A HIERARCHY OF TERTIARY STUDENTS' CONSIDERATION OF VARIATION

Jackie Reid and Chris Reading
University of New England, Australia
jreid@mcs.une.edu.au

In this study a hierarchy of consideration of variation was developed from students' responses to a questionnaire given at the beginning of a tertiary introductory statistics course. The hierarchy was then used to code responses to the same questionnaire post-study. Comparison of student performances showed that the development of consideration of variation differs with the context of the question. The proposed hierarchy could provide a basis for a more general hierarchy of consideration of variation that is applicable across a variety of tasks. It also supports educators in identifying the level of a student's consideration of variation, providing direction for teaching and learning activities that will help that consideration develop.

INTRODUCTION

Wild and Pfannkuch (1999) identified consideration of variation as one of five fundamental types of statistical thinking. Other recent research also emphasized the importance of a sound understanding of variation to the development of students' statistical thinking at all levels of education (e.g., tertiary: Meletiou-Mavrotheris and Lee, 2002; primary and secondary: Reading and Shaugnessy, 2004; Torok and Watson, 2000). Curriculum developers need to be aware of the level of students' consideration of variation, what helps the development of that consideration, and ways of measuring that development. At the tertiary level, there has been little research into the development of students' consideration of variation (Reid and Reading, 2005), and this paper seeks to redress this by describing the development of a hierarchy of consideration of variation. Hierarchies have previously been used for coding student responses at the primary and secondary level (e.g., Torok and Watson, 2000; Watson, Kelly, Callingham and Shaughnessy, 2003).

STUDY BACKGROUND

The *Understanding of Variation* project (Reading and Reid, 2004) aimed to assess the development of students' understanding of variation, and identify those teaching strategies that assisted that development. The project focused on students enrolled in an introductory tertiary statistics course that incorporated variation as a core concept. Hierarchies were developed to assess students' consideration of variation, based on the analysis of responses to various learning activities including questionnaires, minute papers (Reading and Reid, 2004), assignment questions and class tests (Reid and Reading, 2005). This study focused on student responses to the questionnaires and was designed to address the following. What development of consideration of variation do tertiary students demonstrate after completing an introductory statistics course? How can this development be measured? A questionnaire (Figure 1) was developed, focusing on variability (Q1), comparing data sets (Q2), sampling (Q3 and Q4) and probability (Q4).

METHODOLOGY

The same questionnaire was given to students at the beginning (pre-study) and end (post-study) of the course. Thirty-two of the students agreed to participate in the study and had their pre-study responses independently analysed by two researchers. A hierarchy was developed from these responses. Any discrepancies in coding were discussed, helping to refine the hierarchy. Other variation-related hierarchies (e.g., Reading and Reid, 2004; Torok and Watson, 2000) also informed this development. The proposed hierarchy was used to code the post-study responses completed by 23 of the students. The pre- and post-study levels of consideration were compared.

RESULTS

Developing the Hierarchy

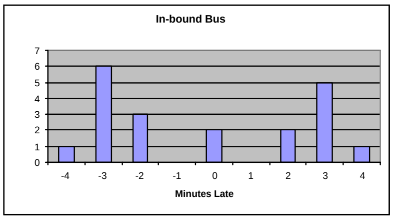
The hierarchy (Figure 2) consisted of four levels: *no*, *weak*, *developing* and *strong* consideration of variation. A summary of typical responses follows.

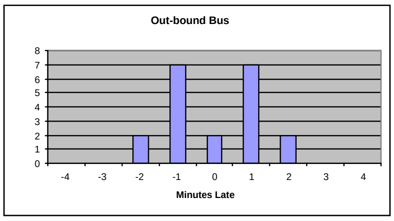
Question 1: What does variability mean to you? (variability)

Two students failed to provide any indication of variability since neither was familiar with the term. Weak responses referred to the concept of change, (e.g., "... a consequence or action is subject to change"), but lacked any clarifying examples. Developing responses included at least one example, while a typical strong response considered different sources of variability (e.g., within- and between-group variation).

Question 1 - What does variability mean to you? Give a verbal explanation and/or an example.

Question 2 - Citizens in an outer suburb were concerned about the reliability of their bus service to the centre of the city. They monitored the in-bound and out-bound service of the buses at Bus Stop 33, and recorded the number of minutes late. Zero minutes late indicates the bus was on time while a negative number of minutes late indicates the bus was early. The data are displayed in the two graphs. Describe and compare the performances of the two bus routes.





Question 3[*] - Every year in New Zealand approximately seven children are born with a limb missing. Last year the children born with this abnormality were located in New Zealand as shown on the map. In New Zealand, it is common knowledge that one-third of the population lives in the top region and one-sixth of the population in each of the other four regions. What do you think? * *Meletiou-Mavrotheris and Lee (2002)*

Question 4[**] - A bowl has 100 wrapped hard candies in it. 20 are yellow, 50 are red, and 30 are blue. They are well mixed up in the bowl. Jenny pulls out a handful of 10 candies whilst blindfolded, counts the number of reds, and tells her teacher. The teacher writes the number of red candies on a list. Then, Jenny puts the candies back into the bowl, and mixes them all up again. Five of Jenny's classmates, Jack, Julie, Jason, Jane and Jerry do the same thing. They each pick ten candies, count the reds, and the teacher writes down the number of reds. Then they put the candies back and mix them up again each time.

a. What do you think the teacher's list for the number of reds is likely to be? Explain why you chose those numbers.

b. If you were asked to choose a response to this question from the following list, circle the one that you would choose. Explain why you chose that one.
A) 5,9,7,6,8,7 B) 3,7,5,8,5,4 C) 5,5,5,5,5,4 D) 2,4,3,4,3,4 E) 3,0,9,2,8,5

c. All the students in Jenny's class watched the demonstration and wanted to take part. The teacher began the trial again, recording the results in a new list, allowing each student in the class of 40 to draw out 10 candies under the same controlled conditions. Describe a list that the teacher would have been likely to record. Explain why you described the list that way. ** *Adapted from Reading and Shaughnessy (2004)*

Figure 1: Questionnaire

Question 2: Bus performance (comparing data sets)

Weak responses focused on a measure of location with reference to a basic measure of spread, e.g., the range. Developing responses quantified the variability in some way and used that for comparison of the two bus performances. Strong responses gave more detail about the distributions, for example comparing quantiles or proportions.

Question 3: Distribution of limb-less births in NZ (probability)

Weak responses accepted the sample, based on one year's data, as representative, and tried to explain the distribution of limb-less births without considering the distribution of the NZ population. Developing responses questioned the difference between the observed data and expectations, based on the NZ population, assuming the births were randomly distributed. Better developing responses acknowledged the need for further data to establish a relationship between regions and birth defects. No responses were coded as strong, but it is anticipated that responses at this level would suggest a need to test for a relationship.

Question 4: Coloured Lollies (sampling and probability)

Some responses did not allow for any variation about the expected outcome, giving C as their response to Q4b. Weak responses allowed for too little or too much variation, (e.g., D or E in Q4b). Developing responses provided a reasonable amount of variation, although not necessarily centered correctly, (e.g., A in Q4b), while strong responses correctly used proportional reasoning to centre the sampling distribution.

No consideration of variation	
Q1:	do not consider any sources of variation
Q2:	may refer to a measure of centre, but not to any measure of spread
Q3:	do not acknowledge any variation about the expected values
Q4:	do not acknowledge any variation about the theoretical or expected outcomes
Weak consideration of variation	
Q1:	discuss one source of variation but expression is poor
Q2:	refer to the range and/ or basic description of shape
Q3:	acknowledge variation and expectations are articulated but not based on given data. Look for extraneous causes of variation
Q4:	allow for variation but amount suggested is low or high. Causes given are extraneous
Developing consideration of variation	
Q1:	describe clearly one source of variation (within-group, between-group, controlling factors, measurement error)
Q2:	refer to measure of location and more detailed description of spread
Q3:	consider variation between expected and observed values and/or identify need for a larger sample or more information
Q4:	provide a realistic amount of variation, but may not be centered correctly. Reasoning may be based on frequencies rather than proportions
Strong consideration of variation	
Q1:	describe clearly more than one source of variation
Q2:	provide further information about the distribution, such as explicit proportions
Q3:	not described since no response coded at this level
Q4:	provide a realistic amount of variation, and proportional reasoning is correctly used

Figure 2: Hierarchy of Consideration of Variation

Applying the Hierarchy to Post-study Responses

Post-study responses were often more sophisticated with respect to the terminology used and the specific references to numerical summaries and statistical tests, although this was not necessarily indicative of true understanding. There was more of an emphasis on measuring and modeling variation rather than simply describing variation, e.g., "the sample will have a mean and a series of values distributed about that mean. The distributed nature of the data is its variability." However, the descriptors for each level were such that they allowed for both formal and informal statistical language. Regardless of the level of language used, or emphasis reflecting the course content, responses still had to indicate an improved understanding of the concepts to be coded at a higher level.

Comparing Student Performance

Students' level of consideration of variation not only varied from question to question, (e.g., one student's responses ranged from *no* to *strong*), but also from the beginning to the end of the course. Table 1 displays the percentage of responses for each level in the pre- and post-questionnaire. Notably, there was a reduction in the proportion of students showing no consideration of variation in Q1, Q2 and Q3. Only Q1 and Q3 demonstrated a marked shift towards better quality responses.

Table 1: Percentage of responses for each level (Pre *n*=32; Post *n*=23)

	Q1 (%)		Q2 (%)		Q3 (%)		Q4 (%)	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post
No	6	0	13	0	19	13	6	13
Weak	56	30	25	65	62	52	50	52
Developing	25	52	56	35	19	35	38	26
Strong	13	17	6	0	0	0	6	9

Table 2 shows the change in levels at the end of the course, where post-study responses had; improved (higher), remained the same (same), or reduced in level (lower). The association between change in level and question is marginally significant (χ^2_6 = 10.63, p=0.1). More than

40% of students demonstrated an improved consideration of variation by the end of the course for Q1 and Q3. However, this is not the case for Q2 and Q4 where more than 25% of post-study responses were coded at a lower level for these 2 questions.

Table 2: Percentage of responses for each type of change in level (n=23)

	Q1 (%)	Q2 (%)	Q3 (%)	Q4 (%)
Lower	22	26	8	30
Same	30	57	48	52
Higher	48	17	44	17

DISCUSSION

Students need to develop a sound consideration of variation. The proposed hierarchy focuses on coding responses to a questionnaire that addresses variation in different contexts, and measures the development of this consideration. Although it would be hoped that this consideration improves by the end of the course, it has been shown that the development is not linear and varies with context. For example, there was an increase post-study in the proportion of students who did not allow for any variation in their response to Q4. This result is consistent with that observed by Reading and Shaughnessy (2004), who used the same question with pre-tertiary students and noted that the tendency to allow for little or no variation is “stronger among older mathematics students.” Students at higher levels tend to focus on expected values, because of a curriculum emphasis on theoretical probabilities. If a large proportion of responses are coded at a lower level following learning experiences, assuming test item validity, further investigation is needed to determine whether elements of the curriculum have had a negative impact. This demonstrates the importance of the hierarchy in curriculum development.

A useful hierarchy has evolved that can be successfully applied at different stages of the course. Although the descriptors for each level are specific to the questionnaire, this hierarchy can contribute to the evolution of a more general hierarchy which can be applied to a diverse range of learning tasks (e.g., Reid and Reading, 2005). This will benefit researchers and educators by identifying the level at which students are operating and providing direction for teaching and learning activities that will help further develop students’ consideration of variation.

ACKNOWLEDGEMENT

Research funded by a UNE Faculty of the Sciences Internal Research Grant.

REFERENCES

Meletiou-Mavrotheris, M. and Lee, C. (2002). Teaching students the stochastic nature of statistical concepts in an introductory statistics course. *Statistics Education Research Journal*,1(2), 22-37.

Reading, C. and Shaughnessy, M. (2004). Reasoning about variation. In D. Ben-Zvi and J. Garfield (Eds.), *The Challenge of Developing Statistical Literacy, Reasoning and Thinking,* (pp. 201-226). Dordrecht: Kluwer Academic Publishers.

Reading, C. and Reid, J. (2004). Consideration of variation: A model for curriculum development. Paper presented at the International Association for Statistical Education Roundtable on Curricular Development in Statistics Education, Lund, Sweden, 28 June to 3 July.

Reid, J. and Reading, C. (2005). Developing consideration of variation: Case studies from a tertiary introductory service statistics course. *Proceedings of the 55th Session of the International Statistical Institute*, Sydney.

Torok, R. and Watson, J. (2000). Development of the concept of statistical variation: An exploratory study. *Mathematics Education Research Journal*, 12(2), 147-169.

Watson, J., Kelly, B. A., Callingham, R. A., and Shaughnessy, J. M. (2003). The measurement of school students’ understanding of statistical variation. *International Journal of Mathematical Education in Science and Technology,* 34(1), 1-29.

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