The development and evaluation of a survey that makes use of student data to teach statistics

David L. Neumann
Michelle M. Neumann
Michelle Hood
School of Psychology, Griffith University, Australia

Journal of Statistics Education Volume 18, Number 1 (2010),
www.amstat.org/publications/jse/v18n1/neumann.pdf

Key Words: Real data; Teaching strategies; Evaluation; Student data.

Abstract

The use of a real data set has the potential to increase engagement and learning in students who enrol in a statistics course at university. The present report describes the development of an approach that uses a real data set, but one that is collected from the students. The questions are designed so that the data set can be used throughout the course to illustrate relevant concepts and methods in the application of introductory statistics. An evaluation was conducted via individual interviews with a random sample of 38 students. Quantitative and qualitative responses indicated that the survey led to in-class participation, was perceived to be a different approach, and contributed to an interest in, understanding of, and appreciation of the relevance of statistics. The creative use of student data is recommended to facilitate the learning of statistics.

1. Introduction

To improve student learning outcomes in statistics courses, a number of teaching approaches have been suggested. The American Statistical Association, for example, recommended incorporating data, using less theory, emphasising statistical thinking, and using active learning (Cobb, 1992). More recently, a focus on the motivational factors that influence the

approach students take to statistics suggests that teaching should engender positive affect and that students should be encouraged to persist even when difficulties emerge (Budé et al., 2007). In addition to the use of general approaches a number of specific teaching strategies have been suggested. The use of computer-based activities (e.g., Morris, Joiner, and Scanlon, 2002), humour (Friedman, Friedman, and Amoo, 2002; Lomax and Moosavi, 2002), and interactive multimedia (e.g., González and Birch, 2000), are among the strategies suggested. One of the most widely recommended strategies, however, is to incorporate real data sets into statistic courses (e.g., Cobb, 1992; Morgan, 2001; Singer and Willett, 1990).

Real data sets may be obtained from various sources and their use can have a number of advantages and disadvantages. Data may be obtained from published journal articles (see Singer and Willett, 1990 for a list). Data may also be obtained from on-line repositories, such as the JSE Data Archive (http://www.amstat.org/publications/jse/jse_data_archive.html), the Data and Story Library (http://www.stat.cmu.edu/DASL/), the Australasian Data and Story Library (http://www.statsci.org/data/), and CAUSEWEB's dataset archive (http://www.causeweb.org/). Statistics textbooks may also come with data sets available though a companion CD-ROM (e.g., Moore, 2009) or on a companion web site, such as the Internet Companion for Statistics, authored by Michael Larsen (http://larsen.duxbury.com/). Finally, individuals may collect their own data by using information published in a newspaper (e.g., obituaries; Morgan, 2001).

It has been suggested that real data sets have intrinsic interest and will capture attention (Singer and Willett, 1990). The student may also act in the role of a researcher to answer questions that were asked when the data was originally collected (Singer and Willett, 1990; Sullivan, 1993). Such benefits will presumably enhance student learning outcomes in statistics courses. The use of real data can also have some drawbacks. For example, there can be low statistical power when small data sets are used and some easily accessible data sets contain only aggregated data or self-selected samples (Singer and Willett, 1990). The high workload in searching for data from published sources can be a further disadvantage.

Some of the disadvantages to using real data sets can be minimised by obtaining the data from the students directly. In the *Instructor's Guide* to an introductory statistics textbook, Moore and Nester (2003) suggested starting a course with a survey that uses the student body as the source of data. Moore and Nester (2003) provided example questions from which qualitative and quantitative data could be obtained (e.g., *how many siblings do you have?* and *what is your favourite type of cheese?*) and some possible ways to explore the data (e.g., make comparisons between males and females). In a similar vein, Sullivan (1993) suggested that instructors can create a survey to obtain information about students in a statistics class. Furthermore, the questions could elicit responses at different levels of response measurement and the concept of measurement be explored in combination with graphical and numerical statistics. Schacht and Stewart (1992) describe the use of vignettes and an IQ test to obtain student data from which measures of central tendency and variability, confidence intervals, and *t*-tests may be calculated. The breadth and scope of statistics makes it conceivable that various applications can be illustrated with a data set collected from students.

In the present article, we describe the development of a survey from which a rich data source is obtained for illustrative purposes in a university introductory statistics course. We also investigated the influence that the use of student data had on learning and engagement. While the use of student data in teaching statistics seems to make sense, it is less clear what effect it will have on student experiences. It has been suggested that using students as the data source

will enhance learning of statistics (Schacht and Stewart, 1992). Because the data is also real, it is possible that using student data will also share the proposed benefits of real data sets, that is, of being intrinsically interesting (Singer and Willett, 1990). To investigate these possibilities in a more systematic way, we conducted interviews with a random sample of students. Qualitative and quantitative data were examined to determine what advantages and disadvantages the use of student data has on learning outcomes.

2. Development of the data gathering survey

The data gathering survey was developed to suit the kinds of research methods, data types, and statistics that are taught in the targeted course. However, many of the questions used are generic, of general interest, and could be applied in any statistics course. Other questions could also be added depending on the specific needs of the course or the characteristics of the student body (e.g., field of study) and instructor.

2.1. The course and student population

The data gathering survey was used during a first year research methods and statistics course taught to behavioural and social science students. The course aims to provide a basic introduction to statistics and has core topics of research methods, graphical descriptive statistics, numerical descriptive statistics, bivariate correlation and regression, probability, sampling distributions, confidence intervals of the mean, and hypothesis testing using *t*-tests and related nonparametric tests. The course is taught across 12 weeks of a semester in a weekly 2 hour lecture and a 2 hour tutorial. The lectures employ audiovisual aids (e.g., computer-based presentation of lecture slides) and tutorials employ both small group and computer-based activities. The statistical software package SPSS is used for data analyses. The course is a pre-requisite for several second year psychology courses and must be taken by students enrolled in the psychology program.

The course has enrolments of approximately 225 students. Demographic information about the students obtained from a survey conducted by the school indicates that the students are mostly aged from 17 to 19 years (60%), with a further 20% aged between 20 to 24 years, and 20% aged 25 years and older. Most students are female (84%). Most are also single with no children (68%), while the remainder are in a relationship with no children (12%) or in a relationship or single with children (20%). A minority are primary income earners (11%) and have career responsibilities. A high majority are domestic students (96%).

2.2. The data gathering survey

The questions used in the data gathering survey are shown in Table 1 and they can be easily formatted so that they fit on a single piece of paper. Some questions are open-ended and some require a forced-choice response. The questions allow the collection of both qualitative and quantitative data and most questions are examined in more than one way throughout the course. In this way, the same data set and question can illustrate various concepts in statistics.

Table 1
Questions used in the data gathering survey

Number	Question
1	Are you MALE or FEMALE?
2	To the nearest centimetre, how tall are you?
3	To the nearest kilogram, what is your weight?
4	What is your date of birth? Please write as DAY / MONTH / YEAR
5	In years, guess the age of the instructor
6	Guess how many children the instructor has[a]
7	On a scale of 1 to 100, where 1 = *no anxiety* and 100 = *very high anxiety*, rate how anxious you are about studying statistics
8	On a scale of 1 to 100, where 1 = *not at all good* and 100 = *very good*, rate your current knowledge of mathematics
9	How high is your self esteem? Please rate your self esteem by circling the appropriate response on the right[b]
10	How motivated are you to achieve a good grade in this course? Please rate your motivation on a scale from 1 to 9 where 1 = *not at all motivated* and 9 = *very highly motivated*
11	Have you ever lied to someone about your age? Please answer YES or NO.
12	What is the diameter of the planet Saturn? Please write your best guess to the nearest whole kilometre e.g., 5970 km[c]
13	Does a dog owner resemble their pet dog in appearance? Please answer YES or NO
14	Of the following, which TV show would you *most* prefer to watch? "Desperate Housewives", "The Bill", "So You Think You can Dance", or "Big Brother"[d]
15	Write down any number between 1 and 10
16	In the past month, have you realised cash through the sale of an item? Please answer YES or NO[e]
17	On a typical day, how long do you spend watching television? Please answer in minutes e.g., 2 hours = 120 minutes

Note: Questions are usually given in random order. They are ordered here to aid discussion of the items. Unless otherwise specified, open-ended responses are made by the student to the right of each question.

[a]The responses are given in two formats, either as an open-ended response or as a forced choice response from the options of "1, 2, 3, 4, or 5".

[b]The response options are *Very high, High, Moderate, Low, Very Low*

[c]There are two versions of this question that differ in the numerical value of the example. given. For instance, the alternative version for this question would state "e.g., 59,700 km". The planet (or moon) used in the question varies from year to year.

[d]The television shows selected can vary according to what are popular at the time and in the region.

[e]The exact question used should be based on an astrology prediction obtained from a newspaper or other source for a particular zodiac sign. The question should be in a yes-no response format for ease of interpretation.

The data obtained from the data gathering survey was used during each of the weekly lectures across 11 weeks of the course (the data was collected from students at the end of the lecture in the first week of the 12 week course). In some lectures, the data from selected questions in the survey formed the focus of the statistical concept being taught and in other lectures the data were used in small example illustrations. To estimate the amount of time the survey was discussed in the course, the percentage of lecture slides that made reference to the survey questions or data was calculated. The mean percentage of slides was 15.04% (*SD* = 13.85). The survey was used most on the topics of measuring association (44%) and graphical descriptive statistics (37%) and least on the topics of confidence intervals of the mean (2.3%) and the standard normal distribution/z-scores (3.1%). This variability suggests that the data collection survey lends itself to illustrating certain statistical concepts more than others.

2.2.1. Research methods.

The data collection process allows students to explore several issues in research. The distinction between qualitative data (e.g., Questions 1, 11, 13) and quantitative data (e.g., Questions 2, 3, 5) and types of measurement scales can be discussed. For instance, the instructor can lead a discussion of whether the measurement scale used in Question 9 is best described as ordinal, interval, or ratio. The items have high face validity and the underlying construct is easy to identify, although the subjective and self-report nature of the data will allow students to consider issues regarding insight or honesty in the answers given (e.g., Question 10 and Question 11) and other validity issues. The rationale behind conducting certain types of analyses can also be discussed in relation to whether they address a real world question. For example, it is possible to calculate correlations between several variables in the survey. However, some correlations will be more meaningful and relevant to real world questions (e.g., correlation between prior mathematics knowledge and level of statistics anxiety) than others (e.g., correlation between prior mathematics knowledge and height).

Two items directly examine the issue of question and answer format. Two forms of the survey are produced in which Questions 6 and 12 are altered. Question 6 is varied between free-response and a forced-choice to explore the issue of whether asking people to provide answers in particular ways influence the answer that is given. Question 12 is varied in the example "answer" to explore whether question wording can influence responses. The phenomenon of anchoring bias (Kahneman and Tversky, 1974) is demonstrated with this question. For example, one group that was given the example of "5,970 km" gave a mean guess of 46,681 km, whereas another group given the example of "59,700 km" gave a mean guess of 141,057 km. The difference of nearly 95,000 km is large enough to appreciate without the need to introduce significance testing.

The major approaches used in research can also be explored. The continuous nature of many of the variables lends themselves to a discussion of the use of correlational research methods. Similarly, differential research can be demonstrated by using the qualitative variable of gender (Question 1) or by converting a continuous scale measure into categories (e.g., converting age from Question 4 into "old" versus "young"). Finally, the manipulations used in Questions 6 and 12 provide an opportunity to discuss experimental research.

2.2.2. Getting to know the other students.

The student can see how their own responses to the questions compare with the group as a whole. This can include demographic characteristics of gender (Question 1), height (Question 2), weight (Question 3), and age (Question 4). Those questions that are more personal in nature, such statistics anxiety (Question 7), self-esteem (Question 9), academic motivation

(Question 10), and lying about one's age (Question 11) can give a student insight into their own psychological traits. For instance, knowing that there are other students who gave a high rating of statistics anxiety could benefit the student by showing that statistics anxiety is common. Finally, there are other questions that are related to social behaviours such as television show preferences (Question 14) and television viewing times (Question 17).

2.2.3. Getting to know the instructor.

The perception of and academic relationship with the course instructor can impact upon the statistics anxiety (Pan and Tang, 2005) and academic motivation (Earley, 2007; Tremblay, Gardner, and Heipel, 2000) of a student. Questions about the instructor might thus have a positive impact upon student experiences. Two items in the survey relate specifically to the instructor (Questions 5 and 6). Question 6, guessing the number of children the instructor has, was deliberately selected because the instructor has significantly more children than average. However, the exact questions used could be easily adapted to suit the individual instructor by, for example, asking students to guess the instructors height, country of birth, years since doctorate degree, favourite sport, or handedness. It is recommended that the correct answer be something unusual or unique so it will generate interest. Various statistics can be demonstrated by using these questions. In the course in the past, graphical and numerical descriptive statistics have been used, responses made by males and females have been compared, and the responses have been correlated with responses from other items on the survey (e.g., the guess of the instructors age has been negatively correlated with self-esteem and statistics anxiety).

2.2.4. Graphical descriptive statistics.

The data from questions about television viewing preferences (Question 14), the resemblance in appearance between dog owners and their pets (Question 13), and lying about ones age (Question 11) can be shown as pie charts or bar charts. The responses to the questions about height (Question 2), weight (Question 3), age (Question 4), television viewing time (Question 17), and others can be explored through frequency distributions, histograms, stemplots, and boxplots.

One demonstration that is particularly instructive is to compare males and females on the question of lying about ones age (Question 11). Students can be asked whether they expect males and females to differ. After a short discussion the instructor shows a bar graph of the data, but the graph only shows the counts in the number of males and females that responded *yes*. There are always more females responding *yes*, and this produces a substantial reaction from the students. Students can be asked why the data shows this apparent difference. While some may say that it confirms their hypotheses, others will recognise that there could be other, statistical reasons. These are that the counts for the *no* responses are not shown and that there may be simply more females in the course. Both these alternative explanations can be explored by showing a bar graph with the counts for both the *yes* and *no* responses and a bar graph showing the percentages of each gender that responded *yes* and *no*. We have found that although there are more females than males in the course, once this baseline dependency is removed there are no gender differences in propensity to lie about one's age.

The histogram, stemplots, and boxplots generated from the items yielding quantitative data can be explored to discuss features of the distribution. For example, we typically find that age produces a positively skewed distribution. An interesting demonstration can be made using the responses to the questions about height or weight. The resulting distributions in a course will be invariably bimodal. The question can be posed as to why a bimodal distribution is

shown. The answer is due to physical differences between males and females. Comparing the separate distributions for males and females side-by-side emphasises this point.

2.2.5. Numerical descriptive statistics.

Similar to the graphical presentation of the data noted above, numerical descriptive statistics for many of the questions will be of interest to the students. The appropriate statistic to use (e.g., mean vs. median vs. mode) can be discussed in the context of the distributional characteristics of the data (e.g., using the median when the data is skewed). Similarly, the influence of outliers that will naturally occur in the data on the numerical descriptive statistics can be explored.

2.2.6. Correlation and regression.

The quantitative nature of many of the variables provides the opportunity to demonstrate scatterplots, correlation, and regression. We have used the data set to examine what variables are correlated with statistics anxiety (Question 7). Key variables that can be explored are age (Question 4), academic motivation (Question 10), mathematics knowledge (Question 8), and self-esteem (Question 9). Prior research has found that statistics anxiety is positively correlated with age and academic motivation (Lalonde and Gardner, 1993; Tremblay et al., 2000), and negatively correlated with knowledge of mathematics (Lalonde and Gardner, 1993; Tremblay et al., 2000) and self-esteem.

The large enrolments in the course also ensure that statistically significant correlations emerge between seemingly unrelated variables. For instance, we have found that the guess of the instructors age (Question 5) is significantly negatively correlated with knowledge of mathematics (Question 8) and weight (Question 3) and that time spent watching television (Question 17) is significantly negatively correlated with the guess of the diameter of Saturn (Question 12). These associations allow the students to explore the distinction between statistical versus practical significance. Moreover, students can consider whether the apparent association reflects a causal relationship, the influence of a third variable, or is spurious.

Similar to the calculation of correlations, many of the variables used can be explored through regression. For instance, regression equations can be constructed to predict the level of statistics anxiety (Question 7) from other variables. The concept of regression may be introduced in a novel way by contrasting it with predictions made from astrology. Question 16 can be constructed immediately before the survey is administered by using a prediction made in a published horoscope. The prediction should be one that can be answered in a yes-no response format for ease of interpretation and it should apply to only one zodiac sign. The extent to which the prediction is correct for the zodiac sign (and others) can be determined because students indicate their date of birth (Question 4). In the past, we have found that this technique has generated considerable student interest and can effectively highlight the unique characteristics of regression, such as it being mathematical, based on objective data, and allowing the calculation of errors of prediction.

2.2.7. Probability and randomness.

An effective way to introduce probability is to ask students to contemplate the probability of events (e.g., coin tossing). Students can be asked what the probability is that two people in the same room share the same day and month for their birthday. Calculations show that the probability in a room of 10, 23, and 50 people is approximately .12, .50, and .97, respectively. Real student data can be applied to this problem using Question 4 in which students indicate their date of birth. A course with moderate to high enrolments will find

many matches, and some dates can have more than two students sharing the same birthday. Students in our course are surprised when shown the results. It highlights to students that not all probability questions are easily answered intuitively, but that probability rules should be applied to aid in the calculations.

Another application of probability in statistics is in the random selection of individuals. It is vital that the selection method used is truly random and that individuals (or numbers) are not selected directly by the person. The importance of this application can be emphasised through Question 15 in which students select a number between 1 and 10. If people are random we might expect to see a uniform distribution in the selections of the numbers. However, we have consistently found that approximately one third of students select the number 7, with the next most popular number being 3. Showing this real data proves the point that to select a random number requires the use of a random digit table or computerised generator.

2.2.8. Inferential statistics.

Several questions can be used to show the application of inferential statistics. Normative data on several variables (e.g., number of minutes spent watching television in Question 17) can be used to demonstrate hypothesis testing with one sample. Gender differences can be explored to examine questions that employ both qualitative and quantitative data. For example, an independent groups *t*-test and effect size analysis can be used to examine whether male and female students differ significantly in psychological variables (Question 7, 8, 9, and 10) and physical features (Questions 2 and 3) among other variables. A contingency table analysis can be used to explore gender differences in propensity to lie about one's age (Question 11) or opinions on whether people resemble their pet in appearance (Question 13). We have consistently found gender differences in statistics anxiety and maths knowledge and because of the differential nature of the research approach, this affords an opportunity to explore issues regarding causality ("Does being female cause you to be anxious about statistics?"), individual variability ("But not all females are highly anxious"), and the influence of other variables ("Could there be a third factor involved in the relationship?"). The administration of the data gathering survey also includes the experimental manipulation noted above for Question 6 and 12.

2.2.9. Assignments or practical class work.

The data can be used to provide a real data set to base an assignment or practical class work on. For example, Questions 7 to 10 have been used successfully in an assignment and in practical class work. Students used the data set to calculate descriptive statistics and correlations between the variables and presented the results in the report. The personal relevance, experience in completing the items, and the high face validity of the questions used facilitates its incorporation into the curriculum. It is also conceivable that other items in the survey could be adapted for assignments or class work.

2.3. Cautions in using the survey

Potential limitations and cautions in using the data gathering survey need to be noted. Some items on the survey request personal information from the students, namely their age, height, and weight. It is conceivable that the identity of a student could be revealed when exploring these variables with some statistical approaches (e.g., through a stemplot, scatterplot, or frequency distribution). As such, the instructor should take care to screen out individuals that have high or low values on the variables (e.g., if there is a much older or heavier student in the class). By necessity, this may not allow the full data set to be used in some

demonstrations. Sullivan (1993) suggested that some questions be avoided in certain situations, such as a question regarding age in a class of non-traditional students.

In the administration of the survey, it should be stressed to the students that answers are provided voluntarily and that they do not need to answer all questions if they do not want. The latter issue can be used to introduce the topic of ethics in research, particularly that concerned with psychological research, clinical trials, or when personal information in collected from people. The ethical nature of data collection could be reinforced by using an informed consent mechanism (e.g., students sign an informed consent form) in the administration of the survey.

Other limitations stem from the anonymous and voluntary nature of the survey. Not all students will necessarily complete the survey. The obtained data will thus not contain the full population of students in the course. However, this fact can be explored with students by discussing the distinction between samples versus populations and potential biases in the actual data that is obtained. Finally, the anonymous nature of the survey can sometimes lead to students giving fake or nonsense answers (e.g., an unrealistic guess of the instructors' age in Question 5). Such attempts can be dealt with by appropriate screening of the data set.

3. Evaluation of the data gathering survey

The evaluation of the data gathering survey focused on how its application influenced student experiences during the lectures and when learning statistics. Semi-structured interviews were conducted with a stratified random sample of students from the course and their responses were coded to identify major themes. The advantage of this approach was that it gave greater open-endedness, greater sensitivity to individual experiences, more depth, and a capacity to connect various experiences (Neuman, 2006). Quantitative data were also obtained on rating scales that were specifically focused on certain themes that had been identified through prior pilot research (e.g., the extent to which the survey aided learning).

3.1. Participants

The participants were sampled from the enrolments in a first year research methods and statistics course at the authors' university. The sampling and interviews were done in the semester following that in which the course was completed. A target sample size of 40 was obtained through a stratified random sample method. Students were first grouped according to the final grade (high distinction, distinction, credit, pass, fail). Students were randomly selected so that the proportion of students with each grade in the final sample was equal to the proportion of the grades in the class as a whole. This was done to ensure that the sample was representative of the full range of abilities in the class. Students were initially contacted via telephone and invited to participate in the study. The interviews were conducted at a later date and participants received a café voucher (total value of $7.00) for their participation. Of the 40 students that originally consented to participate, one student (who obtained a fail grade) could not be subsequently contacted and another student (who obtained a pass grade) withdrew their consent before the interview was conducted. The final sample thus consisted of 38 students (27 female and 9 male) with a mean age of 23.97 years (SD = 7.33). At the time of the interview, the students were enrolled in an undergraduate bachelor degree. Thirty students had no prior tertiary level qualifications, three had completed a certificate, four had completed a diploma, and one had completed a bachelor degree. The students reported that they attended a mean of 91% of the lectures in the course (SD = 13.14).

3.2. Interview procedure

After the initial contact via the telephone, a copy of the project information sheet and consent form was mailed to the student. The student was subsequently contacted via telephone for the interview. In addition to obtaining demographic information and comments on the use of the data gathering survey, the interviews also sought feedback on other aspects of the course and these are reported elsewhere. All interviews were digitally recorded for later transcribing and coding. To ensure confidentiality, the interviewer was a person unknown to the students and had not been involved in the course. The students were informed that their transcribed responses would be coded to ensure that their identity would not be known.

The interviewer began with open-ended questions to obtain feedback on the use of the data gathering survey. Initially, the interviewer described the data gathering survey to the student and provided some examples of the questions used and types of statistics that were calculated. The interviewer next asked questions in a semi-structured format. The student was first asked *What are your thoughts on the use of the data gathering survey in the course?* The interviewer asked the student to elaborate or clarify their comments where relevant. Depending on whether the student had already covered these topics in their previous responses, the interviewer ask the students additional questions of *Did it help you engage with the material? How?*; *Did it help motivate you to learn about statistics?*; *What were some positive aspects to it?*; and *What were some negative aspects to it?*. In each instance, the interviewer asked the student to elaborate or clarify their answers as required.

Following the open-ended questioning, the interviewer asked six questions that required the student to make a rating. The questions were based on pilot research that indicated preliminary themes of how the data gathering survey influenced student experiences during the course. All ratings were made on a 10-point scale where 1 = *not at all* and 10 = *very much*. The exact wording used in each question is shown in Table 3. The questions asked whether the data gathering survey (a) *made the learning of statistics more relevant*, (b) *made the learning of statistics more interesting*, (c) *reduced anxiety about studying statistics*, (d) *made the learning of statistics more enjoyable*, (e) *increase motivation to study statistics*, and (f) *made it easier to apply statistics to new situations or data sets*. The student was next asked to answer "yes" or "no" to the question of *Would you recommend that the data gathering survey be used in the course again?* After these questions, the interviewer obtained demographic information about the student.

3.3. Data coding methods

The student responses during the interview were initially transcribed verbatim from the digital recordings by the interviewer. A second researcher took the transcribed interviews and numbered each unique statement according to the participant code. Next, the resulting 156 statements were grouped into similar themes and given preliminary labels. This initial pass through the data resulted in 12 categories with an additional category containing other comments that could not be categorised or were irrelevant to the purposes of the interview (e.g., comments on teaching style of lecturer). A third researcher examined the categories that were formed and the appropriateness of the fit of each statement in each category. There was agreement on 84% of the classifications made. The two researchers reached 100% agreement following a discussion in which some statements were moved to different categories. In addition, two of the categories were combined with other categories to result in 10 categories at this stage. The next step in the coding involved a fourth researcher checking the fit

between each statement and the category it was placed in. There was 98% agreement in the categorisations. Following discussion, two statements were moved to a different category and one statement was removed.

The 10 categories at this stage were re-evaluated by examining the relationships between them and the number of statements in each category. A category that was tentatively labelled "curious" had only two statements and it was subsequently re-categorised as a subcomponent of the category labelled "interesting". A category tentatively labelled "increased motivation" was removed because it only had four statements representing it. Finally, a category tentatively labelled "icebreaker activity" contained six statements. It was determined that each of the statements contained sufficient overlap with other established categories that they warranted redistribution. As such, this small grouping of statements were distributed to categories that were tentatively labelled as "more at ease" (three statements), "fun" (two statements), and "participation" (one statement). At the conclusion of this stage in the coding process there were seven categories of statements. The final stage in the coding process involved two researchers assigning labels to the categories, deriving an appropriate definition, determining whether the categories contained any subcomponents, selecting exemplar statements, and quantifying how many individual students contributed to each category.

3.4. Results

3.4.1. Categories of statements.

The labels, definitions, and example statements for the seven identified categories are shown in Table 2. The three most common categories were *created interest, increased relevance,* and *helped understanding*. Comments coded in the category of *created interest* included "got me interested in the lectures", was "very clever to engage me", and "made the course more interesting". This category appeared to include subcomponents in which students referred to engendering interest, increasing attention towards statistics, and being curious about the results of the survey. The category of *increased relevance* indicated that the survey helped students to appreciate the significance of statistics. Three subcomponents were apparent. The first was a general reference to the relevance of the survey in that it made the statistics "more relevant", "made it more personal", and gave "something you related to". The second made reference to the real-world aspects of the survey in that it gave a "real life situation to understand it" and gave a "realistic perspective". The practical application illustrated by the survey was the third subcomponent in this category. Comments were made in relation to how the results were shown "in a really practical manner" and also how it gave them "practical knowledge". The category of *helped understanding* suggested that the survey had pedagogical benefits in assisting students to comprehend and learn about statistics (e.g., "gave a better understanding"). Some aspects of the increased understanding appeared to reflect the real-world basis of the data with students indicating the benefits of being able to "relate it back to things we could get our head around" and that some questions could be "turned around to show how statistics work".

Four additional categories were identified during the coding. The category of *fun activity* indicated that the students enjoyed "doing the activities", "it was a fun way to start", and it was "fun to hear the answers and stuff at the end". An unexpected category was that the survey *increased participation* by the students. The increased participation emerged in several areas such as "giving information", "generating a lot of discussion with our peers", and "got us to be involved". *Reduces negative mood* emerged as a category in which students

highlighted that the application of the survey was beneficial in reducing some of the "stress" and "anxiety" associated with statistics and helped to "put us at ease". The final category was labelled as *different approach*. This relatively small category reflected that the application of the survey was a unique approach that had not been encountered in prior courses or in other course materials (e.g., textbooks). It was noted that the approach was also "abstract", "different to the other classes", and showed a "different angle" to statistics.

Table 2
The labels, definitions, percent of students that contributed to each category, and representative comments for the final categories resulting from the qualitative data coding methods

Label	Definition	Percent	Example statements
Created Interest	The survey increased student interest and attention towards statistical concepts	71%	"Made it more interesting because it was our own data that we actually put input into" "They made me more interested it what was going to come next and what we were going to learn about"
Increased Relevance	The survey increased the connection between statistics and real concepts and showed examples of practical applications	63%	"We could get our head around rather than just being numbers we had no relation to them" "It was good to see how it was used practically"
Helped Understanding	The survey helped students to comprehend and learn about research methods and statistics	45%	"It helps me get an understanding of what he was talking about." "It was good to help you get your head around what's involved in the stats course."
Fun activity	The use of the survey was a fun, enjoyable activity	32%	**"**For me I just enjoyed doing the activities" "Something a bit fun"
Increased Participation	The survey functioned as an activity that students could become actively involved in	29%	"Got us involved in the class" "I think it was good that it was something you could participate in."
Reduces negative mood	The survey reduced negative mood states such stress and anxiety about studying statistics and helped put the students at ease	21%	"It made it a bit less daunting seeing that it wasn't that hard like when he broke it down and showed us how he did it. It was a bit less scary" "It made everybody feel more at ease about the stats course"
Different Approach	The survey was a unique way to illustrate statistical concepts to students	16%	"It was something different." "Showed me a different angle."

3.4.2. Quantitative ratings.

Descriptive statistics for the subjective scales are shown in Table 3. The subjective scales included several of the themes identified through the qualitative coding of the statements during the interview. However, since all students provided a response to the ratings, the means obtained indicate the extent of agreement with each theme. The obtained mean values indicate that students showed moderately strong agreement that using the data gathering survey made the learning of statistics relevant, interesting, enjoyable, and helped in applying statistics to new situations or data sets. Students indicated moderate agreement that the data gathering survey reduced their own anxiety about studying statistics and increased their own motivation to study statistics. The subjective ratings thus seem to confirm the impressions from the coding of the interview data in that the main contribution of data gathering survey is in understanding the statistics being taught and the interest and relevance of statistical concepts. Affective influences of the data gathering survey, though present, appear to be more modest.

Table 3
Means (M), standard deviations (SD), and 95% confidence intervals of the mean (CI_{95}) for ratings to each question on the 10 point scale (1 = not at all and 10 = very much)

Question	*M*	*SD*	CI_{95}
Did the use of the survey make learning of statistics *more relevant* for you?	7.39	0.82	6.80 – 7.99
Did the use of the survey make learning of statistics *more interesting* for you?	7.21	1.60	6.69 – 7.35
Did the use of the survey *reduce your anxiety about studying statistics*?	5.45	2.27	4.70 – 6.19
Did the use of the survey make learning of statistics *more enjoyable* for you?	6.68	1.92	6.05 – 7.31
Did the use of the survey *increase your motivation* to study statistics?	5.68	2.12	4.99 – 6.38
Did the use of the survey make it easier to *apply statistics to new situations or data sets*?	7.18	2.00	6.53 – 7.84

4. Conclusions

The data gathering survey described in the present article is easy to administer and provides a rich data source to illustrate many statistical concepts to students. In the evaluation of the approach, several themes were identified to suggest that the survey was related to positive learning experiences. Models exist that relate differences among students in motivation, anxiety, attitudes, and aptitude to achievement in statistics courses (Lalonde and Gardner, 1993; Tremblay et al., 2000). Although the causal relationship between these factors remains to be determined, the data gathering survey was shown to be related to several variables in these models. Statistics anxiety and other negative mood factors were among the themes that

were identified by students in the evaluation. In relation to motivational factors, students reported that the survey influenced their level of interest and the relevance of the statistical concepts they were learning. Finally, students reported that the approach influenced their understanding of statistics. In combination, the results suggest that the data gathering survey may have benefits to student engagement and learning when used in a statistics course. However, it should be noted that the present results are generally descriptive in nature and as such it is not possible to conclude that there is a causal relationship between the use of the data gathering survey and improved student learning and engagement. An experimental approach that uses an appropriate control group and objective measures of learning and engagement (e.g., exam performance, attendance records) would be required to show a causal relationship.

Singer and Willett (1990) describe several desired pedagogical characteristics of real data if they are to enhance learning. The use of data collected from students appears to fulfil many of these characteristics. The data set is *authentic* because it is collected from the students themselves. The data set has *interest* and *relevance* because it is based on variables that have general appeal (e.g., instructor's age, statistics anxiety, favourite television show) and are relevant to many topics in psychology (e.g., academic motivation, self-esteem). Feedback from students obtained in the interviews confirmed that the data was of interest and had relevance. The application of the data gathering survey also leads to *substantive learning* as students can learn more about themselves and how they compare to the class as a whole. The students can also learn about relationships between variables, gender differences, and biases that people have (e.g., anchoring bias and non-random selection of numbers). The *availability of multiple analyses* is also present in the data set to allow students to explore how several types of analyses can be applied to the same variables. Finally, because the *raw data* is available, along with *case identifiers*, summary statistics can be calculated and issues of outliers and missing data can be explored by relating them to actual cases in the data set.

The findings that emerged from the qualitative analysis of the interviews and the quantitative ratings are generally consistent with the suggestions made by other researchers that there are benefits in using real data sets when teaching statistics. The results support the claim that real data is intrinsically interesting to the students (Singer and Willett, 1990). The present sample consisted of psychology students. Prior research has indicated that psychology students report only a moderate level of interest in mathematics, and certainly a lower level of interest than in psychology itself (Tremblay et al., 2000). The use of the data gathering survey might thus help address the low interest in statistics experienced by these students. In addition, the results indicated that the survey illustrated the practical application of statistics. This could reflect that real data provides an opportunity to show students how to deal with common problems such as outliers, missing values, and response bias (Morgan, 2001; Singer and Willett, 1990). Finally, there was some indication that the students enjoyed examining the data resulting from the data gathering survey. Although this does not relate directly to pedagogical aspects, it does have the potential to address the generally moderate level of interest in learning about statistics reported by students (Tremblay et al., 2000).

An unexpected finding during the coding of the interview data was that students reported that completing the survey encouraged them to become involved in the class. Schacht and Stewart (1992) note that the data collection methods that they used to obtain data from students (e.g., use of vignettes) required participation from the students. Moreover, the techniques lose their effectiveness if the class is reluctant to participate. It is important to create a class environment that encourages participation prior to administration of any data collection tool

(Schacht and Stewart, 1992). The use of short questions, intrinsically interesting variables (e.g., statistics anxiety, guess of instructor's age), and the anonymous nature of the survey may be factors that facilitated student participation. Each student must also understand the relevance of their participation, which is why the data gathering survey was introduced as an educational activity.

While the results from the interview were generally positive in regards to the application of the data gathering survey, there were some limitations to the approach used. While the evaluation data indicated that the survey helped understanding of statistical concepts, no information was elicited during the interviews to examine what specific concepts the survey data contributed most towards understanding. Future research might ask students directly what examples in the use of data they recall or what statistical concepts they learnt better as a result of the survey. In addition, the approach taken in the interviews may have biased the results towards a more favourable evaluation of the data gathering survey. For example, the question about the positive aspects of the survey was always asked before the question about the negative aspects (in future these should be counterbalanced). Finally, it would be desirable to use a larger sample size to ensure that the full diversity of student responses can be accounted for.

The present form of the data gathering survey can be modified to suit the teaching context or types of statistics that are taught in a course and be expanded upon to examine new questions[1]. For example, in addition to asking students to pick a number between 0 and 9, a second question could ask for a number between 0 and 99. The correlation between the two numbers could be examined and whether there is a tendency to answer with a rounded number for the latter question (e.g., answer with 50 or 55 instead of 53). In addition, the present data gathering survey was developed so that it could be completed in class and through self-reports. An alternative approach could be to develop a survey that students take home with them to complete by taking measurements on objects in their homes (e.g., number of computers, size of bedroom), on themselves (e.g., heart rate at rest and after a physical activity), or observations of the behaviour of other people or animals (e.g., number of times pet dog barks in a set period of time). Such activities can also promote discussion about measurement scales, operational definitions of variables, and the reliability and validity of measurements. It is also possible to use the data from the survey to illustrate how it may be possible to use statistical software to conduct statistical analyses that make no sense. For example, most software will readily calculate the mean of a qualitative variable (e.g., gender), yet such a statistic is nonsensical. This and similar statistics could be calculated to promote discussion about using appropriate statistical analyses for the type of variables or research questions that are examined.

In conclusion, we recommend the novel use of data collected from students such as through the data gathering survey that we have employed. While there are some limitations to its use, the present study showed some significant benefits to student learning, motivation, and engagement in the course. The questions that we describe are not meant to be prescriptive, but can be easily adapted or changed to suit the needs and context of a statistics course at any educational institution. For instance, the bias towards psychological variables in the present survey naturally lends itself for application to statistics courses in the social sciences. Statistics courses in other disciplines could easily modify the questions used so that they

[1] The input from anonymous reviewers for some of these suggestions is acknowledged.

focus on more appropriate variables. In addition, many of the variables can be explored with the basic statistics that are taught in the secondary school curriculum. An important consideration in any application of the data gathering survey is to avoid using too many questions and to use questions that can be examined in various ways throughout the course. In this way, students can appreciate the power that statistical methods have to examine variables that are of interest and relevance to them.

References

Budé, L., Van De Wiel, M. W., Imbos, T., Candel, M. J. J. M., Broers, N. J., and Berger, M. P. F. (2007), "Students' achievements in a statistics course in relation to motivational aspects and study behaviour", *Statistics Education Research Journal, 6,* 5-21.

Cobb, G. (1992), Teaching statistics. In L. A. Steen (Ed.), *Heeding the call for change: Suggestions for curricular action MAA Notes No. 22* (pp. 3-43), Washington DC: Mathematical Association of America.

Dilevko, J. (2000), "A new approach to teaching research methods courses in LIS programs", *Journal of Education for Library and Information Science, 41,* 267-361.

Earley, M. (2007), "Students' expectations of introductory statistics instructors", *Statistics Education Research Journal*, 6, 51-66.

Friedman, H. H., Friedman, L. W., and Amoo, T. (2002), "Using humor in the introductory statistics course", *Journal of Statistics Education*, 10 (http://www.amstat.org/publications/jse/v10n3/friedman.html)

González, G. M., and Birch, M. A. (2000), "Evaluating the instructional efficacy of computer-mediated interactive multimedia: Comparing three elementary statistics tutorial modules", *Journal of Educational Computing Research, 22,* 411-436.

Kahneman, D., and Tversky, A. (1974), "Judgment under uncertainty: heuristics and biases", *Science, 185,* 1124-1131.

Lalonde, R. N., and Gardner, R. C. (1993), "Statistics as a second language? A model for predicting performance in psychology students", *Canadian Journal of Behavioural Science, 25,* 108-125.

Lomax, R. G., and Moosavi, S. A. (2002), "Using humor to teach statistics: Must they be orthogonal?", *Understanding Statistics*, 1, 113-130.

Moore, D. S. (2009), *The Basic Practice of Statistics 5th Edition,* New York: W. H. Freeman.

Moore, D. S., and Nester, D. K. (2003), *Instructor's Guide to The Basic Practice of Statistics 2nd Edition,* New York: W. H. Freeman.

Morgan, B. L. (2001), "Statistically lively uses for obituaries", *Teaching of Psychology*, 28, 56-58.

Morris, E. J., Joiner, R., and Scanlon, E. (2002), "The contribution of computer-based activities to understanding statistics", *Journal of Computer Assisted Learning*, 18, 114-124.

Neuman, W. L. (2006), *Social research methods: Qualitative and quantitative approaches*, New York: Pearson Education Inc.

Onwuegbuzie, A. J., and Wilson, V. A. (2003), "Statistics anxiety: nature, etiology, antecedents, effects, and treatments—a comprehensive review of the literature", *Teaching in Higher Education*, 8, 195-209.

Pan, W., and Tang, M. (2005), "Students' perceptions on factors of statistics anxiety and instructional strategies", *Journal of Instructional Psychology*, 32, 205-214.

Schacht, S. P., and Stewart, B. J. (1992), "Interactive/user friendly gimmicks for teaching statistics", *Teaching Sociology*, 20, 329-332.

Singer, J. D., and Willett, J. B. (1990), "Improving the teaching of applied statistics: Putting the data back into data analysis", *The American Statistician, 44*, 223-230.

Sullivan, M. M. (1993), "Students learn statistics when they assume a statistician's role", *Paper presented at the 19th Annual Conference of the American Mathematical Association of Two-Year Colleges*.

Tremblay, P. F., Gardner, R. C., and Heipel, G. (2000), "A model of the relationships among measures of affect, previous achievement, and performance in introductory statistics", *Canadian Journal of Behavioural Science*, 32, 40-48.

David Neumann
School of Psychology
Griffith University
Queensland, 4222
Australia
D.Neumann@griffith.edu.au

Michelle Neumann
School of Psychology
Griffith University
Queensland, 4222
Australia
Michelle.Neumann@student.griffith.edu.au

Michelle Hood
School of Psychology
Griffith University
Queensland, 4222

Australia
Michelle.Hood@griffith.edu.au
