Distribution Division: Making It Possible For More Students to Make Reasoned Decisions Using Data

Anthony Harradine
Noel Baker Centre for School Mathematics, Prince Alfred College
Australia

Abstract

The Roundtable theme selected by the International Association for Statistical Education (IASE) organizing committee (2004) focuses on the ability to make "reasoned decisions based on sound statistical thinking." This string of words rolls easily off one's tongue; as an outcome, however, it remains elusive. This paper examines the journey taken by one school, from 1997 to 2004, in an effort to empower its students to make reasoned decisions based on sound statistical thinking. It outlines the background reasons for the journey, and presents a general description and some critical commentary on some of the learning experiences used. The last twelve months of the journey are addressed in some detail. Six phases of teaching and learning that may assist in realizing the goal of reasoned decision-making based on sound statistical thinking are discussed. Complementing these phases is a preliminary approach to data analysis called "distribution division". The effects of this approach on the middle school curriculum are also discussed.

Introduction

In 1993 the Senior Secondary Assessment Board of South Australia (SSABSA) introduced an externally-examined, pre-tertiary (final year of high school) course called Quantitative Methods (QM). Approximately fifty percent of the content was statistical in nature. The overview of the Statistics section reads:

> The aim of this section is to illustrate statistical investigations, from inception to report. Students will complete elementary statistical investigations of their own and comment upon the statistical investigations of others. Ideas of statistical inference will be introduced for proportions and means but only for a single variable with simple random sampling, whereas relationships between two variables will be examined using graphical techniques and simple descriptive statistics. The emphasis will be on interpretation and the use of statistics to solve problems rather than the mechanics of drawing graphs and making calculations, and consequently access to electronic calculators and computer packages is essential (SSABSA, 1992).

Each student was required to devote four weeks to completing a major project that was worth fifteen percent of his or her final mark. The expectation of a major project was daunting; most teachers were without a framework that allowed them to guide students in performing a simple, but sensible, statistical investigation. It may be no surprise that from 1993 to 2002 the course only attracted around 200 candidates per year. Informal discussions with teachers indicated that some reasons for this were that teachers had little or no background in statistics, so the course demands seemed greater than they really were; mathematics classes had very limited access to electronic technology; a textbook had not yet been written; and local universities did not support the course. These challenges notwithstanding, the projects of some of the QM students were outstanding.[1]

In 2003 the QM was replaced with a similar course called Mathematical Methods. Among other topics, Mathematical Methods includes elementary differential calculus (which was not an aspect of its predecessor) as well as statistical concepts beyond those in QM. In this course, however, students are not required to carry out a major project. In 2003 the course had over 500 candidates; a rapid rise in candidature is expected in future years.

The Journey at Prince Alfred College (PAC)

I taught QM from 1993 to 2000. It was hard but rewarding work. Students loved the course. It is in no small part the experience of teaching this course from 1993 to 1997 that shaped my thinking about what might be possible in the middle school (statistics) curriculum at Prince Alfred College (PAC). PAC is a private boys' school in Adelaide, South Australia that delivers educational programs to boys from ages 5 to 18 years. Over the last seven years, work carried out at PAC has aimed to increase 12 to 15 year-old students' ability to make decisions based on sound statistical reasoning.

Prior to 1998, Statistics was not taught seriously, or at all, in most mathematics classes at PAC. Most teachers saw Statistics as the topic one taught at the end of the year if you had covered all the more important topics. This changed in 1998 when mandatory Statistics sections were added to the PAC middle school courses (for boys aged between 12 and 15 years). Each section was of three to four week's duration. Teachers' immediate questions were: "How can we spend *that long* on Statistics?" (They had not taught QM) and "What are we *not* going to teach from other topics?"

Most textbooks available at the time included a chapter on Statistics, but the Year 8 chapter differed little from the Year 9 chapter and so on. The texts tended to offer a mix of skills for making graphs and performing calculations, all of which were rarely put to any logical or interesting use. I decided to write learning materials for each of the Years 8, 9 and 10 courses. The next section describes the major influences on the content of these materials.

Early influences on the materials

The three greatest influences on the material written were:

- the phrase *from inception to report* (SSABSA 1992)
- Moore and McCabe's book, *Introduction to the Practice of Statistics* (1996), and
- an unpublished data handling matrix developed by Robert Hall, a Senior Lecturer in Statistics at the University of South Australia.

The expectation for students to understand what they read in the media was being pushed from all directions. One example of this is the expectation that students be able to understand a graph presented in a newspaper. Some responded to this expectation by providing a flood of media examples for students to read and interpret. My experience teaching QM led me to believe it to be vastly more difficult for a young mind to do this if they have never experienced a simple statistical investigation from inception to report. Through this experience it became apparent to me that it is difficult to appreciate the wealth of knowledge hidden behind a graph, table or statement unless you have had firsthand experience in developing such materials from raw data.

[1] In one of the more memorable projects a student investigated how two different procedures for castrating young male sheep affected their weight at sale time. As a result of this investigation, the student's father changed his farming practice.

If students are to appreciate and understand the sense (or nonsense) of plots and statements made by others, they first need to come up with a problem of their own, or have one posed to them. They need to collect some data pertaining to the problem, calculate statistics, and produced graphs from the data. Moore and McCabe's book, *Introduction to the Practice of Statistics* (1996), was the text book, in our field of vision, whose approach to the learning of statistics most closely approximated an *inception to report* format. As a result, this text was an important influence on the materials developed for the Year 8, 9, and 10 courses.

Robert Hall's data handing matrix also played an influential role. It provides a way for novice teachers (and students) to think about statistics and gives them a chance to use statistical techniques to solve simple problems. The matrix categorizes problems by both the number of variables and the status (i.e. response or explanatory) of the variables. Variable-response categories in Hall's matrix include:

- Single variable, nominal/ordinal
- Single variable, interval
- Two variables, nominal/ordinal response and nominal/ordinal explanatory
- Two variables, interval response and nominal/ordinal explanatory
- Two variables, nominal/ordinal response and interval explanatory
- Two variables, interval response and interval explanatory.

For each category the matrix outlines how to gather and organize the data. It also indicates the appropriate graphical displays and summary statistics, which hypothesis tests are appropriate, and how to create tables suitable for publication. Some statisticians, however, dislike this matrix due to its procedural approach that does not openly encourage creative thinking by students when faced with unfamiliar situations.

The 1998 materials

In 1998 two sets of learning materials were implemented at PAC. In Year 8 boys (aged 13 to 14 years) focused on small problems that required the analysis of categorical data. Year 9 students focused on small investigations requiring the analysis of data measured on an interval scale. The materials for Year 9 students aimed to promote students ability to:

- read *within the distribution of a single variable*, *between two distributions of the same variable but for different categories, and beyond the sample distributions* , appreciating that sample data may provide a hint as to what was happening in the population from which the data came,
- *support a conclusion* using facts and statistics resulting from their data analysis.

Both teachers and students received the materials very well and they have continued to be used by teachers both in and outside of PAC. However, it is questionable whether the more subtle, and most important, aims of the materials were achieved. The following reflection, from a teacher who has used these materials a number of times, illustrates this.

> One of the aims of these units is the acquisition by the student of statistically related skills like the use of percentage, the drawing of graphs and graphical interpretation. As a teacher this aim is easily attained. Students readily learn (or revise) these skills and can trot them out when asked to. The learning here comes easily, but a sense of achievement is lacking. The students work happily in their comfort zone, revising mathematical skills that feature strongly in primary school curriculum. Many of them feel that it is 'Mickey Mouse' maths and rightly feel that little of worth has been achieved.
>
> The main aim of the Statistical Investigator units is the gaining by the students of an appreciation that these statistical skills do something powerful, the idea that statistics is concerned primarily with the answering of questions and the solving of problems. One

> focus of the first of these units is on the concept of sampling, its power and its potential flaws, and the implication it has for the solutions to the problems that we obtain. In this area the learning comes with more difficulty but the potential for achievement is far greater. Students struggle to put things into words, and when they succeed they don't always appreciate the significance of what they have done. Students struggle to understand the core concepts and the way that they interact. Students will have little chance of grasping these ideas if their teachers do not fully understand the significance of these concepts. Whilst the learning is harder, the sense of achievement and sense of the power of this vital field of mathematics makes it worth the effort.
>
> When I first used the first Statistical Investigator unit I was an inexperienced teacher of statistics who had only ever been asked to deliver a skills-based textbook-focused curriculum. As such I glossed over the concepts I should have emphasised. Since then I have taught inferential statistics at a leaving level, and I now feel that I can teach statistics at an entry level with an understanding that I lacked earlier. I now see these units as inferential statistics without the tests and intervals. The problems we solve should really be put on the classroom wall, to be revisited 5 years later when we have the skills to prove our conjectures. (Lupton, A., 2004).

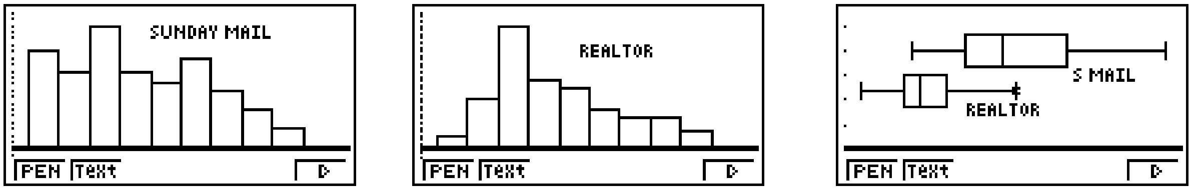
So, as the author of these units, I had some nagging doubts about exactly what students who used them were learning, especially from the unit that focused on the analysis of data measured on an interval scale. One thing not in doubt was that approaching the learning of statistics with a small number of questions to be answered or problems to be solved was the best way to proceed. My personal experiences and those of other teachers left little doubt that students enjoyed trying to solve problems. A sample problem involving the analysis of interval-scaled data is provided in the next section.

A problem from the 1998 materials

Consider the following problem taken from the PAC materials for Year 9 students. It is the type of problem students were expected to cope with at the end of the unit. The materials provided a scaffold to this level, based on both learning about and using the summary statistics and other necessary statistical ideas.

> A home-owner was interested in whether the Sunday Mail (newspaper) and The Realtor (a real estate sales paper produced weekly) contained houses that were for sale for reasonably similar prices in general or whether one paper contained houses of generally higher prices. To investigate this he randomly selected 100 homes advertised in the Sunday Mail and 101 from the Realtor over a period of three weeks and recorded the asking price for each house.

Students are guided to produce a graphical display (a pair of Stemplots/Histograms) to compare the shapes of the two distributions, and then compare the centres of the distributions, the spread of the distributions and finally to make box plots to see what these reveal. They are directed to summarize their findings in a table and form an argument based on their analysis that supports their answer to the question or solution to the problem. A model output (taken from the materials) is shown in Figures 1 – 3, Table 1 and the text below.



Figures 1-3. Histogram and box plot representations of the real estate data as produced by a graphic calculator; histograms have a common scale.

Table 1
Summary Table of Analysis of Real Estate Data

	Sunday Mail	*Realtor*
Outliers	none	none
Shape	Approximately uniform	Skewed to the high
Median	$129,950	$84,950
IQR	$54,950	$24,750
Boxplot story	Over three quarters of Sunday Mail prices were higher than three quarters of Realtor prices, around 50% of Realtor prices were less than around 90% of the Sunday Mail prices.	

No abnormally high or low house prices were found in the asking prices of homes collected from either the Sunday Mail or Realtor. The Sunday Mail's price distribution is reasonably uniform, while that of the Realtor is skewed high. The median asking price in the Sunday Mail sample is $129,950 compared to a much lower $84,950 for the Realtor sample. The asking prices in the Sunday Mail sample show far more variation than the Realtor. The interquartile ranges are $54,950 and $24,750 respectively. It is also worth noting that over three quarters of Sunday Mail prices are higher than three quarter of realtor prices and that around 50% of Realtor prices are less than 90 percent of the Sunday Mail Prices.

The analysis of our samples support the hypothesis that the asking prices of houses in the Sunday Mail are, in general, considerably higher than houses advertised in the Realtor.

Potential problems regarding the materials

For most students this was their first serious look at analyzing this sort of data and there were many new things to learn, such as stem plots, histograms, the concept of distribution, shapes of distributions, median, IQR, box plots, and so on, not to mention the skills required to construct a sound argument.

In reality, the intended focus (the ability to construct a sound argument) came at the end of a rather long chain of other skills which had to be learned. For all but the most capable students, the learning of the mechanical skills dominated. There was also evidence that teachers glossed over the main focus for reasons which have not been fully investigated.

It was possible for students to follow a formula to form an argument based on the framework and examples they were given. Too many students tried to apply a learned procedure and it was evident from their attempts that no real statistical reasoning had taken place. Some teachers encouraged this formulaic approach as it was worth marks in the test.

All the problems posed required the students to read within an individual data set, between a pair of data sets and ‘beyond the data’ (Curcio, 1987). That is, the student had to make a comparison between the two sample data sets and then hypothesize about what that may mean about the population from which the data were drawn. As an initial expectation this seems to be too much to ask of many students.

Most examples in the material provided data sets of which the students had no ownership. This made it difficult for most students to read beyond the sample data and think about what it may mean in terms of the population the data came from; they had little familiarity, for example, with house prices or with the target audience of the two newspapers. Just building an understanding of what the population is for a given sample data set, is not a trivial task. Did the statement "The analysis of our samples supports the hypothesis that ….." mean much to the students, or was it just the thing you had to write to get a mark?

What was it that actually promoted the need for students to learn and use things like histograms, box plots, the mean, the IQR, and so on? It seemed there was nothing apart from the fact that we were telling them they were useful tools.

Despite these issues, the materials were and still are received very well. I suspect that says more about the materials teachers were using previously, or what they previously thought of statistics, than it does about the actual quality of the 1998 materials. However, the findings presented above led to a series of questions that had to be answered before a set of materials could be written and tested to replace those presently being used.

Addressing issues with the materials

The first question that needed to be addressed in order to improve the materials was the following: "Is there a better sequence of learning that could be employed?" In response, I have developed the following four phases that form a learning sequence that largely relates to data measured on an interval scale.

1. Students work with a single variable that requires them only to 'read the data' (Curcio, 1987) for description purposes (i.e., reading beyond it is not possible, the 'population' is before you). This process, from here on, will be referred to as *reading within a distribution*.
2. Students work with two variables where they are required to read only within the distribution of the variable measured on an interval scale and between two distributions of this variable but for two different categories, but *not* beyond (i.e., reading beyond it not possible, the 'population' is before you).
3. Students work with a single variable that requires them to read within a distribution and beyond it as well (i.e., inferring to a population because acquiring data from a sample of the population is all that is reasonably possible).
4. Students work with two variables (as in phase two above) that require them to read within each distribution, between the two distributions and beyond the distributions.

Whether or not this sequence is an improvement is still to be determined, but it clarifies the learning trajectory implicit in the materials and, as a result, makes it easier to conceptualize each of the learning activities in the context of the course. This is a higher-order issue that helps with revising the materials. This clarification may also help teachers stay grounded regarding the course goals and stay focused on the appropriate learning issues while teaching with the materials. Towards these ends, the following content clarifies what types of tasks relate to each of the phases presented above.

Sample tasks for each phase

Phase One (reading within a single distribution) and Phase Two (reading between two distributions of categorizations of one variable) present a problem/question where the student has *all* of the data associated with the problem posed, as opposed to a sample from the population. This allows the student to singularly focus on describing what is seen. An example of a suitable task for Phase One is: "Describe the batting performance of the Australian cricket team during the 1990s."

Phase Two consists of tasks that are two-dimensional forms of the questions from Phase One. This provides added interest and the chance to make a more powerful use of some simple tools. The students would be comparing and contrasting, and developing an argument to support differences or similarities. Such a task is: "The cricket team of which country, Australia or England, had the better batting performance during the 1990s?"

Phase Three (reading within and beyond) and Phase Four (reading within, between, and beyond) provide a challenge that leads to the second question, “How can we supply students with sample data sets for Phase Three and Four tasks that are of a sensible size (not too small), of a type that students will feel a sense of ownership over, and that are drawn (preferably by the students) from a population that the students can relate to with little effort?” For a busy teacher it is very difficult to find data that meets these requirements. Databases such as the *SeniorSchoolCensus-online* can significantly ease the teacher's burden.

In 2003, *SeniorSchoolCensus-online* offered all students in South Australia the chance to take part in a secure online census (HREF, 2004). The survey had 34 questions, some providing categorical data, others providing data measured on an interval scale. The survey can be viewed by visiting the project website (www.censusonline.net). Over 21,500 students took part. The respondents form a *population*. Some of the population parameters have been released, but most have not. This offers students the chance to genuinely experience the problems that face a real statistician. Using the web-based *Sampler*, students can draw simple random samples of up to 255 cases. Most importantly, students can define the attributes of the individuals sampled using the unique interface. Figure 4 shows the set up used to draw a simple random sample of 200 Year 8 males who attend public schools and live in the central suburbs of South Australia.

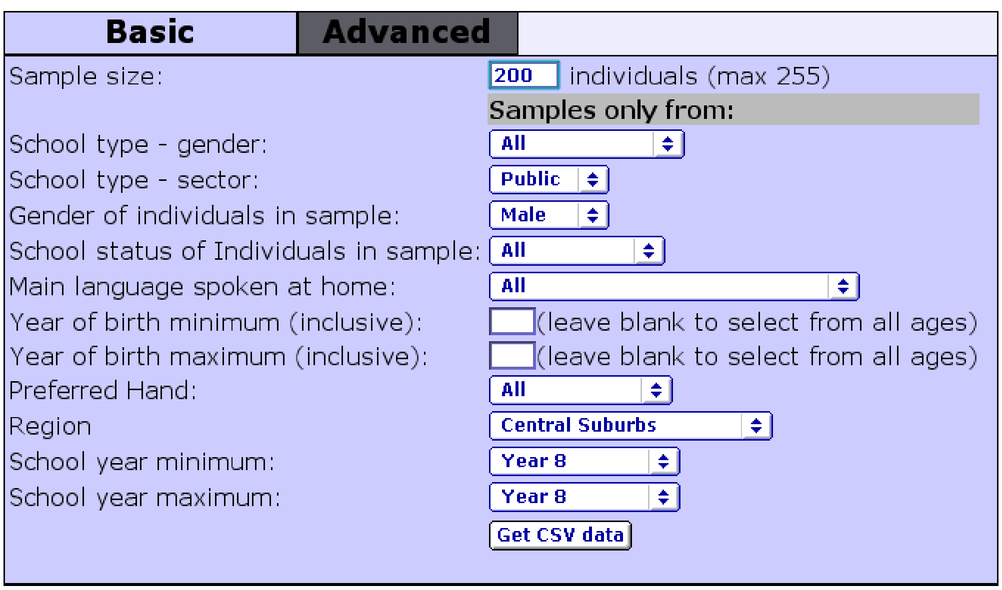
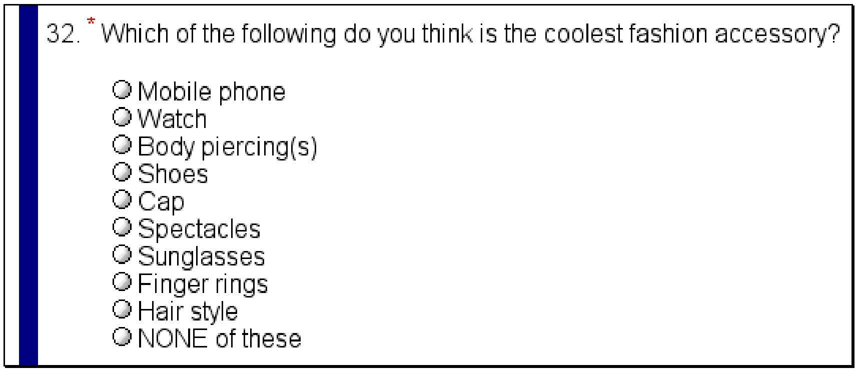


Figure 4. The Simple Random Sampling Device of *SeniorSchoolCensus-Online*

This simple random sampling capability allows students to investigate all sorts of interesting things about the population, a population to which they belong, and parameters about which they have considerable knowledge. One question asked in the survey from the *SeniorSchoolCensus-online* project is shown in Figure 5.

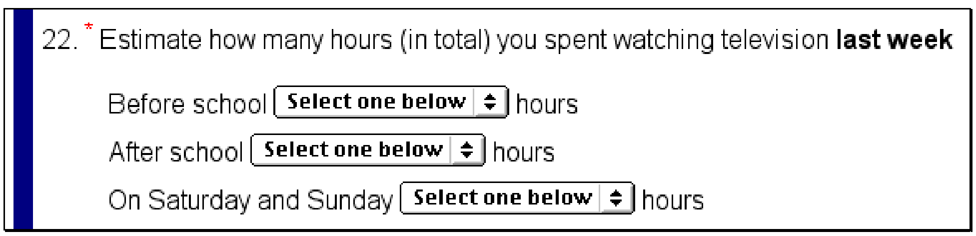

32. * Which of the following do you think is the coolest fashion accessory?

- Mobile phone
- Watch
- Body piercing(s)
- Shoes
- Cap
- Spectacles
- Sunglasses
- Finger rings
- Hair style
- NONE of these

Figure 5. Question 32 from *SeniorSchoolCensus-Online*, as seen in the survey

This task is suitable for use during Phase Three of the learning sequence. The following question could be asked of students analyzing these data: “What do Year 8 boys tend to think is the coolest fashion accessory?” Students can take a simple random sample from the population and proceed with the analysis of their sample (which will of course be somewhat different from other students’ samples), reading within and then beyond the data.

Figure 6 presents another question asked in the survey.


22. * Estimate how many hours (in total) you spent watching television **last week**

Before school [Select one below] hours

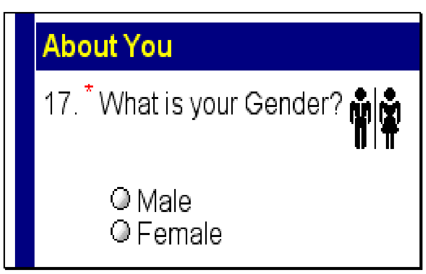
After school [Select one below] hours

On Saturday and Sunday [Select one below] hours

Figure 6. Question 22 from *SeniorSchoolCensus-Online*, A as seen in the survey

The following question could also be asked of students analyzing these data during Phase Three: “How much television did Year 8 students watch on Saturday and Sunday last week?”

Also asked as a part of the survey is the question seen in Figure 7.


About You

17. * What is your Gender?

- Male
- Female

Figure 7. Question 17 from *SeniorSchoolCensus-Online*, as seen in the survey

A task suitable for use during Phase Four of the learning sequence could involve the data generated from the survey question in Figure 7 along with the following question: “How, if at all, do Year 8 boys differ from Year 8 girls in terms of what they regard as the coolest fashion accessory?” Students can take a simple random sample of Year 8 boys and then a simple random sample of Year 8 girls from the population and proceed with the analysis of their samples.

Another question asked in the survey is seen in Figure 8.

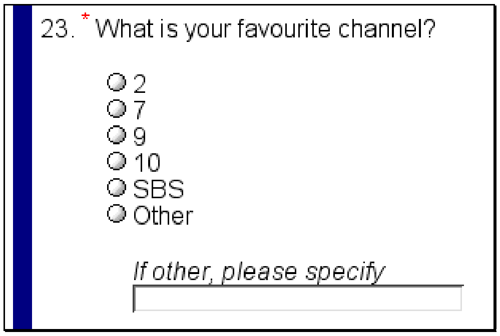


Figure 8. Question 23 from *SeniorSchoolCensus-Online*, as seen in the survey

Channels 2, 7, 9, 10, and SBS are free-to-air channels. The students who responded 'Other' claimed a pay TV channel was their favorite channel. So, the following question could be asked of the students analyzing these data during Phase Four of the learning sequence: "Did Year 8 students who stated their favorite TV channel was a pay TV channel watch more or less TV on the Saturday and Sunday last week compared to those who stated a free-to-air channel was their favorite?"

Having addressed the first two questions regarding the learning sequence and sources of appropriate data, a series of other questions followed. The questions include (questions one and two are re-stated for completeness):

1. Is there a better sequence of learning that could be employed?
2. How can we supply students with sample data sets for Phase Three and Four tasks that are of a sensible size (not too small), of a type that students will feel a sense of ownership over, and that are drawn (preferably by the students) from a population that the students can relate to with little effort?
3. If we accept that we want students to use data to answer questions and solve problems, then is it realistic to have students 12-14 years of age (or younger) forming arguments that support conclusions and, in some cases, conjectures about populations as a result of reading within, within and between, and so on, in a manner that is not driven by some formulaic process?
4. If we suppose the answer to the previous question is yes, then what is the minimum set of tools required to equip students to form satisfactory conclusions and supporting arguments? The existence of a minimum tool set may provide the time for students to come to grips with the process of concluding and arguing and allow them time to understand more deeply the context of the problem at hand and how to solve it. The main outcome would then not be on the end of a long line of other outcomes.
5. When should these phases be implemented?

I propose that the minimum set of tools required by a student to be able to produce useful conclusions and supporting arguments to questions similar to those posed earlier are:

- a functional understanding of percentages, and
- the ability to be able to quickly and easily produce a simple pictorial representation of the distribution, and slice a distribution into (student defined) sections for which they can determine the percentage of the data points falling in the sections.

Clearly the slicing and percentage calculation process, if done manually, would be too labour intensive. However, recent developments in statistical software have made this a simple process. The software product *Tinkerplots* (Konold & Miller, 2004), provides the tools (among many others) that make these tasks simple and quick. Students can rapidly slice up distributions, consider what information that particular slicing configuration conveys and then try other alternatives. The application of these tools in this way is called *distribution division*. I believe distribution division should precede the use of the more traditional summary statistics like the mean, median, interquartile range, and so on as tools to summarize and describe distributions.

To illustrate how these tools may be used I look at some of the tasks posed earlier. Recall the task: "Describe the batting performance of the Australian cricket team during the 1990s." Using the minimum set of tools found within the *Tinkerplots* environment we could proceed as seen in Figure 9 to analyze the total runs scored in every match played by Australia in the 1990s.

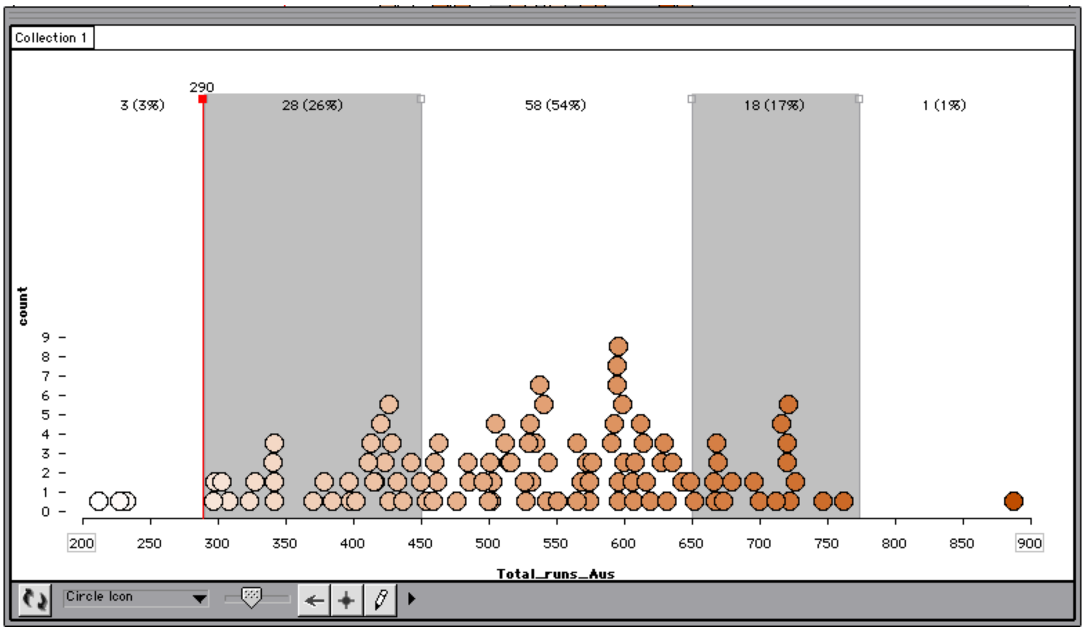


Figure 9. A graphical representation of the distribution of total runs scored in each match played by Australia in the 1990s cut into five divisions.

After creating a distribution of scores, the student can place dividers over the distribution which they can freely position anywhere they like. To these divisions they can also add labels that show the counts and/or percentages lying within each division. The red reference line is another option which helps students read off the location of a divider or case with respect to the horizontal axis. Choosing to slice the distribution into five divisions provides the following description:

> 108 matches were played in total. In 54 percent of the matches, Australia scored between 450 and 650 runs. In 26 percent of the matches between 290 and 450 runs were scored. Greater than 650 runs but less than 770 were scored in 17 percent of matches while there were 4 extreme results.

The number of dividers used and their placement will differ from person to person. An alternative approach would be to consider the overall run rate (runs per over) for each match played, as seen in Figure 10.

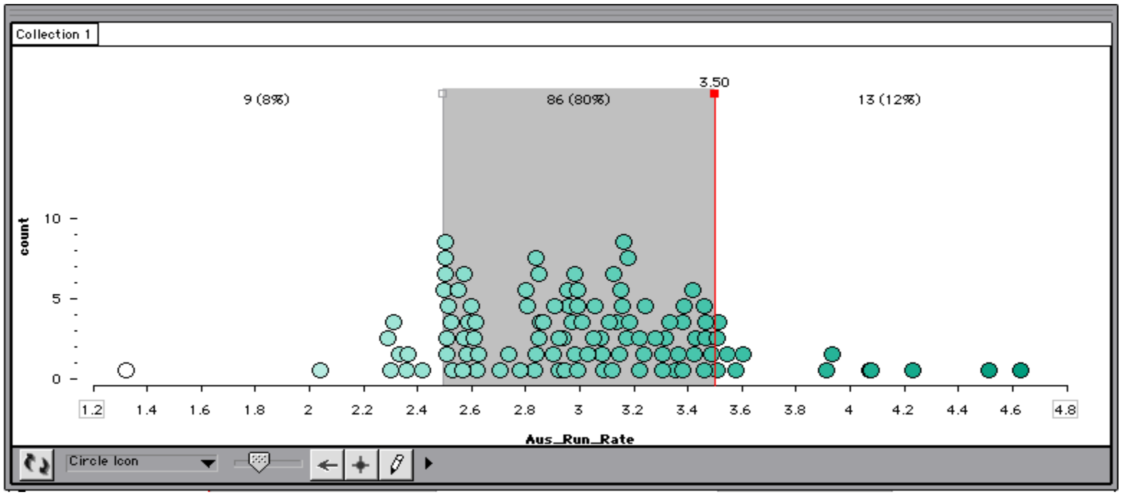


Figure 10. A graphical representation of the distribution of run rate for each match played by Australia in the 1990s cut into 3 divisions.

Here we see that cutting the distribution into three divisions provides good information to describe the performance using this attribute. I have chosen to set the central division to have limits of 2.5 and 3.5 runs per over respectively. Some understanding of the game is needed to make this choice. It should be noted here that the investigation of other variables would be sensible in attempting to perform the task: "Describe the batting performance of the Australian cricket team during the 1990s." It is not the purpose of this article to perform a full analysis to answer this or other questions posed.

Now recall the task: "How much television did Year 8 students watch on Saturday and Sunday last week?" Using a simple random sample of 200 Year 8 boys from the population formed in the *SeniorSchoolCensus-online*, we can proceed as seen in Figure 11.

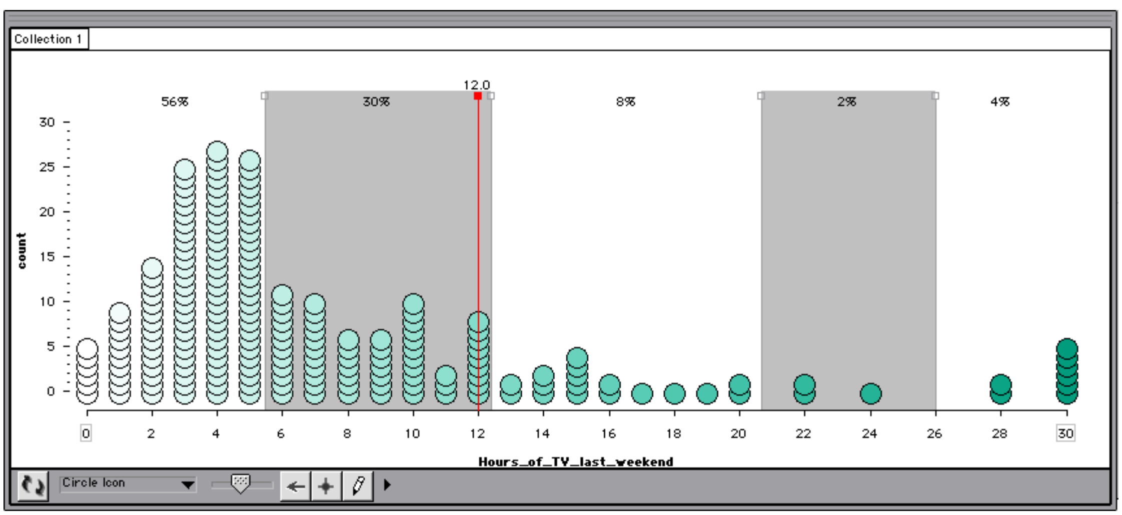


Figure 11. A simple graphical representation of the distribution of hours of TV watched cut into 5 divisions

Choosing to cut the distribution into five divisions provides a good summary of the distribution, clearly highlighting the fact there are eight extreme responses. The choice of the number of sections and positioning of the cuts is obviously critical and requires some experimentation, during which a student would become more familiar with the distribution. It is now open for the student to consider what this sample may tell us about the population, giving them the opportunity to form a conjecture about the population. Research is needed to investigate how students would approach this task and how successful they would be.

Now recall the question: "The cricket team of which country, Australia or England, had the better batting performance during the 1990s? Using the minimum set of tools (within Tinkerplots) we can proceed as seen in Figure 12.

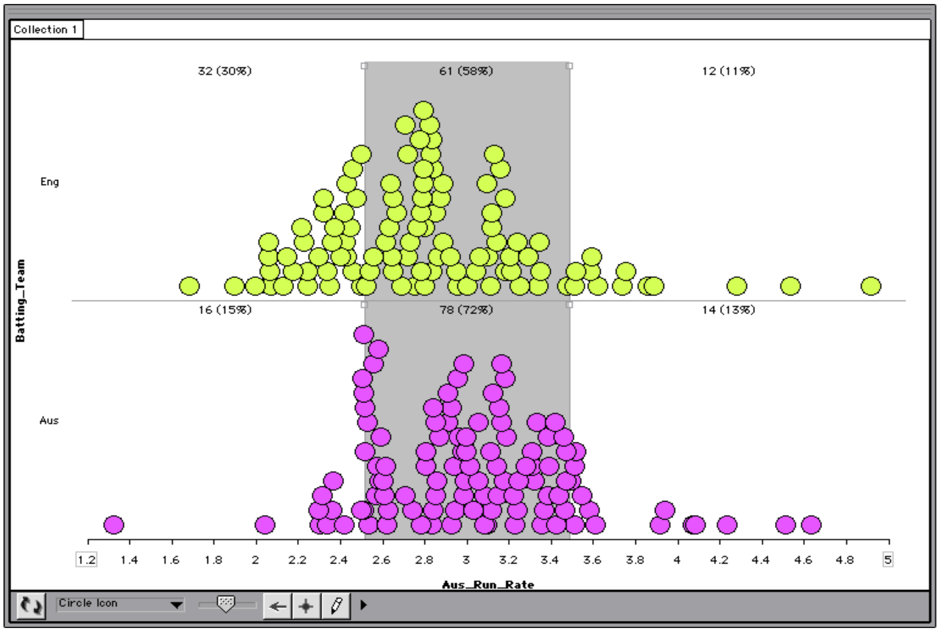


Figure 12. A graphical representation of the distribution of run rate for each match played by Australia (below) and England (above) in the 1990s

Choosing three divisions and setting the cuts to be the same as in the Australia only analysis (Figure 10) allows us to see that 14 percent more of Australia's run rates were between 2.5 and 3.5 runs per over. Also, Australia had 2% more run rates in the upper range than England. So one could ask where are the 16 percent of the English run rates that would be in between these limits if the teams' performances were generally the same? It is clear that they are in the lower zone. This supports the argument that Australia had the better batting performance in the 1990s.

Now recall the question: "Did Year 8 students who stated their favorite TV channel was a pay TV channel watch more or less TV on the Saturday and Sunday last week compared to those who stated a free-to-air channel was their favorite?" Using a random stratified sample of 400 students, comprising 200 Year 8 boys and 200 Year 8 girls we can proceed as seen in Figure 13.

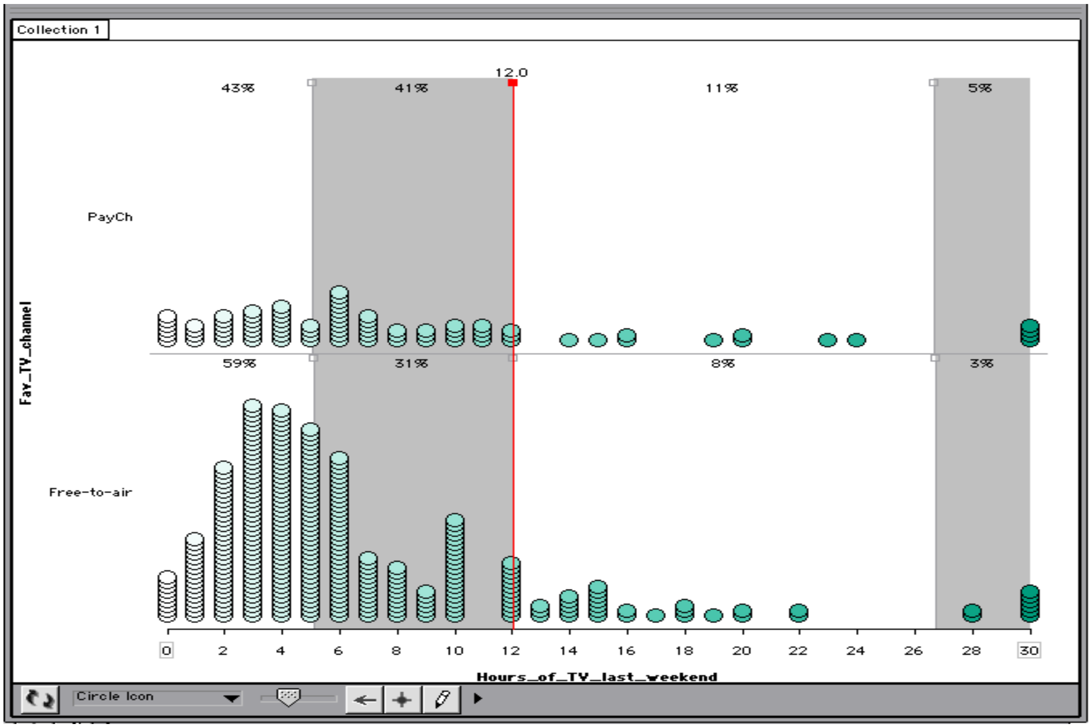


Figure 13. A graphical representation of the distribution of hours of TV last weekend for a sample of 400 students broken down by form of favorite channel

The choice of number and placement of divisions in this case was based on dividing the pay TV channel distribution in a satisfying way and then comparing the free-to-air distribution when divided the same way. We can use a similar analysis to that used in the previous example, noting that the pay TV distribution has 10 percent more students in the 5 to 12 hour range than does the free-to-air distribution. We can then look to see where that 'missing 10 percent' are, and so on, suggesting that a difference does exist between these distributions and that this difference may support the conjecture of a difference in the population.

Obviously, this type of analysis is not possible without an interface like that of *Tinkerplots*. It also requires considerable perseverance from students in experimenting to achieve a division structure. I do not see this as a problem and suggest that it is a good thing. At this stage, my classroom experiences support the claim that this sort of analysis is a reasonable expectation of young students and a better way to initiate their data analysis journey.

One interesting question, specific to Tinkerplots, is whether or not the colour gradient feature, when representing variation in the response variable, would help or hinder the division process. When a certain attribute is selected, the icons representing each case are coloured. The level of colour (light to dark) of a given icon indicates to the student the relative value of the attribute for this case. You will note that in all figures except for Figure 12 the colour gradient represents variation in the response variable. Figure 12 shows an alternative use of the colour gradient feature.

Having said this, the color feature in Tinkerplots is a powerful one. Its real use is in making visible a variable from the dataset that is not shown otherwise. If this is a numerical variable, it will be represented by the colour gradient. If it is a categorical variable, it will be shown by discrete levels of colour. Software that enables us to visualize several variables at once, and that works interactively, has become increasingly available in the workplace since the 1980s. The strength of software like Tinkerplots is that it makes these features accessible to, and attractive to, children.

Distribution Division as a complementary tool in Phases 1-4

I propose that distribution division provides a method that allows students to make the process of building an argument the focus of their learning rather than it being on the end of a long chain of other skills which have to be learned as it was in the 1998 materials. It provides some level of structure for students to work within, but not so much that students are able to follow some pre-learned formula, as was also the case in the 1998 materials.

I also propose that the distribution division method provides the opportunity for students to really appreciate the concept of *distribution* – something so poorly understood by many students (and adults). This approach may also work towards breaking down the tendency students have to think of distributions in three parts, the majority in the middle, low values and high values (Bakker & Gravemeijer, 2004; Konold et al., 2002). The chance of this would increase if more of the numerical variables with which students worked displayed non-normal tendencies. In addition, if students have a greater appreciation of the concept of distribution, the distribution division approach should provide an excellent foundation that will aid in the understanding of modal clumps, outliers, symmetry and asymmetry, and probability distributions that will be faced by students later in their learning.

Finally, this approach employs the same methods used in analysing categorical data and data measured on an interval scale, making the need for the distinction at this stage unnecessary. This raises the question as to when the distinction between different forms of data should be made. It must be made, but I suggest not prior to Phase Four-type investigations. The icon colouring feature of *Tinkerplots* may well lead students to this distinction. Distinctly different colours are used to represent the response levels of categorical data as opposed to the single colour gradient used for data measured on an interval scale.

If students are to function well at higher levels of school and university, they must move on from the analysis techniques outlined in this article. The ability to locate modal clumps, identify outliers, identify symmetry or asymmetry and apply summary statistics like the mean, median, interquartile range and standard deviation all need to be developed (Phases Five and Six). Therefore, we need to consider what sort of learning experiences would motivate this 'moving on'.

Complementary tools beyond Phase 4

The reason for needing different tools to those in the minimum set must be clear to students. Summary statistics like the mean, standard deviation and so on could be described as more 'efficient' tools in the sense that, to a mature user of these tools, two numbers and a description of the shape of a distribution can provide a fair indication of what the distribution is like. This could be one reason for introducing them. However, a far more powerful approach is to provide students with a problem for which the minimum tool set becomes cumbersome. One such task could be: "The cricket team of which country, Australia, England, India, South Africa or New Zealand had the best batting performance during the 1990s?"

This task is considered a Phase Five task in which students use the more traditional tools in situations where *all* the data is before them. Attempting a division type analysis would be a serious challenge (five explanatory levels to compare and contrast) and the thought of it, or attempting it, would hopefully have students asking, "Is there a better way?" The need for different tools would then clear to the students. Here, the introduction of a tool like the mean provides a simpler way to compare the distributions. Other tools, their use and relative merits could follow.

A Phase Five task like that described above, together with tasks involving two categorical variables where one has numerous levels, provides the logical motivation to explore the differences in the forms of data that students have previously been working with. Hence, the nominal, ordinal, interval distinction can be made and alternative forms of analysis can be developed. In addition to this, students should be able to begin to appreciate and identify the structure of the data sets they are using, realizing that most data sets they use are in fact multi-variate. Robert Hall's matrix represents one attempt to draw attention to this most important feature of the data sets with which we work.

A Phase Six task would include the opportunity for the students to think beyond the sample they have to a population. Starting with familiar populations and then moving to non-familiar situations would be desirable. *SeniorSchoolCensus-online* offers the opportunity to create numerous tasks of this nature.

Learning materials based on the phases proposed in this article and the process of distribution division will be used by students throughout the latter part of 2004 and early in 2005. After this has occurred, conclusions about the success (or otherwise) of these ideas in enabling students to make reasoned decisions will be made.

Effects on curriculum

The potential effects on the curriculum of the material and ideas outlined in this article are considerable. The timing of the introduction of the phases is an important consideration. I suggest that Phase One- and Two-type activities be implemented as soon as students are operational with percentages. In Australia this would be when students are approximately 12 years of age. I would suggest that Phase Three and Four type activities be implemented in the following one to two years. Problems involving categorical data could be included during this time as well as problems like those seen earlier.

This would mean that students aged approximately 15 years would be ready to move on to forms of analysis that require them to locate modal clumps, identify outliers, identify symmetry or asymmetry, apply summary statistics like the mean, median, interquartile range, standard deviation and make use of more complex graphical displays like the box plot. It would also be at this age that students could begin to make formal distinctions between the types of variable they are analyzing and how the analysis differs for each.

The major impact on the curriculum for students of age 12 years or thereabouts is that the main outcome of their learning can be the development of the skills required to form an argument to support a conjecture. This has been an elusive outcome for a very long time.

Summary

The six phases discussed in this article offer a developmental pathway that could have students making decisions based on sound statistical reasoning. Distribution division allows students to work with very simple tools, the basics of which would have been previously learned, which in turn allows them to build an intimate knowledge of the distribution. It also allows them to focus on the formation of a supporting argument, rather than being expected to do this on the end of a long chain of new learning. The only new learning in Phases One to Four is the art of distribution division, the application of the skills of comparing and contrasting and forming arguments, based on fact, that support a conclusion or conjecture. Traditionally our students have been initially taught to apply Phase Five and Six concepts to problems from Phases One to Three; this may well have been compounding their statistical illiteracy.

The approach outlined in this paper offers the possibility that students will have a sound foundation on which to build when first meeting concepts and problems from Phases Five and Six. If we suppose the proposed phases are desirable, then phases prior to these six exist. Some thought and some minor teaching experiments have been done in this area, but it is not the point of this article to elaborate on further.

References

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

Curcio, F.R. (1987). Comprehension of mathematical relationships expressed in graphs. *Journal for Research in Mathematics Education, 18*: 382-393.

HREF: (2004), *SeniorSchoolCensus-online*. Available on the World Wide Web at www.censusonline.net

IASE organising committee, (2004), IASE Roundtable theme.

Konold, C., Robinson, A., Khalil., Pollatsek, A., Well, A.D., Wing, R., et al. (2002). Students' use of modal clumps to summarise data. In B. Phillips (Ed.), *Proceedings of the International Conference on Teaching Statistics [CD-ROM]*, *Cape Town*. Voorburg, the Netherlands: International Statistics Institute.

Konold, C., & Miller, C. (2004). *Tinkerplots*. Data analysis software for middle school curricula. San Francisco: Key Curriculum Press.

Lupton, A. (2004). *Learning and achieving through statistical investigation*. Unpublished.

Moore, D., & McCabe G. (1996). *Introduction to the practice of statistics*, NY:Freeman.

Senior Secondary Board of South Australia, (1993). *Quantitative Methods syllabus*. Adelaide, Australia: SSABSA.