Developing assessment items to measure tertiary students' reasoning about explained and unexplained variability

Jackie Reid (School of Science and Technology), Dr Chris Reading (School of Education) and Bernard Ellem (School of Science and Technology)

Abstract

As one of the principal forms of fundamental statistical thinking, consideration of variation impacts on all aspects of statistics. There has been extensive research about students' reasoning about variation and numerous assessment items have been developed to assess students' reasoning about many aspects of variability in data. However, there has been little research focusing on helping tertiary students model variability as a combination of explained and unexplained variability, and few items have been developed to assess this construct. An instrument is needed to assess this reasoning, before and after learning experiences, so that the change in level of reasoning can be measured. This can then inform curriculum development and the design of learning experiences.

This paper presents a measurement instrument that was developed to assess tertiary students' intuitive reasoning about variation, focusing on explained and unexplained variability. The theoretical framework that provided the structure for the development of assessment items is discussed, followed by presentation of the results from a pilot study designed to test newly-developed items, and lastly the inclusion of previously-developed items is justified. Analysis of the responses led to clarification of those items that were misinterpreted by students or did not measure the relevant construct. The refined assessment items will be incorporated with previously-developed items in a measurement instrument to be administered to tertiary students enrolled in an introductory service statistics course. The direction for future research is also presented.

Introduction

> "The idea of data as signal and noise, …, is perhaps the most fundamental conceptual model for reasoning statistically."
>
> Konold & Pollatsek 2004:196

> "If our primary goal is to teach statistical thinking, rather than statistical techniques, then we should look to the noise, and not the signal."
>
> Gould 2004:15

Recently, there has been an increase in research on developing students' understanding of fundamental statistical concepts that lead to inference, such as; the Central Limit Theorem, sampling distributions, confidence intervals and the p-value, as well as research that emphasises the need to help students develop an appreciation of stochastic thinking (Meletiou 2000; Pfannkuch 2005; Wild & Pfannkuch 1999). The importance of "an understanding of how variation arises" and of the "uncertainty caused by unexplained variation" for inference is widely acknowledged (e.g., Pfannkuch & Wild 2004:18). However, the concept of total variation as a combination of explained and unexplained variation is rarely emphasised in introductory statistics courses. There is also little research into students' reasoning about the statistical model as a representation of the total variation in the data combining explained (signal) and unexplained (noise) variation. This lack of research is not surprising since, as Shaughnessy (1997, cited in Meletiou, 2000:10) indicated, education research often focuses on themes from curriculum content.

Traditionally, the concept of the mean as a representation of the signal in "the 'noise' of individual data points" has not formed part of the statistics curriculum (Konold & Pollatsek 2004:170). Hammerman and Rubin (2003:18) argued that students often struggle to accept the mean as a representative value of a distribution, and need to be helped to "conceptualize the mean as a relatively stable measure of the signal within the noise". A random phenomenon produces uncertain individual outcomes but exhibits a regular pattern of outcomes after many repetitions (Moore 2003:283). This duality of uncertainty and pattern causes difficulties for many students (Metz 1997:225), and there are many examples in the literature demonstrating students' tendencies to focus on either the pattern (signal) or the uncertainty (noise) rather than an integration of the two (e.g., Hammerman and Rubin 2003:20; Metz 1997:226). In Hammerman and Rubin's study, teachers, new to the ideas of statistical analysis, were given the opportunity to explore various data sets. They often focussed on either the signal (most frequently) or the noise, to the exclusion of the other. For example, if an obvious trend (signal) could be found in the data, the teachers would not pay attention to the variability. When an individual focuses on the pattern without consideration of the underlying uncertainty the available information is given exaggerated weight. Conversely, a focus on the uncertainty inherent in a phenomenon to the exclusion of pattern recognition will underestimate the importance of the available information (Metz 1997:226). Some students even have difficulty distinguishing signal and noise as separate constructs (Bakker 2004:166; Kelly, Sloane & Whittaker 1997:89). This preference for deterministic thinking has often been observed (e.g., Hoerl, Hahn & Doganaskov 1997, cited in Meletiou-Mavrotheris & Lee 2002:23; Kahneman, Slovic & Tversky 1982, cited in Metz 1997:230; Pfannkuch & Wild 2000:147). Furthermore, as Hammerman and Rubin (2003:3) observed, "…one way to reason in the context of variability is to break up a data set into subsets about which one can reason deterministically, even if such a technique ignores the power of statistics to understand variability in data."

The importance of encouraging a view of measures of centre as the signal in a noisy process is the main argument made by Konold and Pollatsek (2004). Others (e.g., Biehler 1994) have argued that the distribution should remain the primary focus of analysis and that the average is only one of its many properties. However, Konold and Pollatsek (2004:171-172) claimed that "the central idea of distribution comes into better focus when it is viewed as the 'distribution around' a signal" and that the "most basic questions in analyzing data involve looking at group differences to determine whether some factor has produced a difference in the two groups. Typically, the most straightforward and compelling way to answer these questions is to compare averages."

This comparison, however, cannot be done in isolation. Any meaningful comparison must compare the amount of explained variation modelled by the signal with the unexplained/ random variation, that is, the noise. Although Konold and Pollatsek's paper (2004) addressed the idea of the statistical model, and they claimed (2004:173) that "the notion of an average understood as a central tendency is inseparable from the notion of spread", their work focused on helping students find the signal amongst the noise, without emphasising the need to consider both when dealing with comparisons of groups. For example, their argument (2004:18) that "viewing an average as a central tendency provides a strong conceptual basis for… using averages to compare groups" fails to make explicit the need to view the differences in means in the context of the variability inherent in the distributions. Recognition of the impact of variability on the appropriate use of an average as a representative statistic is considered for non-standard cases (2004:188-190), for example, when variability is "enormous", but even then it is unclear how students would define "enormous", or even if they would have the statistical reasoning skills to think about what may be considered as typical (and consequently non-typical) variability.

Konold and Pollatsek's work (2004) made a strong case for teaching the idea of statistics as a study of noisy processes, with the view to detecting signals in those processes, but they do acknowledge that the focus of their analysis is on measures of centre (2004:171) . There are some valid reasons why curriculum emphasis has been on how means change. Historically, there is a sound theoretical basis for simple models where only the means change when moving from group to group. Furthermore, inference for means is much more robust under departures from the normality assumption than other distributional elements such as variation (Wild 2006). However, although this focus avoids some of the more complex modelling issues for beginning students, it is important that students develop the ability to model both explained and unexplained variation.

Future research should extend Konold and Pollatsek's work to give more emphasis to helping students not only view the mean as a measure of the signal in a noisy process but also to make explicit the need to link the two components of signal and noise (i.e., explained and unexplained variation) when comparing groups.

Defining a statistical model of variability

What is meant by the statistical model? The term "statistical model" can encompass many notions. Some take a broad view of the concept. Wild and Pfannkuch (1999:230), for example, defined their use of the term statistical model as "all of our statistical conceptions of the problem that influence how we collect data about the system and analyse it." However, in an introductory statistics course it is important to reinforce the links between variation, statistical models representing the total variability in data, and inference (informal or formal). Consequently, a *statistical model of variability* is defined here as a mathematical construct used to model the total variability in data, which is hopefully a "relevant [and useful] abstraction of the reality" (Pfannkuch & Wild 2000:138). Total variation in the data can be modelled as a combination of:

explained variation + *unexplained variation*

Alternative representations for modelling variability (explained and unexplained) are variously described as deterministic and stochastic; systematic and random variation; special-cause and common-cause variation (Wild & Pfannkuch 1999:239) or between-group and within-group variation. These different representations provide slightly different perspectives for students, some more useful than others. In particular, referring to the unexplained or random variation as 'noise' may lead students to consider it as a nuisance, with the possibility that they may focus purely on the signal, without due consideration for the variability in the data. Gould (2004:7) provided useful examples highlighting models where consideration of the variation in the data is of at least equal importance as the identification of a simple trend (signal), for example, in time series data and principal component analysis. Other terminology may also cause problems for students. Terms such as “random” or statements such as “the observed difference is due to chance” may confuse students who believe that, when dealing with biological data, for example, every observed difference can be attributed to some environmental impact (Wild and Pfannkuch 1999:241). Students find the concept of “unexplained individual variation” more accessible than that of “random variation” (Moore 1999: 252). The concepts of explained and unexplained variation form a sound basis for exploring and modelling variation (Wild & Pfannkuch 1999:242).

Furthermore, ability to model variability is fundamental to all stages of the investigative cycle: Problem, Plan, Data, Analysis, and Conclusions (Wild & Pfannkuch 1999:226). The modelling of variability needs to be emphasised throughout the curriculum (e.g., in exploratory data analysis, probability, sampling distributions and inference) and not just at the analysis stage. For example, during exploratory data analysis, students should be examining the data, asking questions such as: what is explained variation and what is unexplained variation? What are the causes of the pattern in the data? Context knowledge plays a very important role in helping students identify potential causes of explained variation. Commonly, statistics curricula emphasise formal statistical inference with an underlying emphasis on probabilistic thinking, which may discourage causal thinking and the seeking of explanations for phenomena. However, students should be encouraged to look for patterns and trends, to seek causes and explanations for the variation, while recognising that not all variation can be explained (Pfannkuch 2005:88).

One goal for statistics educators should then be to consider ways in which the concept of the model of variability (total variation = explained variation + unexplained variation) can be made a focus of all topics (e.g., exploratory data analysis, probability, sampling distributions and inference) in an introductory statistics course.

Components of reasoning about variability

Wild and Pfannkuch (1999:226) described four components of consideration of variation: noticing and acknowledging variation; measuring and modelling variation for the purposes of prediction, explanation, or control; explaining and dealing with variation; and, using investigative strategies in relation to variation. This was used as a basis for the following framework (Table 1), which describes five components of the construct, “reasoning about variability”, that are considered important if students are to reason about explained and unexplained variability.

Table 1: Components focussing on reasoning about explained and unexplained variability

Component	Description
noticing and acknowledging variability	At the most basic level, students must be able to acknowledge that variation exists in data.
acknowledging variability around the signal	Before students can consider comparisons or even the simplest informal inference, they must recognise variation around a signal (mean, trend etc).
looking for causes	Students often try to look for causes even where none may exist, i.e., where the situation would be considered as purely random. However, looking for and identifying causes can enable students to appreciate the concepts of explained and unexplained variation
controlling causes	Consideration of how one might control various causes of variation enables students to focus on sources of explained and unexplained variability and how they should be included in experimental design and analysis.
linking explained and unexplained variation for inference	The final goal is to link this information in such a way that inference, conclusions or comparisons can be made from the data.

These components are not hierarchical but necessary if students are to be able to model variability as a combination of explained and unexplained variability.

The Study

If reasoning about explained and unexplained variability is important then educators need to be able to measure the development of students' reasoning. A review of the literature revealed many existing items that could be used to assess students' reasoning about variability in general. However, there were few items specifically designed to assess reasoning about explained and unexplained variability. Such items were needed for an assessment instrument to be used for future research that focuses specifically on this reasoning, and includes items that assess the five components listed above (Table 1). Some existing items have been chosen to be included in the assessment instrument but some new items needed to be developed.

The main objective of this study was to develop items that specifically address the last four components in Table 1 (acknowledging variability around the signal, looking for causes, controlling causes, linking explained and unexplained variation for inference) to complement the items already available that focus on students' ability to notice and acknowledge variability in general. Of course, items that address these four components will also require students to acknowledge the existence of variability. This study involved trialling newly developed items with students for possible inclusion in the assessment instrument. Two questions were of interest: did students understand each item; and did the item assess the components of interest? The development of assessment items that focus on reasoning about explained and unexplained variability, and the analysis of student responses to those items, are presented in the following sections.

Methodology

Five items that focussed on students' reasoning about explained and unexplained variability were developed. The items included a number of different contexts and assessed various components of reasoning about variability, but with a particular focus on the last four components listed in Table 1. The following outlines the development of those items.

In a bid to assist the development of students' ideas of the signal-noise perspective of various statistical measures Konold and Pollatsek (2004) suggested using the following tasks: repeated measures; explorations of stability such as drawing multiple samples from a known population, comparing the relative accuracy of different measurement methods, and observing a distribution as the sample gets larger; simulations; group comparisons; and, conducting experiments. Tasks that involve group comparisons, which require students to consider not only explained variability but also to view that variability in relation to the unexplained (within-group) variability, will reinforce the concept of the statistical model of variability.

Other tasks can help emphasise the importance of consideration of the "noise" in the data, rather than simply viewing it as a nuisance that is obscuring the signal. For example, Gould (2004) gave instances where the consideration of the unexplained variability is as important as consideration of the measures of centre. These included a simple example of comparing distributions of blood lead levels in two groups of children, one of which was exposed to lead contamination, while the other was a control group. Examination of the within-group variability raised issues that a simple comparison of means would not, lending support to a possible conclusion of causation. Gould (2004) included further examples that highlight the importance of consideration of the "noise", such as time series analysis, where consideration of the variation may lead to identification of, for example, seasonal trends.

With these various contexts in mind, five items were developed that addressed the various components of reasoning about variability. The data and scenarios used in the development of the items were drawn from different sources. It is recognized that the setting for a question may impact on students' responses. Students may find an item more challenging, not because of the statistical content but because of its placement in an unfamiliar setting. However, the scenarios for the five items (rainfall, waiting times, weight gain, classroom disruptions, and speed of light, respectively) were considered familiar settings for most, if not all, participants in the study. The five items are presented in the Appendix, and Table 2 shows how each item addressed particular components (rows) and contexts (columns). All items addressed the first two components of reasoning about variability (noticing and acknowledging variability, and noticing variability around a signal), while the other three components (looking for causes, controlling causes, linking explained and unexplained variability) were each assessed in at least two of the items.

Table 2: Items by component and context

	Repeated measures	Time series	Group comparisons	Conducting experiments
Noticing variability	Q5	Q2 Q4	Q1 Q3	Q3
Variability around a signal	Q5	Q2 Q4	Q1 Q3	Q3
Looking for causes	Q5	Q4		Q3
Controlling causes		Q4		Q3
Linking explained & unexplained	Q5		Q1 Q5	

The five item questionnaire was administered to science students enrolled in a first year mathematics course. This cohort of 30 students was selected because they had received no statistical education beyond high school and would have a similar background as a typical student enrolling in an introductory statistics course. The questionnaire was given to students in the second half of a lecture, held in the last week of semester. Due to a drop in attendance in that week only five students completed the questionnaire. All students completed the questionnaire within 35 minutes and the questionnaires were collected at that time. The responses were analysed to determine if the students had understood each of the items, and whether each item had assessed the component of reasoning that it had been designed to assess.

Results and Discussion

The analysis of student responses resulted in changes to all items. In the following sections, modifications made to the items, as a result of analysing student responses, are presented. Firstly, cases where an item was revised as a consequence of students having misunderstood the question are considered. Secondly, cases where items were revised because they did not assess the component of the construct that they were designed to assess are discussed. One item was deemed redundant since responses added no further information about students' reasoning evident from responses to the other four items.

Students' Understanding

In most instances student responses indicated that they had understood the questions. However, some responses to questions 2 and 4 suggested otherwise. The ideal response to question 4 would identify the overall pattern as demonstrating a reasonably constant high number of disruptions before intervention, followed by a marked drop immediately after intervention and then a reasonably constant number of disruptions after intervention. The random day-to-day fluctuations would be identified as the variation around that pattern. However, in the response to (b), one student wrote: "*I don't fully understand the question – be more specific".* Furthermore, in some responses to this question it was clear that students identified the random fluctuations as a pattern. For example, one response to Q4(a) stated: *"there is a specific cyclic pattern for each period (before and after)".* The term *pattern* has a more specific meaning for students at the end of an introductory statistics course where it is used to denote systematic or explained variation, but for students with little statistical training *pattern* may be interpreted in various ways. Therefore it was decided to replace *pattern* with *trend*, with the expectation that it may make the question clearer for the students.

The time series barchart in question 2 was modified since some students struggled to interpret the graph. This was a student generated representation of the geyser data and was non-standard. Some students struggled to make sense of it and so it was decided to replace it with a more typical time series plot.

Assessing the construct

In addition to ensuring that each item was correctly interpreted by the students it was also necessary to ensure that each item assessed the component of interest. In question 1, the objectives of the question were to see if students could identify the need for historic data to make relevant comparisons (acknowledging variation around a signal), and if they could outline an informal test of hypothesis (linking explained and unexplained variation). Of interest is*:* how much greater than the historical average should last month's average rainfall be for it to be considered above normal?

Not all students achieved the second of these objectives. Students recognized the need for historical data but did not explicitly outline the decision making process. The replacement of "*why"* with *"explain how you would use that information to check the validity of the statement"* is expected to elicit more detail of the comparison between last month's average with previous years.

Part (b) of question 3 did not elicit the expected responses from students. It was hoped that student responses would consider the inclusion of variables to account for different sources of variability (e.g., different species, environmental conditions). Student responses focussed on sample size or alternatively they suggested eliminating, rather than allowing for, possible sources of variation (for example, ensure the same exercise regime for all rabbits). Consequently, the wording for question 3(b) was modified to encourage students to allow for different causes of variability.

Although Question 5 presented students with graphical representations (dotplots and boxplots) that were different from those in the other items, the questions asked ("*Comment on the variability* …", "*Give reasons for the apparent variability*.", "*Which representation did you find most useful*…") were very similar to those asked in the other items. Responses from students to Question 5 provided no additional information about students' reasoning beyond that achieved by the other 4 items. Taking into account that the items from the pilot study were to form part of a larger questionnaire, Question 5 was discarded to avoid unnecessary duplication.

This study has highlighted the importance of trialling new assessment items. The trial resulted in item refinement, ensuring that each item is correctly interpreted by students and that each one assesses the relevant component of reasoning about variability it was designed to assess. For items that are to be given at the beginning of an introductory course, it is important that the context and terminology are at a suitable level for students who have had no previous tertiary statistical training.

Future Research

A questionnaire that combines the four newly developed items from this study with existing items will be used to measure the development of students' reasoning about variability. The three existing items have been previously used to measure the development of students' reasoning about variability and complement the four developed items. One item taken from Reid & Reading (2006) asks students to define variability, and provide an example. This item will be the first in the questionnaire and gives students the opportunity to think in general terms about the existence of variability in data. The second item, taken from the ARTIST (2007) item bank, asks students to decide if high or low variability is desirable under different circumstances (e.g., age of trees in a national forest, diameter of tyres from a production line). This item highlights to students that variation is not necessarily a bad thing, and hence does not always have to be eliminated. The third item asks students to compare the performance of two bus services (Reid & Reading 2006), and was chosen since it allows a comparison of distributions, where there is no difference in means, but unexplained variability differs and is crucial to the comparison of the two distributions.

The seven item questionnaire will be administered pre- and post-study to students enrolled in an introductory statistics course in second semester 2007. The responses will be coded according to their level of reasoning and a comparison of pre- and post-study codings will allow a measurement of students' development. Of various theories that may be used to explain cognitive growth, the Structure of Observed Learning Outcomes (SOLO) Taxonomy (Biggs & Collis 1982) has been identified as a powerful tool in the assessment of mathematical reasoning (Pegg 2003). It is expected that the SOLO Taxonomy framework (Biggs & Collis 1982) will underpin the coding of student responses. In addition, selected students will be interviewed to gain further insight into their understanding of the question and their reasoning about variability.

A redeveloped introductory statistics course, which has an increased emphasis on modelling variability as a combination of explained and unexplained variability, is planned for 2008. As before, the questionnaire will be administered pre- and post-study to measure the development of students' reasoning. Furthermore, a comparison will be made between the two cohorts (2007 and 2008) to determine whether the redeveloped curriculum provided a better support for students to develop reasoning about variability.

References

Assessment Resource Tools for Improving Statistical Thinking (ARTIST). Retrieved 31 July 2007 from https://app.gen.umn.edu/artist/

Bakker, A. 2004, Design research in statistics education on symbolizing and computer tools. Unpublished Doctoral Thesis, Utrecht University.

Biehler, R. 1994, 'Probabilistic thinking, statistical reasoning, and the search for causes - Do we need a probabilistic revolution after we have taught data analysis?', in *Research Papers from ICOTS4*, ed J. Garfield, University of Minnesota, Minneapolis, pp. 20-37.

Biggs, J.B. & Collis, K. F., 1982, *Evaluating the quality of learning: The Solo Taxonomy*, Academic Press, New York.

Data and Story Library (DASL), Retrieved 31 July 2007 from http://lib.stat.cmu.edu/DASL/

Gould, R. 2004, 'Variability: One statistician's view', *Statistics Education Research Journal*, vol. 3, no. 2, pp. 7-16.

Hammerman, J.K. & Rubin, A. 2003, Reasoning in the presence of variability. Paper presented at the *Third International Research Forum on Statistical Reasoning, Thinking and Literacy*, Lincoln, Nebraska.

Hoerl, R., Hahn, G. & Doganaksoy, N. 1997, 'Discussion: Let's stop squandering our most strategic weapon', *International Statistical Review*, vol. 65, no. 2, pp. 147-153.

Kahneman, D., Slovic,P. & Tversky, A. 1982, *Judgement under uncertainty: Heuristics and biases*, Cambridge University Press, Cambridge.

Kelly, A.E., Sloane, F. & Whittaker, A. 1997, 'Simple approaches to assessing underlying understanding of statistical concepts', in *The Assessment Challenge in Statistical Education*, eds I. Gal & J.B. Garfield, IOS Press, Amsterdam, pp. 55-90.

Konold, C. & Pollatsek, A. 2004, 'Conceptualizing an average as a stable feature of a noisy process', in *The Challenge of Developing Statistical Literacy, Reasoning and Thinking*, eds D. Ben-Zvi & J. Garfield, Kluwer Academic Publishers, Dordrecht, pp. 169-199.

Meletiou-Mavrotheris, M. & Lee, C. 2002, 'Teaching students the stochastic nature of statistical concepts in an introductory statistics course', *Statistics Education Research Journal*, vol. 1, no. 2, pp. 22-37.

Meletiou, M. 2000, Developing Students' Conceptions of Variation: An Untapped Well in Statistical Reasoning. Unpublished Doctoral Thesis, University of Texas.

Metz, K. 1997, 'Dimensions in the assessment of students' understanding and application in chance', in *The Assessment Challenge in Statistics Education*, eds I. Gal & J.B. Garfield , IOS Press, Amsterdam, pp. 223-138.

Moore, D.S. 1999, 'Discussion: What we teach beginners?', *International Statistical Review*, vol. 67, no. 3, pp. 250-252.

Moore, D.S. 2003, Introduction to the Practice of Statistics (4 Ed), Freeman, New York.

Pegg, J. 2003, 'Assessment in mathematics: A developmental approach', in *Mathematical Cognition*, ed J. Royer, Information Age Publishing, Greenwich, pp. 227-259.

Pfannkuch, M. 2005, 'Thinking tools and variation', *Statistics Education Research Journal*, vol. 4, no.1, pp. 83-91.

Pfannkuch, M. & Wild, C. 2004, 'Towards understanding of statistical thinking', in *The Challenge of Developing Statistical Literacy, Reasoning and Thinking*, eds D. Ben-Zvi & J. Garfield Kluwer Academic Publishers, Dordrecht, pp. 17-46.

Pfannkuch, M. & Wild, C. 2000, 'Statistical thinking and statistical practice: Themes gleaned from professional statisticians', *Statistical Science*, vol. 15, no. 2, pp. 132-152.

Reid, J. & Reading, C. 2006, 'A hierarchy of tertiary students' consideration of variation', in Proceedings of the *Seventh International Conference on Teaching Statistics*, [CDROM], International Association for Statistical Education & International Statistical Institute, Salvador, Brazil.

Shaughnessy, J.M. 1997, 'Missed opportunities on the teaching and learning of data and chance', in, *Research Papers on Stochastics Education*, eds J. Garfield and J. Truran, pp. 129-145.

Shaughnessy, J.M. & Pfannkuch, M. 2002, 'How faithful is Old Faithful? Statistical Thinking: A story of variation and prediction', *The Mathematics Teacher*, vol. 95, pp. 252-259.

Teague, D.J. 2006, ‘Experimental Design: learning to manage variability’, in *NCTM 2006 Yearbook: Thinking and reasoning with data and chance*, ed G. F. Burrill,. The National Council of Teachers of Mathematics, Inc., Reston, VA, pp. 151-169.

Wild, C. 2006, ‘The concept of distribution’, *Statistics Education Research Journal*, vol. 5, no. 2, pp. 10-25.

Wild, C. & Pfannkuch, M. 1999, ‘Statistical thinking in empirical enquiry’*, International Statistical Review*, vol. 67, no. 3, pp. 233-226.

Appendix

1) Suppose that a weather report stated that a city's rainfall for last month was "well above normal".

   a) What would you infer from the statement "well above normal"? Explain your response.

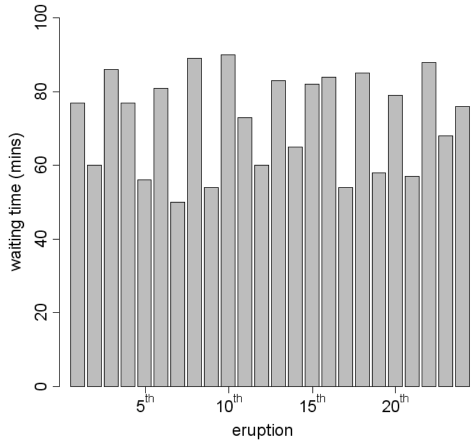
   b) Suppose that you were asked to check the validity of the statement "well above normal". What information would you need and why?

2) Waiting times (in minutes) between 24 successive eruptions of the 'Old Faithful' geyser in Yellowstone National Park were given to students. They were asked to summarise the data.

77	60	86	77	56	81	50	89	54	90	73	60
83	65	82	84	54	85	58	79	57	88	68	76

   (i) Lachlan calculated the mean waiting time between eruptions as 72.2 minutes and stated that the times ranged between 50 and 90 minutes.

   (ii) Chris produced the following graphical summary, where waiting times are plotted chronologically.



   (iii) Stephanie produced the following histogram of waiting times.

a) Comment on the usefulness of each of the summaries for demonstrating the variability in the waiting times.

<u>Lachlan's:</u>

<u>Chris's:</u>

<u>Stephanie's:</u>

b) Which one of the three summaries do you prefer? Explain why.

(adapted from Shaughnessy and Pfannkuch 2002)

3) An experiment is designed to compare the effect of two different diets on the *weight gain* of rabbits. Weight is measured at the age of two weeks and six weeks, and the *gain* calculated. The research question is:

Is the average weight gain greater on one diet than the other?

a) When the measurements are made, not all weight gains will be the same. What might cause the variability in weight gain?

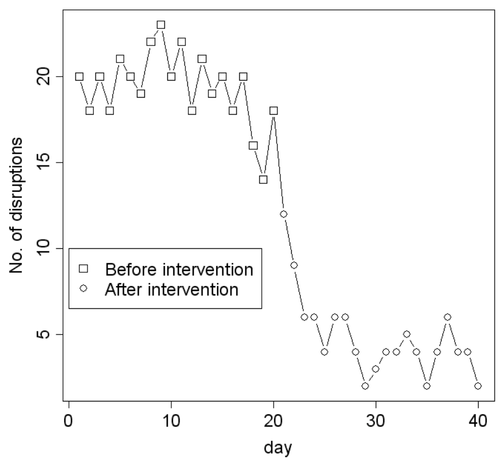
   i) List the most important possible causes of variability. Please limit your list to a maximum of three causes.

   ii) Briefly explain how each cause may affect the variability in weight gain.

b) For this experiment, how would you plan the data collection to take into account the sources of variability listed in (a)?

(adapted from Teague 2006)

4) The daily number of student disruptions in a classroom was recorded for 40 days. Half way through this period, remedial intervention was introduced to try to help reduce the number of disruptions. The figure below plots the daily number of disruptions *(No. of disruptions*) against time (*day*). The different symbols represent data collected either before or after the intervention.



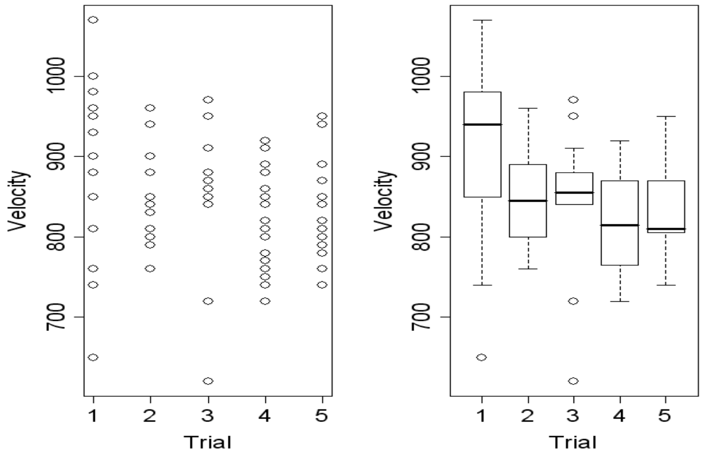
a) Describe any pattern(s) apparent in the plot.

b) Describe any variation from the pattern(s).

(data from Data and Story Library, http://lib.stat.cmu.edu/DASL/)

5) In 1879, A. A. Michelson made 100 determinations of the velocity of light in air using a modification of a method proposed by the French physicist Foucault. These measurements were grouped into five trials of 20 measurements each. The numbers are in km/sec, and have been coded by having 299,000 subtracted from them.

Because the speed of light is a physical constant, we know (to a close approximation) the "true" coded value that Michelson was trying to measure is 734.5. The two plots (A & B) below give different representations of the same coded data.



A B

a) Comment on the variability in the data apparent in the plots.

b) Give possible reasons for the apparent variability.

c) Which plot (A or B) did you find most useful for commenting on the variability? Explain why.

d) Compare the five trials.

(data from Data and Story Library, http://lib.stat.cmu.edu/DASL/)