Neck circ. (cm)	35.5
Reaction (secs)	0.44
Sleep (hours)	10
Bedtime	9pm

Figure 1. Example of a data card used.

The focus of the first lesson was to get students to start to think about the different types of questions that could be asked, in particular, summary, comparison and relationship. The teacher gave an example of each type of question and then gave the students time to generate their own questions for each type based on the data cards. The starter, *I wonder…,* was used extensively as part of the posing question process. Questions from the class were collected and

collated under the three different types, with the students deciding which type each question was. During the unit of work the teacher continued to bring each aspect back to the fact that they were trying to answer a particular question.

In the post-test students were asked specifically to write three summary questions and three comparison questions in relation to a given multivariate dataset. The dataset had 30 Year 10 students randomly selected from the 2005 Census at School New Zealand database. The variables given for each student were: gender, age, ethnicity, height (in cm), right foot length (in cm), arm span (in cm), wrist circumference (in cm), person they look up to, how they usually travel to school, time it usually takes to travel to school (in mins), reaction time (in secs) and hours sleep last night. The students posed their questions in the lesson before the rest of the assessment as the teacher needed to check the questions. Students had to choose a question to use once the teacher had checked that the questions were suitable. If all their questions were not suitable the teacher gave them a question to use. This model of posing the question in a lesson before reflects the national assessment for qualification in Year 11.

RESULTS

A qualitative analysis by the researcher of the questions posed in the pre- and post-tests and during the first two lessons extracted eight different categories of summary questions and seven different categories of comparison questions (Figures 2 and 3). In the summary questions categories 1 to 6 and 8 were evident in the questions posed by the teacher and students and for the comparison questions categories 1 to 5a were evident. The additional categories (7 for summary and 5b for comparison) were conjectured using Graham's (2006) definition. Therefore, categories 7 and 8 are considered as summary questions, categories 4 to 6 are considered as pre-summary questions and categories 1 to 3 as non-summary questions. The analysis of the data suggests that there is a hierarchical order overall for the categories of questions posed. In the comparison question categories, however, 4 and 5 are considered to have no hierarchical order.

Summary question category	Student example
1. Nonsense or not a summary question.	Is the Asians circled above related?
2. A question that is partially related to the data, but not answerable by the given data.	What region is the most preferred to live in?
3. A question that asks about an individual case.	I wonder which person has the biggest neck size?
4. A question that asks how many of a particular category.	I wonder how many people look up to their family?
5. A question that asks how many within a specified range.	I wonder how many Year 6 girls go to bed in between 8:30pm and 10pm?
6. A question that asks for the most popular or most common.	I wonder what is the most popular ethnicity?
7. A question that asks about the overall distribution of the data or what is typical.	*What is the typical way students usually travel to school?*
8. A question that asks about the overall distribution of the data or what is typical and reflects the population for which conclusions can be drawn.	I wonder what a typical time a Year 10 student goes to bed?

Figure 2. Summary question categories. (*Italics indicate a hypothesised question.*)

The students' questions from all three situations and the teacher's questions from the first two lessons were then analysed using the classifications described in Figures 2 and 3. There were 58 summary questions and 44 comparison questions that were considered in the analysis.

Comparison question category	Student example
1. Nonsense or not a comparison question.	Is there more 15 year olds or more 14 year olds?
2. A question that is partially related to the data, but not answerable by the given data.	I wonder if the students in Wellington take longer to get to school than students in Auckland? (Region was not a variable in the dataset.)
3. A question that is related to the data but not answerable due to sample size issues.	I wonder if Asian girls have a longer arm span than Indian boys? (Two Asian girls and two Indian boys in the dataset.)
4a. A question that is answerable by the data.	I wonder if boys are taller than girls?
4b. A question that is answerable by the data and requires recategorisation of categories to be compared.	I wonder whether people who have the longest time travel get less sleep than those who travel for less time?
5a. A question that is answerable by the data and reflects the population for which conclusions can be drawn.	I wonder if Year 10 boys have a longer arm span than Year 10 girls?
5b. A question that is answerable by the data and requires recategorisation of categories to be compared and reflects the population for which conclusions can be drawn.	*I wonder if Year 10 NZ European students are taller on average than Year 10 students of other ethnicities?*

Figure 3. Comparison question categories. (*Italics indicate a hypothesised question.*)

Eleven students had both a pre- and post-test classification for summary questions which are summarised in Table 1. In the pre-test the students posed one question and in the post-test three questions. The question category chosen for the post-test was the majority category for 8 students (which was their best category in all instances) and for the remaining three students (who had three different categories) their best category was chosen.

Table 1. Pre- and post-test summary question categories.

		Post-test summary categories								
		1	2	3	4	5	6	7	8	Total Pre
Pre-test summary categories	1	1		1	3					5
	2				1	1				2
	3									
	4				1	1	1			3
	5									
	6									1
	7				1					
	8									
	Total Post	1		1	6	2	1			11

Seven questions posed in the pre-test were classified as non-summary questions and four as pre-summary questions. In the post-test two questions were classified as non-summary and nine as pre-summary. The majority of the questions in the post-test were “how many” questions. Eight students “improved” the category of their summary question from the pre-test to the post-test, two remained the same, and one regressed.

Twelve students sat the post-test. The category of the question the students completed the remainder of the post-test are summarised in Table 2. In addition all of the questions that were posed in the post-test are summarised in Table 2.

Table 2. Post-test comparison question categories.

Post-test comparison category	Number of students who used this category of question to complete the post-test	Number of questions in the post-test that were this category
1		10
2		1
3	1	8
4a	9*	14
4b	2	2
5a		1
5b		
Total	12	36

* one student posed unsuitable questions and was given a question by the teacher to complete the test

Since the teacher pre-marked the students questions, most students in the post-test used a category 4a question to complete the test. Ten students had at least one suitable question within the three that they posed, one student had no suitable questions (all category 1) and was given one by the teacher, the other student also had no suitable questions, but the teacher did not pick this up and had marked category 3 questions as suitable. When considering all the questions posed by the students in the post-test, just under half were suitable (categories 4a, b and 5a) and just over half were unsuitable (categories 1 to 3).

The teacher gave examples of one category 4 summary question in the first lesson and two category 8 summary questions in the second lesson. The teacher gave examples of two comparison questions (4a and 5a) in the first lesson. In the second lesson the teacher reviewed the questions recorded from the first lesson and looked at rewriting the questions so that they would come up with a summary question to investigate. In the following excerpt she takes a category 4 summary question from the previous day and changes it to a category 8 summary question and it can be noted that she didn't address why the original question was unsuitable.

> T: today we are going to look at how we can answer these questions, but before we do that we want to make sure we have a good question to start with so we are going to rewrite the questions just a little bit. Let's look at the question here, I wonder if all Year 5 students go to sleep at 7pm? Can anyone think how we could reword that to make it a little easier to investigate or think of another way we could say that? Maybe we want to know what a typical time year 5 students goes to bed, let's see how we rewrite this. How about we wonder what a typical time a year 5 student goes to bed is?

DISCUSSION

The research question for this paper addresses the different categories of investigative questions (summary and comparison) students are posing at the beginning of a statistical investigation based on a given multivariate dataset. The findings suggest that the students can pose comparison questions but are not successful yet at posing summary questions. Of the 12 students who sat the post-test 10 of them posed at least one suitable comparison question of the three comparison questions they posed. None of the students posed a suitable summary question in the three they posed in the post-test, though seven of them posed predominantly pre-summary questions. These questions varied from the how many of a particular category through to what was the most popular category.

The teacher modelled both a pre-summary question and two suitable summary questions within the two lessons that were analysed, however, she didn't address what made a good summary question with the students. It would appear that modelling the questions without addressing what made a good question was insufficient to facilitate a major change in the category of summary questions that students were posing. Students have been exposed to summary-type questions for a number of years. Predominately these have been textbook examples which usually involve a given data display and then a series of questions of the pre-summary type to answer. There are no examples in current New Zealand textbooks requiring

students to pose the question to answer, or even the situation where the given question to answer is a summary question. The types of questions asked in textbooks are "read the data" and "read between the data" (Friel et al., 2001, p. 130) types of questions. These questions match with the pre-summary posing questions. There is no evidence of "reading beyond the data" (Friel et al., p. 130) types of questions that match with the summary investigative questions.

All these questions are valid within the statistical investigation cycle. However, there is a distinction between question posing and question asking (Friel et al.). The summary investigative questions involve an aggregate view and need to be *posed* both by teachers and students. Whereas the individual view or non-summary questions and the pre-summary questions need to be *asked* by students when they are writing their analysis.

The teacher modelled suitable comparison questions in class. Generally this was the category of questions that the students used in the post-test. Comparison questions are new for nearly all of the students at this level. It would appear that having less previous exposure to comparison questions means the students predominately use the model they have seen. The one issue that came up with the comparison questions was students selecting a comparison group that had too small a sample size. This was the situation for one student, though the teacher didn't identify these categories of questions as being not suitable. The student's question was "I wonder if NZ Europeans sleep longer than Others?" Other was one of the categories for ethnicity. This question generated three problems: Other had a small sample size (n = 5); the student was unable to "recategorise" and put the Other with Asians, Maori etc. to include all other ethnicities; when the student answered the question he reverted to answering a "how many" question and to noticing individual cases. The student didn't appear to understand the original question he had asked. Therefore the ability to pose a question at a certain level does not mean that a student can answer the question.

When comparing the findings of this study with Franklin and Garfield's (2006, p. 353) framework most of the student posed summary-type questions would be at Level A: "beginning awareness of the statistical question distinction", whereas most of the comparison-type questions would be at Level C: "students can make the statistics question distinction". This framework provides a holistic view of the statistical investigative cycle. In terms of using it to determine the categories of the questions posed it is too broad. For example, it does not distinguish between questions that can be answered with the given data, and questions that cannot. It does not appear to be based on how students think.

Posing questions is at the heart of the statistical investigation cycle and asking questions is an integral part of all aspects of the cycle. Clarity around the purpose of the question at each stage of the cycle is critical as is the actual question that is posed for investigation. Ideas in regard to the survey question, such as, students "formulate one or more questions that can be answered with data" (Franklin & Garfield, 2006), and making a "question that is specific enough so they can collect relevant data yet make sure that in the process they do not trivialize their question" (Konold & Higgins, 2003) have already been considered, but the *investigative* question has not.

The key finding from this research is the proposal of a descriptive framework for the classification of the *investigative* questions (summary and comparison) that students pose. In addition to the framework suggested there are a number of moderating questions that are conjectured would be useful to consider when posing an *investigative* question from a given multivariate dataset. For example: What was the original question that was used to collect the data (survey question)? What type of data is being used? What graph or display of the data will be made? What hypothesis can be made about the data? Is the question interesting? Who would be interested in the answers to this question? Is there enough data available to answer the question (issues around sample size)? What background information is available about the data (how it was collected, who it was collected from, when it was collected etc.)? Is the variable of interest in the dataset? All of these issues might help to answer the question: Is the *investigative* question "right"?

Further work is needed to explore the relevance of the issues mentioned above in relation to posing questions for investigation. More work is needed to consolidate the framework of the types of investigative questions students can pose (summary and comparison) and to build a framework for relationship questions. Other questions to consider include: Are the data

continuous or discrete and does this matter? Is the question going to be about the sample or about the population? What sort of modelling by the teacher is required? What help do the teachers need before being able to do this? What are the links between the questions posed by the students and their analyses and conclusions? As Konold and Higgins (2003, p. 212) said "Helping students raise questions that interest them and that they can productively pursue is a challenge for the teacher."

REFERENCES

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

Bakker, A., & Gravemeijer, K. P. E. (2004). Learning to reason about distribution. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 147-168). Dordrecht, The Netherlands: Kluwer.

Bishop, G. & Talbot, M., (2001). Statistical thinking for novice researchers in the biological sciences. In C. Batanero (Ed.), *Training researchers in the use of statistics* (pp 215-226). Granada, Spain: International Association for Statistical Education and International Statistical Institute.

Burgess, T. (2007). Investigating the nature of teacher knowledge needed and used in teaching statistics. Unpublished doctoral thesis, Massey University, New Zealand.

Franklin, C. A., & Garfield, J. B. (2006). The GAISE Project. Developing statistics education guidelines for grades pre-K-12 and college courses. In G. F Burrill & P. C. Elliot (Eds.), *Thinking and reasoning with data and chance* (pp. 345-375). Reston, VA: National Council of Teachers of Mathematics.

Friel, S. N., Curcio, F. R., & Bright, G. W. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education, 32*(2), 124-158.

Graham, A., (2006). *Developing thinking in statistics*. London: Paul Chapman Publishing.

Konold, C., & Higgins, T. L. (2003). Reasoning about data. In J. Kilpatrick, W. G. Martin & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 193-215). Reston, VA: National Council of Teachers of Mathematics.

Lehrer, R., & Romberg, T. (1996). Exploring children's data modeling. *Cognition and Instruction, 14*(1), 69-108.

Ministry of Education. (2007). *The New Zealand Curriculum*. Wellington, New Zealand: Learning Media Limited.

Pfannkuch, M., & Horring, J. (2005). Developing statistical thinking in a secondary school: A collaborative curriculum development. In G. Burrill & M. Camden (Eds.), *Curricular development in statistics education: International Association for Statistical Education 2004 Roundtable*, (pp. 204-218). Voorburg, The Netherlands: International Statistical Institute.

Rubick, A. (2000). The statistical thinking of twelve year 7 and 8 students. Unpublished masters thesis, The University of Auckland, New Zealand.

Russell, S. J. (2006) What does it mean that "5 has a lot"? From the world to data and back. In G. F Burrill & P. C. Elliot (Eds.), *Thinking and reasoning with data and chance* (pp. 17-29). Reston, VA: National Council of Teachers of Mathematics.

Watson, J. (2006). *Statistical literacy at school: Growth and goals*. Mahwah, NJ: Lawrence Erlbaum Associates.

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

Zayac, S., (1991). Improving open-ended problem solving: lessons learned in industry. In D. Vere-Jones (Ed.) *Proceedings of the third international conference on teaching statistics. Volume 1 School and General Issues* (pp 79-83). Voorburg, The Netherlands: International Statistical Institute.