CREATING COGNITIVE CONFLICT IN A CONTROLLED RESEARCH SETTING: SAMPLING ®

Jane M. Watson
University of Tasmania, Australia

This paper reports on research that created a controlled environment for interviewing individual students on the topic of sampling, allowing for cognitive conflict from other students. At various points in the interview the student was shown video extracts with contrasting views to those expressed and ask for a reaction. Outcomes are discussed with respect to (a) the outcomes for 37 students, in terms of their reaction to the cognitive conflict presented, and (b) the methodology, in terms of modeling cognitive aspects of a classroom environment in a controlled setting.

INTRODUCTION

The importance of cognitive conflict as an influence on learning has long been acknowledged by educators. From the time of Piaget *assimilation* has been used to describe the integration of new ideas with old and *accommodation* has been used for the revision necessary when conflict is created as a result of the new ideas. For learning to occur, however, students must feel some dissatisfaction with current ideas and the new ones must be intelligible and appear plausible (Strike & Posner, 1992), and hence students need to express their own ideas clearly and have opportunities to see where these conceptions fail (Shaughnessy, 1985). From the perspective of research into accommodation of new ideas, classroom group contexts provide difficulties because of the inability to control the input provided by high level thinkers or the initial expressions of understanding from the targeted students. Hence replication is impossible.

One of the topics which students find difficult and hence which lends itself to study in this context is *sampling*. Student understanding of sampling is often gleaned from outside the classroom (Jacobs, 1999; Watson & Moritz, 2000) and because understanding is descriptive rather than based on mathematical calculations, the topic does not often get much attention in school mathematics classes. The research of Watson and Moritz acknowledged the influence of out-of-school contexts on student understanding and classified responses within a three-tiered hierarchy of statistical literacy (Watson, 1997) applied to sampling. Tier 1 was a basic understanding of sampling terminology. Tier 2 was an understanding of sampling language and concepts when they are embedded in the context of wider social discussion. Tier 3 was a questioning attitude one can assume when applying concepts to contradict claims made without proper statistical foundation of the sampling involved.

Watson and Moritz (2000) interviewed 62 students in grades 3, 6, and 9 using the protocol in Figure 1. The answers to these questions combined with responses to earlier survey items provided the basis for defining six levels (L1 to L6) of development in relation to sampling, situated within the statistical literacy hierarchy. These were summarised in the following fashion.

T*ier 1: Small samplers with idiosyncratic or no selection (L1)* – provide examples such as food, describe sample as "a small bit" but rarely as a "test", agree to sample size less than 15, suggest idiosyncratic or no method of selection. *Small samplers with primitive random selection (L2)* – as above but suggest selection "by random" or "choose any".

Tier 2: Small samplers with preselection of results (L3) – as above but describe sample as both "a small bit" and a "test or try", suggest selection of people by weight, either for a spread of fat and skinny, or for "normal" weight. *Equivocal samplers (L4)* – provide examples and descriptions of samples, may show indifference to sample size, mix either small sample size and appropriate selection methods or large sample size with inappropriate selection methods. *Large samplers with random or distributed selection (L5)* – provide adequate definitions, often mention average, suggest sample size of at least 20 or a percent of the population, suggest selection by a random process or geographical distribution.

Tier 3: *Large samplers sensitive to bias (L6)* – as for previous group and often mention surveys, express concern for selection to avoid bias, recognise sampling bias in newspaper articles.

Of the 62 students, 7 were Equivocal Samplers, 12 (all grade 3) were Small Samplers Without Selection, and 12 (11 grade 9, 1 grade 6) were Large Samplers Sensitive to Bias.

1.(a) Have you heard of the word "sample" before? Where? What does it mean? (b) A news person on TV says: "In a research study on the weight of Grade 5 children, some researchers interviewed a sample of Grade 5 children in the state." What does the word "sample" mean in this sentence? 2.(a) Why do you think the researchers used a sample of Grade 5 children, instead of studying all the Grade 5 children in the state? (b) Do you think they used a sample of about 10 children? Why or why not? How many children should they choose for their sample? Why? (c) How should they choose the children for their sample? Why?	3. The researchers went to 2 schools: 1 school in the centre of the city, and 1 school in the country. Each school had about half girls and half boys. The researchers took a random sample from each school: 50 children from the city school, 20 children from the country school. One of these samples was unusual: it had more than 80% boys. Is it more likely to have come from ❒ the large sample of 50 from the city school, or ❒ the small sample of 20 from the country school, or ❒ are both samples equally likely to have been the unusual sample? Please explain your answer.

Figure 1. Interview Protocol for Sampling (also used in Watson & Moritz, 2000)

THE CURRENT STUDY

The current study sought to extend the previous research on sampling by interviewing a new group of students using the protocol in Figure 1 and to overcome some of the complexities experienced in the classroom in an attempt to understand the effect of cognitive conflict in situations where students are given the opportunity to express their initial ideas as clearly as possible. It was explained to students that the researcher was interested in their views and their reactions to the views on the same topic of other students about their same age. Views of teachers or of the researcher were never canvassed. Of interest from the perspective of studying the effect of cognitive conflict on students' understanding of sampling was whether the interviewee subsequently (i) retained the original response, (ii) accepted the alternative offered in the prompt, or (iii) was unable to decide between the alternatives.

In a related study Watson and Moritz (2001) asked the same cohort of students to respond to a chance measurement problem involving a choice between two boxes with the same ratio but different numbers of marbles of two colours. Of 32 students who could improve their level of response, half made no change in response to higher level prompts from other students. Of the other half, 14 improved their level of response with either active or passive agreement with the prompt, whereas 2 students at first agreed with the prompt but then reverted to their original response. It is hence of interest to compare results of a question involving chance measurement with questions related to sampling in terms of students' reaction to other students' prompts.

SAMPLE, PROCEDURE, ANALYSIS

Thirty-seven students answered the first two parts of the protocol in Figure 1: 6 in grade 3, 13 in grade 6, and 18 in grade 9. One grade 9 student who offered a high level response to Part 2 was not prompted. Twenty students answered Part 3 of the protocol and all were given at least one prompt from another student. The prompts were gathered from 92 student responses obtained earlier (Watson, Collis, & Moritz, 1995; Watson & Moritz, 2000). Figure 2 contains eight of the most commonly used prompts, two for sample size, three for sample selection, and three for Part 3. Other prompts suggested "10% or 20%," "10 because it's an even number," and "100 because most surveys are like that" for sample size, and "a few schools, 5 out of each" and "people who have been working well" for method of sample selection.

Students were chosen by their teachers as articulate and willing to be challenged in their mathematical thinking, as it was important to have subjects who would not be intimidated by the experience. It is hence acknowledged that the students could be judged to be of above average intelligence for their grade levels. They were interviewed by a researcher in a quiet setting away from the classroom and the sessions were video taped. After each of Parts 2 and 3 of the protocol students were shown digitised extracts on a laptop computer of other students answering the question in a different fashion than the student had done. A hyper-text transcript of the prompt

was also on the screen allowing the student to read it during or after viewing the video. The prompts will be referred to by the pseudonyms for the students giving the prompts. After each prompt the interviewer asked, "What do you think of her/his idea?"

<table>
<tr><td colspan="3">Sample Size: Here are some comments from other students about how many people they should pick.</td></tr>
<tr><td colspan="2">Martine said
S: About 40.
I: Why would you say that?
S: With 10 children, they could be fat, skinny... they could have a lot more or a lot less weight, than more children put together would.
I: And if you take 40, what happens then?
S: There's more children, so there'd be more the same weight.</td><td>Michael said
S: They should use a lot... most of the Grade 5 children in Tasmania.
I: Is there any advantage to only doing a few, rather than looking at them all?
S: Well you wouldn't know the average, would you, if they only sampled a few? So probably better off with the whole.</td></tr>
<tr><td colspan="3">Selection: Here are some comments from other students about how they should choose the people.</td></tr>
<tr><td>Jane said
S: If they saw a really small person, bring them up and then measure some of the other members of the class with them that they also thought were small. And if the other people are around their height, they could choose them. And if there was a really tall person, they could measure some of the others with them, and they're around that height, then you could take those 2.</td><td>Jessica said
S: They shouldn't choose them all from the same primary school. I'd go to different areas, country South Australia and city South Australia, and different wealth and backgrounds and things.</td><td>Jack said
S: Say take a kid from each school. Just take some--just pick a kid from random order. Look up on the computer, don't even know what the person looks like, and just pick that person.</td></tr>
<tr><td colspan="3">Part 3: Here are some comments from other students about which sample is more likely to have more than 80% boys.</td></tr>
<tr><td>Brian said
S: The large sample would have more chance of being the one with the more boys because there's more children to choose the boys from.</td><td>Bert said
S: I think you would be more likely to get it from the country school, because there are a lot less children, so if you had perhaps a few more, you would bring the percentage up a lot quicker than with this sample because it takes a lot more children to get 80%, so I think it would be more likely to come from the country school.</td><td>Barbara said
S: I think it's equally likely because it was a random sample but also that it's a percentage not a definite number, so it could have been from any of the schools.</td></tr>
</table>

Figure 2. Prompts Used to Create Cognitive Conflict

The digitised video of each interview was linked via a spreadsheet both to the typed transcript of the dialogue of student with interviewer and through it to the other video clips that were shown to the student during the interview. This provided an efficient basis for the analysis. Based on Part 2 of the protocol, the initial student responses were classified according to the hierarchy devised by Watson and Moritz (2000), except that the highest category, Large Samplers Sensitive to Bias (L6), was not used. This was because the biased newspaper articles were not presented to students as part of the protocol. Further, it was possible to distinguish within the Large Samplers group, those who suggested either random or stratified methods only (L5a), or both (L5b). Part 3 of the protocol was analysed separately.

Table 1
Levels of Response for 30 Consistent Students on Sample Size and Selection

Grade	Small None (L1)	Small Prim Random (L2)	Small Pre Select (L3)	Equivocal (L4)	Large Ran or Strat (L5a)	Large Both (L5b)
3	2	0	1	0	0	0
6	1	2	3	1	2	2
9	0	1	0	4	8	3*

* one student not prompted

RESULTS

With five levels of appreciation of sampling for Part 2 of the protocol, 15 students responded at L5 but only 5 of 37 initially responded as Large Samplers with both Random and Stratified Selection (L5b). This meant that 32 students had an opportunity to improve their category of response following a prompt from another student. In fact only seven students did so.

A summary of the levels of response, by grade, is given in Table 1 for the 30 students who did not change their views during the interview (including the five in the highest category).

Two students at both grades 6 and 9 and three at grade 3 improved their responses after hearing one or more prompts from other students. One grade 3 girl was a Small Sampler with an Idiosyncratic method (L1), who agreed with both Michael and Jack but with primitive reasoning.

S: I'd just choose like the people who were working the most, best. I'd find out that they would know a lot and the people that weren't working don't know a lot, so I'd choose about like 4 people that are working, 1 person that isn't working. *Michael's prompt.* Yes I think he's got a good idea because some people would miss out and they'd feel that they'd have to do work so I think that's a good idea. *Jack's prompt*. That's a really good idea just looking them up on the computer and say they just don't want to see their picture and just say one person was a horse rider or really intelligent and never forget to do anything … and I don't want to see what they look like, just pick them.

Another grade 3 moved from the Small Sampler with No Selection group (L1) to the Equivocal group (L4), suggesting a sample size of 30 (after hearing Martine's suggestion of 40) but no other selection methods. A third grade 3 who was a Large Sampler with some Stratified Selection (L5a) took on the idea of Random selection (from Jack) to move to the highest category (L5b).

A grade 6 student also changed from the lowest to the Equivocal category (L4) but stayed with a small sample size (rejected Michael's and Martine's ideas) while accepting random and stratified methods. A second grade 6 moved from the Small Sampler with Preselection category (L3) to the Large/Stratified category (L5a). A grade 9 moved from the Small Sampler with Preselection group (L3) to the Equivocal group (L4) by rejecting random methods and accepting a sample size of 40 (Martine's idea) and some stratification (Jessica's idea) but retaining a view to select for "appearing average." The other grade 9 was a Large Sampler with Random Selection (L5a) who took on stratification (Jessica's idea) to reach the highest group (L5b).

As 25 of the 32 students (78%) did not change their views when prompted with more statistically appropriate responses, it is of interest to look at some of the reasoning displayed in defending their initial views. One grade 3 student initially suggested a sample size "about half of 10, around that, about 7 and under", and as a selection method, "they should choose them if they are large or small or just medium, they would choose like an amount of each one" (L3). He was then shown Martine's suggestion.

S: Yeah well that's true except if you had 40 there would be a bit more variety, but if you had 40 people it would still be a waste of time because you would be measuring their weight and heaps of them would probably be the same or just under. Yeah so I reckon 7. *Jack's prompt.* Well that would probably, like going around picking a few kids from each school would probably be a good idea, just choosing randomly would be good I suppose, so yeah I reckon that's a pretty good idea. *I: [Is it] better to not know what the person looks like ... or better to ... pick some small and some big and some medium size?* I would stick with mine just to like keep, like pick some people that were small sized, then medium, then large.

Hence although considering the alternative this student stayed with his own idea. This was typical of several students in different grades. A grade 6 boy who was also a Small Sampler with Preselection (L3) suggested the following, including his personal experience.

S: Well they choose an average of like 5 tall kids, 5 middle kids, 5 small children, yeah. *I: And why should they do it like that?* Well you wouldn't be correct if you chose all the heavy kids because like it wouldn't add up all the way around there, I don't reckon, if you put all the small ones in. So I reckon it should be around, like there's a lot of kids in our class about 30/35 and a lot of kids about 40 and a few kids like 50 or over, so I reckon like the average would be about my weight, about 40-41. *Michael's prompt.* Well I reckon you could do a lot of them but not like every single one, because like then if you get every single one I reckon practically it wouldn't make any difference all up. *Jack's prompt.* Well if you did random order you don't actually know like you could get all heavy kids or all the really light kids. He's thinking that he doesn't really know so you have all, really 30, and just under 30 and really heavy, you could just get all the heavy people and then you'd think okay the average is 50, like every 30 kid is … oh geez I must be skinny then. *I: So you*

think you would be better off to know[what the person looks like]? No you just need to know the weight, like it doesn't matter about what they look like.

Twenty students were asked Part 3 of the protocol. Of the five grade 6 students none were correct and none changed their minds after prompts. Of the 15 grade 9 students, three changed their minds (one from (a) to (=), one from (a) to (b), and one from (=) to (b)). Four students initially said (b), of whom two wavered when prompted with the three responses in Figure 2. Eight others were consistent in saying (=) before and after prompts.

Of the 13 students who initially chose (=) only one was persuaded by another student's prompt to change to (b). It is of interest then to consider some of the comments made by students in response to the prompts. Ten of the students who did not change their minds were shown the (a) and (b) prompts in Figure 2 in the order *Brian* then *Bert*. Typical of students with a strong belief in percent as an absolute concept was the following grade 6 girl. The following exchange was the initial discussion of the question.

S: Both I suppose are likely because, it depends why they, how they chose the sample… *I: If they were randomly chosen would that make it?* It doesn't really matter because 80% it doesn't matter whether it comes from 1000 or 20 it is still 80% and so it could have come from either. *Brian's prompt.* It was randomly chosen so there is more children to choose the boys from, but it was a randomly chosen sample so it could equally have come from the small sample and you could argue that same thing about the small sample because it is 80% not like 15 children, yeah. *Bert's prompt.* It would, it would bump the percentage up a lot, but 80% is 80% it's not like I said 15 people so it could easily come from the [city] school.

Other students did not have as firm views on percent but appeared to base their arguments on the balance of girls and boys in the schools. The following grade 9 boy responded as follows to the question and prompts.

S: They are both equal. *I: And why do you say that?* Because they've both got an even number of boys and girls. *I: And so because they are drawing from that they are equally likely to get 80% boys?* Yes. *Brian's prompt.* No, because once again there's more girls to choose from as well, so it doesn't really work. *Bert's prompt.* That makes sense but well one child would be worth 5% so you would need 16 boys, yeah I still think there's an even chance.

It is interesting that one grade 9 boy chose (b) for an intuitive reason, wavered after the first prompt, but in end accepted the correct argument with confidence.

S: I would say the small sample of 20 from the country school. *I: And why do you say that?* Well because it's small and there might not be that many girls there. *I: Yes, so does the fact that it's small mean that, or the fact that it is country, or a bit of both?* A bit of both. *Barbara's prompt.* Actually that makes it a bit more equal now. It's making me a bit towards her answer now. *I: Which bit do you like about hers, that it's a random sample, or that it's a percentage not a definite number?* Both of them. *I:Both of those, so she's changed your mind, or?* I'm not sure. *I: Well we'll have a look at what some other students have said. Brian's prompt.* I don't like that one, no. *I: No?* No, there's probably the same amount of girls there as well. *Bert's prompt.* Yes I agree with him, yes.

The other student who wavered in belief started off with stronger reasoning and ordered the likelihood of his answers, first (b) and then (=).

DISCUSSION

The discussion will be presented in two parts: what is learned about students' understanding of sampling and willingness to accept conflicting views, and the use of the methodology associated with the research. Although it might be considered discouraging that so few students appeared to improve their understanding after prompting, especially for Part 2, it was encouraging that none, when offered views of a less appropriate nature, accepted them. The fact that a smaller percent of students raised their levels of response than in the study of Watson and Moritz (2001), raises the question of the topics that are the foci of the protocols. The question in Watson and Moritz involved ratio in a closed probability problem as did Part 3 of the current protocol; but Part 2 involved no calculations and hence students may have felt more confident in defending their opinions in a descriptive setting. The comments of the first student about choosing all available for a sample so no one felt left out reflects views on sampling reported by

Jacobs (1999). As in other studies of problems like Part 3 (e.g., Kahneman & Tversky, 1972) the most popular distracter was the (=) response. Some students had a basic understanding of percent but had difficulty distinguishing 80% as an abstract concept from 80% of 50 = 40 as a concrete quantity. This points to the need to do more applications with percent when it is introduced in the middle school. Also some who chose the (=) option appeared to stick with it due to the presence of half boys and half girls in the schools. There is need for sampling activities where proportions in the population are equal but samples may not represent this equality in the short term.

The methodology for this study was chosen to reflect some characteristics of classrooms where students are able to exchange conflicting ideas and the lack of opportunity to include hands-on sampling tasks during the interview may be a short-coming. Although genuine debate was not possible, after each comment from another student the interviewer asked for an opinion on the comment and sometimes followed up by clarifying the implications of the prompt. Students being interviewed knew they were going to be presented with alternative views to their own; whether they were more or less defensive of their own views in an interview setting than they would be in a classroom is not known. With the general lack of interest in taking on others' views on the topic of sampling, there may be implications for teachers in this descriptive area of the curriculum: they need to be aware that more than stated views are likely to be required to get students to accept higher level views. Exposing students to views of "equals" rather than an authority figure gives a better idea of the acceptability of the views expressed. If expressed by a teacher, the views might be accepted without full understanding and later forgotten or perhaps not applied appropriately.

Although not replacing research into classroom interactions, for example including experimentation with sampling, the methodology of this study has allowed for the examination of student reaction to higher and lower level views. It offers promise for those investigating student change in the face of cognitive conflict in other areas of statistics. In the area of sampling, it is encouraging that no students accepted a less viable explanation than their own, and although only 22% who could improve their level of response on Part 2, did, the study points to the feasibility of allowing student interaction with each other in the classroom. Supplemented with teacher intervention and planned sampling activities, the overall benefits are likely to be positive.

Acknowledgements. Jonathan Moritz conducted the interviews and set up the spreadsheet environment in which the analysis took place. Judith Deans prepared the transcripts. The research was funded by the Australian Research Council, grant number 79800950.

REFERENCES

Jacobs, V.R. (1999). How do students think about statistical sampling before instruction? *Mathematics Teaching in the Middle School, 5*, 240-246, 263.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgement of representativeness. *Cognitive Psychology, 3,* 430-454.

Shaughnessy, J.M. (1985). Problem-solving derailers: The influence of misconceptions on problem-solving performance. In E.A. Silver (Ed.), *Teaching and learning mathematical problem solving: Multiple research perspectives* (pp. 399-415). Hillsdale, NJ: Erlbaum.

Strike, K.A., & Posner, G.J. (1992). A revisionist theory of conceptual change. In R.A. Duschl & R.J. Hamilton (Eds.), *Philosophy of science, cognitive psychology, and educational theory and practice* (pp. 147-176). Albany, NY: State University of New York Press.

Watson, J.M. (1997). Assessing statistical literacy using the media. In I. Gal & J.B. Garfield (Eds.), *The assessment challenge in statistics education* (pp. 107-121). Amsterdam: IOS Press and The International Statistical Institute.

Watson, J.M., Collis, K.F., & Moritz, J.B. (1995). *The development of concepts associated with sampling in grades 3, 5, 7 and 9*. Paper presented at the Annual Conference of the Australian Association for Research in Education, Hobart, Tasmania.

Watson, J.M., & Moritz, J.B. (2000). Developing concepts of sampling. *Journal for Research in Mathematics Education, 31*, 44-70.

Watson, J.M., & Mortiz, J.B. (2001). The role of cognitive conflict in developing students' understanding of chance measurement. In J. Bobis, B. Perry, & M. Mitchelmore (Eds.), *Numeracy and beyond* (pp. 523-530). Sydney: MERGA.