REASONING ABOUT COVARIATION WITH TINKERPLOTS

Noleine Fitzallen

BSc, BTeach, BEd (Hons), Grad Cert (ICT T&L)

Submitted in fulfilment of the requirements for the Degree of Doctor of Philosophy

Faculty of Education, University of Tasmania

July, 2012

Declaration of Originality

This thesis contains no material which has been accepted for a degree or diploma by the University or any other institution, except by way of background information and duly acknowledged in the thesis, and to the best of the my knowledge and belief no material previously published or written by another person except where due acknowledgement is made in the text of the thesis, nor does the thesis contain any material that infringes copyright.

Signed:

Date: July 13, 2012

Authority of Access

This thesis may be made available for loan and limited copying in accordance with the *Copyright Act 1968.*

Signed:

Date: July 13, 2012

Statement of Ethical Conduct

The research associated with this thesis abides by the international and Australian codes on human experimentation, as set out in the *National Statement of Ethical Conduct in Human Research* (2007) and interpreted by the Human Ethics Committee of the University.

Ethical approval for the research was gained from the *Southern Tasmania Social Sciences Human Research Ethics Committee* at the University of Tasmania in 2006 – Ethics Approval Number H8778. The committee adheres to the guidelines outlined in the *National Statement on Ethical Conduct in Human Research.* The research also had permission and approval from the Department of Education, Tasmania, and satisfied department criteria for *Conducting Research in Tasmanian Government Schools and Colleges* (2006).

Signed:

Date: July 13, 2012

Statement of Co-authorship

The following people and institutions contributed to the publication of the work undertaken as part of this thesis:

Paper 1 *Noleine Fitzallen (80%), Natalie Brown (20%)*

University of Tasmania

Fitzallen, N., & Brown, N. (2006b). Evaluating data-analysis software: Exploring opportunities for developing statistical thinking and reasoning. In N. Armstrong & C. Sherwood (Eds.) *IT's up here for thinking. Proceedings of the Australian Computers in Education Conference, Cairns,* October 2-4, 2006. [CDROM]

Details of the Authors' roles:

Noleine Fitzallen contributed to the idea, its formalisation and development.

Natalie Brown assisted with the refinement and presentation.

Paper 2 *Noleine Fitzallen (80%), Natalie Brown (20%)*

University of Tasmania

Fitzallen, N., & Brown, N. (2007). Evidence-based research in practice. In P. J. Jeffrey (Ed.), *Engaging Pedagogies. Proceedings of the Australian Association for Research in Education 2006 International Education Research Conference, Adelaide, Nov 26-30, 2006.* [CDROM]

Details of the Authors' roles:

Noleine Fitzallen contributed to the idea, its formalisation and development.

Natalie Brown assisted with the refinement and presentation.

Paper 3 *Jane Watson (60%), Noleine Fitzallen (40%)*

University of Tasmania

Watson, J. M., & Fitzallen, N. E. (2010). *Development of graph understanding in the mathematics curriculum. Report for the NSW Department of Education and Training.* Sydney: State of New South Wales through the Department of Education and Training. Retrieved September 7, 2010, from http://www.curriculumsupport.education.nsw.gov.au/primary/mathematics/assets/pdf/dev_graph_undstdmaths.pdf

Details of the Authors' roles:

Jane Watson contributed to the idea, its formalisation and development.

Noleine Fitzallen contributed the *Historical Developments in Graphing* and the *Current Graphing Curricula* sections. She also contributed to the refinement and presentation of the report. The report is independent of the thesis and does not include any data used in the thesis. The *Historical Developments in Graphing* section is reproduced in part in this thesis. The authorship of other information from the report included in the thesis is duly acknowledged.

We the undersigned agree with the above stated "proportion of work undertaken" for each of the above published (or submitted) peer-reviewed manuscripts contributing to this thesis:

Prof Jane Watson	Assoc Prof Karen Swabey
Supervisor	*Head of School*
School of Education	*School of Education*
University of Tasmania	University of Tasmania

Signed: Signed:

Date: Date:

Acknowledgements

My life's journey has taken many twists and turns, providing me with the opportunity to experience a number of different careers and meet a diverse range of people. My partner once quipped "You haven't made up your mind what you want to be!" This may be the case but to me it does not matter. I want to explore new things, places and ideas. I want to continue to grow and develop and embrace new opportunities. This thesis is a reflection of my personal experiences and achievements and is influenced heavily by my family, people I have worked with and people I have not met, yet know through their research. Some people assisted greatly in the production of the actual artefact but many more had an impact on why I was in this place, at this time, to be able to embrace the opportunity of conducting research when it was offered. I would like to acknowledge that this thesis is a culmination of all these influences.

The story in this thesis is about 12 students. They graciously let me into their lives to extend, question, and examine their thinking. It was a very rewarding experience both professionally and personally. I am extremely grateful for the opportunity to work with them. My personal thanks go to Blaire, Jake, James, Jessica, Johnty, Kimberley, Mitchell, Natalie, Natasha, Rory, Shaun, and William.

Special thanks are extended to my supervisors Jane Watson and Natalie Brown. Among other things, Jane allowed me the freedom to be creative, found the time to debate points of view to extend my thinking, and had the patience to assist me to improve my writing. Natalie valued my ideas, challenged me to think in different ways, and supported me to publish my work.

I must also thank the designers of *TinkerPlots*, Cliff Konold and Craig Miller. The creative and innovative approach taken to develop *TinkerPlots* has resulted in the

development of innovative data analysis software that has these features imbedded in its design. Consequently, students using *TinkerPlots* are empowered to be creative and innovative in return.

This research was funded by an Australian Postgraduate Award – Industry Scholarship associated with the Australian Research Council Linkage Project LP0560543 and the industry partner, Department of Education, Tasmania (DoET). The linkage project, *Providing the Mathematical Foundation for an Innovative Australia within Reform Based Learning Environments* (MARBLE) (Watson, Beswick, & Brown, 2011), provided the context for this thesis.

Abstract

Covariation is recognised as an important aspect of statistical thinking and reasoning and is used to explore the relationship between two attributes. Often, covariation is determined from the interpretation of scatterplots that display the correspondence of two numerical attributes and is described as a trend in the data. Scatterplots are utilised when conducting exploratory data analysis (EDA). EDA strategies are useful for interpreting the data as they allow the data to be manipulated in order to construct graphical representations that facilitate making sense of the data. The translation of EDA strategies into innovative software packages, such as *TinkerPlots: Dynamic Data Exploration*, has placed student learning about data analysis in technological environments and there is a need to investigate the way in which students learn in these contexts.

This inquiry had two objectives. The first objective was to further understanding of the factors that influence student learning when working with software packages. This is through the development of a conceptual framework for learning in EDA graphing environments that aligns with and extends current research about student understanding of graphing and data analysis. The second objective was to explore the intersection between the students' thinking and reasoning about covariation and the influence of *TinkerPlots* on that process, as students explore data sets to determine the relationship between variables and identify trends. To realise these objectives the following research questions are explored:

1. How can the learning behaviours of students as they engage with exploratory data analysis software be characterised through a framework that can then be used to explore and analyse students' understanding of covariation using *TinkerPlots*?

2. How do students interact with the exploratory data analysis software, *TinkerPlots*, to represent data in a variety of forms when exploring questions about relationships within a data set?

3. How do students demonstrate an understanding of covariation in the exploratory data analysis software environment afforded by *TinkerPlots* and use these understandings to provide informal justification for their conclusions about the relationships identified?

The inquiry employed an educational design research methodology within a pragmatist paradigm to facilitate the development of a systematic iterative study. The methodology was chosen to encapsulate the way students learn about the interpretation of graphical representations, more specifically related to covariation, in the technological software environment afforded by *TinkerPlots*.

The inquiry was enacted across seven stages. Stage 0 involved the development of the research design. Stage 1 involved the development of a conceptual framework for learning in EDA software environments that incorporated four aspects of graphing and data analysis skills – *Generic knowledge, Being creative with data*, *Understanding data,* and *Thinking about data*. Stage 2 involved an evaluation of *TinkerPlots* to determine its usability as a teaching and learning tool. Stage 3 involved the development and evaluation of an assessment tool to determine the prior learning of students in relation to the interpretation of graphs, and select the participants for the data collection stage of the inquiry. Stage 4 involved the development and implementation of a sequence of learning experiences. The activities in the sequence of learning were based on recommendations from the research on the development of graphing and data analysis skills. The sequence of learning experiences was implemented with 12 students working in pairs, twice a week for 45 minutes, over a period of 6 weeks. In addition, the data generated from individual interviews with the 12 students conducted at the end of the sequence of learning were included in this stage. The data from the student interviews are presented as Student Profiles that encapsulate the way in which they used *TinkerPlots* to develop not only an understanding of covariation but also develop other data analysis skills and strategies. Stage 5 involved the presentation of the results for the Research Questions, with the discussion of the findings, implications of the inquiry, and recommendations for future research included in Stage 6. The presentation of the thesis follows this chronological order.

Through the evaluation of *TinkerPlots* and its subsequent implementation in the inquiry, it was identified that *TinkerPlots* provides a powerful learning environment for supporting students' understanding of covariation. In terms of student understanding of covariation, the inquiry identified that young students are able to reason about covariation and display three levels of reasoning. The results also suggest that students adopt three different strategies when accessing the features of *TinkerPlots* while creating and interpreting graphs. These strategies are: *Snatch and Grab*, *Proceed and Falter*, and *Explore and Complete*.

Outcomes of the inquiry are presented in relation to the thesis-developed *Model of Learning Behaviour in EDA Graphing Environments*. Within the framework of the model the students' development of covariation reasoning is revealed and discussed in terms of the potential of the results to inform the teaching and learning of covariation within EDA software environments and future curriculum development. Consideration was also given to the merits of the *Model of Learning Behaviour in EDA Graphing Environments* and its application throughout the inquiry process. Unexpected insights into the students' thinking and reasoning about association are also discussed to demonstrate the utility of the thesis-developed model and to highlight the need to further research in the area of student understanding of association.

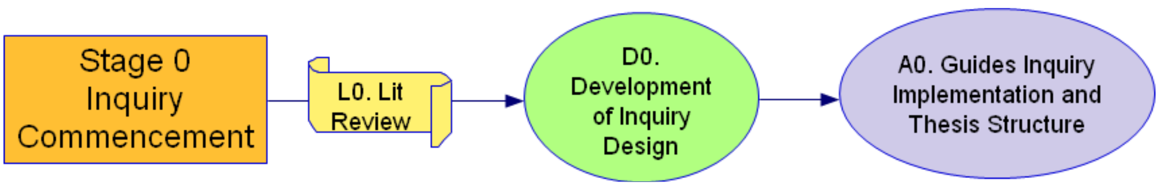
Table of Contents

List of Tables

List of Figures

Stage 0

Inquiry Commencement



Covariation is recognised as an important aspect of statistical thinking and reasoning and is used to explore the relationship between two attributes. Often, covariation is determined from the interpretation of scatterplots that display the correspondence of two numerical attributes and is described as a trend in the data. Scatterplots are utilised when conducting exploratory data analysis (EDA). EDA strategies are useful for interpreting the data as they allow the data to be manipulated in order to construct graphical representations that facilitate making sense of the data. The translation of EDA strategies into innovative software packages, such as *TinkerPlots: Dynamic Data Exploration*, has placed student learning about data analysis in technological environments and there is a need to investigate the way in which students learn in these contexts.

New educational technologies not only are innovations in themselves but also have the potential to empower students to be innovative. They transform student learning by providing a platform for building knowledge and an arena for showcasing creativity (Finger, Russell, Jamieson-Proctor, & Russell, 2007; Scardamalia & Bereiter, 2006). The promises they offer are exciting yet pose many challenges for teachers. The rate of development of new technologies out paces the rate at which teachers can: learn how to use the technologies

(Perso, 2006); develop an understanding of the way in which the technologies influence student thinking (Finger et al.); and develop effective learning experiences for students (Fitzallen, 2005a, 2005b; Fitzallen & Brown, 2006a). Similarly, education research into teaching and learning with technologies is facing the same difficulty and has been of little assistance in guiding teachers towards adopting effective pedagogies. As Roblyer (2005) observes, “Technologies change so quickly that it is difficult to build a body of findings over time on any given application” (p. 193).

To utilise a new technology well, it is necessary for teachers to have a good understanding of what the technology brings to learning experiences and combine it with an understanding of the pedagogies that enhance student learning experiences (Finger et al., 2007). In the case of mathematics education, it is also necessary for teachers to understand the way in which technologies conceptualise the mathematics and support the mathematical thinking required for students to integrate their learning with technologies with their prior knowledge (Scardamalia & Berieter, 2006). This is necessary in order for students to build knowledge of new and complex ideas.

When planning educational experiences with technologies, having an understanding of what particular software packages bring to the learning experience is essential as it enables teachers to become more purposeful in the way they develop and plan learning experiences for students. This understanding allows learning experiences to be designed to exploit the potential afforded by the technologies to influence student learning, rather than merely facilitate the duplication of established learning experiences in a new medium (Roblyer, 2005). To be able to do this well, teachers need to have a well established knowledge of educational technologies as well as understand how to teach specific content with those technologies (Angeli & Valanides, 2009).

In many cases, teachers are neither confident nor comfortable with adopting new technologies. In many cases the teachers feel they have an obligation to include it in their learning programs for students but lack the expertise to do so effectively (Fitzallen, 2005a, 2006a; Fitzallen, Brown, Booth, & Howells, 2008). It is, therefore, understandable that statements such as “you can learn it along with the students” and “if you have problems the students will tell you what to do” are used to compensate for teachers’ lack of understanding of the way in which technology can facilitate learning opportunities (Finger et al., 2007). It can be argued that teachers who adopt pedagogical strategies that rely on assistance from the learners in their classrooms are operating from a deficit view of the potential of the technology utilised and are not in a position to exploit fully the learning opportunities afforded by the technology.

Inquiry Approach

This inquiry investigates the application of the software package, *TinkerPlots: Dynamic Data Exploration* (Konold & Miller, 2005) to student learning in the area of statistics. The inquiry has two main foci – students' development of the statistical concept covariation, and students' interaction with *TinkerPlots* to develop an understanding of covariation. It adopts an innovative research approach to capture the complexity of student learning of covariation, which is influenced heavily by the context of the technological learning environment, student prior knowledge, and instructional design (Moore, 1997). These influences demand the inquiry creates a complex research/learning environment that recognises that the research process and the learning intervention are intrinsically intertwined – one constantly influencing the other. Embedded within this is the idea that the students' learning about statistical concepts and learning about data analysis software are intertwined in similar ways.

In order to investigate students' interactions with *TinkerPlots,* as they work towards developing an understanding of covariation, the inquiry approach needs to be set up as a teaching experiment (Steffe, 1991). This approach assists the researcher to learn about the students' knowledge of covariation and how the students construct their understanding using *TinkerPlots,* as well as experience the students' mathematical thinking and reasoning in the process (Steffe & Thompson, 2000). The complexity of the inquiry setting commands that a rigorous scientific inquiry process is adopted. This type of inquiry facilitates the application of qualitative methodologies that foster theory building by developing a synergy with theoretical frameworks so they guide the inquiry and the inquiry informs the further development of those frameworks (Seeto & Herrington, 2006). Within this milieu, evidence collected in this inquiry provides threads of insight that are unravelled and interrogated individually to reveal the way in which students develop an understanding of covariation within the learning environment afforded by *TinkerPlots*. The intention is also to unveil students' misconceptions or incomplete thinking.

To accommodate these factors, two objectives underpin the inquiry:

- First, a theoretical objective, to extend understanding of the factors that influence student learning when working with software packages, through the development of a conceptual framework that characterises learning in exploratory data analysis (EDA) graphing environments that also aligns with and extends current research about graphing and data analysis.
- Second, a practical objective, to contribute to an understanding of the complex nature of the teaching and learning of data analysis skills, particularly covariation, within the EDA software environment afforded by *TinkerPlots*.

The following research questions are posed:

1. How can the learning behaviours of students as they engage with exploratory data analysis software be characterised through a framework that can then be used to explore and analyse students' understanding of covariation using *TinkerPlots*?

2. How do students interact with the exploratory data analysis software, *TinkerPlots*, to represent data in a variety of forms when exploring questions about relationships within a data set?

3. How do students demonstrate an understanding of covariation in the exploratory data analysis software environment afforded by *TinkerPlots* and use these understandings to provide informal justification for their conclusions about the relationships identified?

In order to ensure that the objectives of this inquiry are realised and evidence is gathered to answer the research questions, this inquiry adopts an educational design research methodology (Akker, Gravemeijer, McKenney, & Nieveen, 2006) that facilitates a knowledge building approach (Scardamalia & Bereiter, 2006) through a qualitative exploratory perspective (Creswell, 2003). The educational design methodology selected facilitates an iterative research approach that increases the internal validity of the inquiry, which in turn contributes to the rigorous nature of the inquiry. The inquiry design is developed in Stage 0 of the thesis and the inquiry stages are used as a structure for this thesis. Figure 0.1 shows the over-arching inquiry objectives with links between each of the research questions and the stages of the inquiry that contribute to answering the questions.

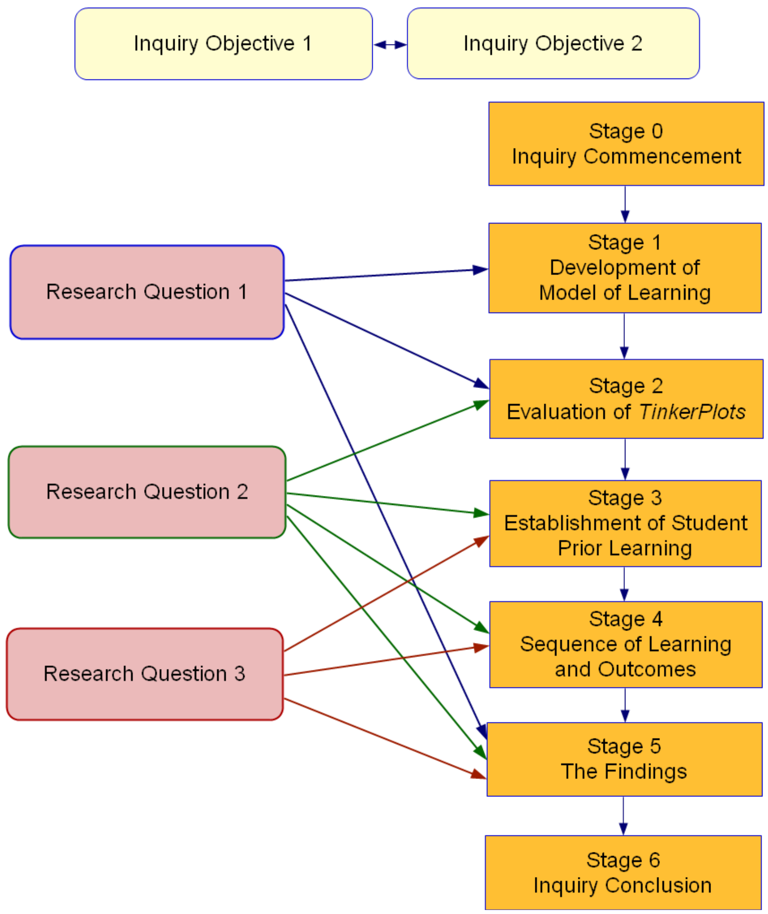


Figure 0.1. Links between research questions and the seven stages of the inquiry.

Inquiry Origins

This inquiry was funded by an Australian Postgraduate Award – Industry (APA(I)) scholarship associated with an Australian Research Council (ARC) Linkage Project and Industry Partner, the Department of Education, Tasmania (DoET).

In 2004, the Department of Education, Tasmania and the Tasmanian Catholic Education Office worked as industry partners with the mathematics education research team at the University of Tasmania to develop a collaborative quantitative literacy project

(Watson, Beswick, & Brown, 2011). Researchers from the three educational institutions were successful in gaining a Linkage Project grant (LP0560543) from the ARC to provide professional learning in mathematics education and evaluate the professional learning programs, in terms of teacher change and student outcomes. Although the main objectives of the project, *Providing the Mathematical Foundation for an Innovative Australia within Reform Based Learning Environments* (MARBLE), focused on the development of professional learning for teachers, the overarching purpose of the project was to improve the teaching and learning of mathematics. Key Curriculum Press provided the software, *TinkerPlots*, to the schools participating in the project.

This PhD inquiry was developed as a companion project to MARBLE. It was supported by one of the Industry Partners – the DoET – as an APA(I) scholarship. As this inquiry is linked to MARBLE, it is proposed that the outcomes will provide the DoET with insights that could inform the development of professional learning programs for teachers. Common to both projects is the context of the middle years of schooling in rural district schools and also aspects of the Chance and Data Strand of the *Tasmanian Curriculum: Mathematics – Numeracy* (DoET, 2007) facilitated by *TinkerPlots*.

Inquiry Methodology

> "…research is not just about descriptions of what happens in good classrooms."
>
> (Conference paper referee, 2007)

The above quotation was taken from a peer reviewed article submitted for a mathematics education research conference. This comment, from one researcher to another, demonstrates the difficulty early-career researchers – and possibly experienced researchers – face when their work is critiqued. Obviously, there are conflicting views of what constitutes quality education research, making it imperative that the choice of methodology is justified effectively to ensure the research is considered legitimate and valid. In this section the methodology adopted for this inquiry is outlined to demonstrate its applicability to the aim and objectives of the inquiry.

Inquiry Aim

The aim of this inquiry is to explore the phenomenon of student learning about the statistical concept covariation, in the complex technological learning environment afforded by *TinkerPlots*. To uncover the significant factors that characterise student learning and their interaction with this technology requires that the focus of the inquiry be on the student (Merriam, 1998, p. 29). In this way the inquiry is positioned to provide research-based insights that have the potential to improve practice and contribute to the building of theory from evidence-based research (Burkhardt & Schoenfeld, 2003; Seeto & Herrington, 2006). To meet the aim of this inquiry, a methodology that is best placed to provide the evidence needed to answer the research questions empirically is required (Barab & Squire, 2004). Selecting an appropriate methodology presents many challenges, particularly when the inquiry needs to produce the best available evidence to address the concerns expressed about what is quality research (Slavin, 2002; Shavelson, Phillips, Towne, & Feuer, 2003).

The debate about evidence-based research challenges what is valued as education research. Mary Kalantzis (2006), in her keynote address, *The Radford Lecture*, delivered at the *Australian Association for Research in Education 2005 International Education Research Conference*, encouraged the education community to consider defining education as a science and as such to hold a view that quality research encompasses more than what is traditionally considered good scientific method. During the address a broad view of what a science of education may be was endorsed. Kalantzis also promoted the idea of education research encompassing the practices and research methodologies developed in education and other social sciences, such as psychology.

A review of the literature reveals two opposing views of evidence-based research. As Kalantzis (2006) alluded to, the first is based on traditional scientific experimental design. The second is based on qualitative research and is termed correlational and descriptive research. Correlational and descriptive research is non experimental research that employs qualitative research methodologies (Slavin, 2002). What is not in dispute is the value of both kinds of evidence-based research to inform government policy and guide the organisation of schools (Erickson & Gutierrez, 2002; Feuer, 2005; Olson, 2004; Slavin, 2002).

Scientific experimental design research. Ongoing calls for education research to adopt scientific principles (Eisenhart & Towne, 2003; Feuer, 2005; Feuer, Towne, & Shavelson, 2002; Kalantzis, 2006; National Research Council [NRC], 2002; Slavin, 2002, 2005, 2008) have created much debate in the education research community. Although there is a general consensus on the value of evidence-based research to guide educational programs and policies, researchers have the challenge of conducting research that is regarded as rigorous and legitimate, in terms of a scientific inquiry.

Slavin (2002) recognises the potential evidence-based research has in developing education policy and the role it plays in transforming educational practices. He also holds a very strong view of the importance on adopting scientifically-based experimental design methodologies when conducting research. Slavin's stance is a reflection of the United States government policy expressed in the No Child Left Behind initiative instigated in 2001, which emphasises and virtually mandates the application of experimental or quasi-experimental designs in education research (Slavin, 2002, 2005).

The experimental design research advocated by Slavin (2002) involves large-scale quantitative studies utilising experimental and control groups. The process adopts the random assignment of the participants or schools into one of the groups and ideally, provides results that can be replicated and generalised. He argues that conducting research in this way will make it easier to determine if teaching interventions are effective. Slavin maintains that research utilising rigorous scientific experimental principles to conduct studies similar to clinical trials will go a long way towards reinstating the confidence in education research. He argues that experimental research is rigorous, scientifically-based, and best placed to inform policy development.

In contrast, experimental design studies are characterised as being "iterative, process focused, interventionist, collaborative, multileveled, utility orientated, and theory driven" (Shavelson et al., 2003, p. 26). Shavelson et al. go on to describe each of these aspects of design studies and draw attention to an important feature of design studies, which is termed "retrospective analysis" (Cobb, McClain, & Gravemeijer, 2003). This involves a situated narrative of the evolving design process. Cobb et al. argue that retrospective analysis in the form of a narrative provides rigorous evidence of the research experience. Shavelson et al.

take a more critical stance and pose a multitude of questions that challenge the veracity of narratives due to their subjectivity and lack of scientific rigor. They do, however, acknowledge that narratives may assist in developing an understanding of the way in which outcomes of education research may be applied in practice as they potentially provide evidence of the issues that impact on the implementation of classroom interventions.

Correlational and descriptive research. Correlational and descriptive research exploits qualitative research methods such as case study and ethnography to explore the relationship between contextual factors within complex learning environments (Erickson & Gutierrez, 2002; Slavin, 2002) and is not to be confused with the statistical connotations of the term "correlational." The focus of this type of research methodology tends to derive from a more practice-based perspective that takes into account the contexts of classrooms and learning environments. It often utilises sustained observation and interviewing to explore the realities of classroom practice (Creswell, 2003) and provide information that goes beyond determining causal factors on the effectiveness of education programs (Slayton & Llosa, 2005). The application of qualitative methods in education research can offer a great depth of information about the success or otherwise of education programs. The exploration of interesting questions can also enhance the information gleaned from a quantitative study evaluating the effectiveness of an education program. Worthy of consideration are questions like: How was the program integrated into the curriculum? What did the teacher do in the classroom to facilitate the implementation of the program? and Why did the teacher focus on particular elements of the program?

Implications for the inquiry methodology. It is clear the two perspectives, scientific design research and correlational and descriptive research, have value and merit in their own right as appropriate research methodologies. They gather different information in different ways and each has a distinctive role to play in developing understanding of curriculum, students, teachers, and learning environments. Some researchers, however, express concern about the potential restrictions imposed by the selection of a particular methodology. For instance on this matter, Feuer (2005) cautions:

> I hope that in our pursuit of more rigor and more science we do not inadvertently narrow the range of questions and the breath of scholarship that is relevant to the

> accumulation of knowledge about teaching, learning, and the organization of schooling.... I am convinced that good science – which, by the way, must not be confounded with any particular methodology used by scientists – is fundamental to continuous learning about and improvement in public education. (para. 12)

St. Pierre (2002) also expresses concern about the ways in which education research may be judged, noting: "Unfortunately, it is often the case that those who work within one theoretical framework find others unintelligible" (p. 26). This comment implies that personal bias and preference for particular research methodologies may influence heavily the opinions of researchers and may hinder the adoption of an objective view of the work of fellow researchers. Fortunately, this is not always the case.

Although the questioning of research methodologies introduces healthy scepticism into the debate over what is quality research, it also provides a basis for the privileging of some methodologies over others. Siegel (2006) suggests that this critical questioning is necessary as there is the need to distinguish between the results of a variety of research approaches.

> [W]hat would be the point … if the results of all research were either worthless or equally legitimate? In these circumstances there would be no point in conducting research, because any result would either be worthless or stand on an equal epistemological footing with any alternative method. (p. 6)

Siegel goes on to suggest that the scrutiny of research should be based on relevant criteria. At this time, when there are calls for evidence-based research that would be considered rigorous and legitimate in terms of scientific inquiry (Kalantzis, 2006), it would be useful to use criteria that recognise the underlying principles of scientific inquiry. The NRC (2002, p. 5) provides six guiding principles that may be useful when developing, critiquing, or evaluating research. The six principles are:

1. Pose significant questions that can be investigated empirically.
2. Link research to relevant theory.
3. Use methods that permit direct investigation of the question.
4. Provide a coherent and explicit chain of reasoning.
5. Replicate and generalise across studies.
6. Disclose research to encourage professional scrutiny and critique.

The six guiding principles are not intended to be a set of standards for conducting and evaluating research but have been developed to support scientific inquiry in education research (NRC, 2002). The NRC note that there are a variety of scientific research designs utilised in education research and caution that the design of research does not necessarily make it scientific.

> To be scientific, the design must allow direct, empirical investigation of the research question, account for the context in which the study is carried out, align with a conceptual framework, reflect careful and thorough reasoning, and disclose results to encourage debate in the scientific community. (p. 6)

Methodological and Theoretical Underpinnings

The orientation of this research aligns with a pragmatist paradigm. Pragmatism seeks to link theory and praxis through the exploration of the research problem (Greenwood & Levin, 2003; Mackenzie & Knipe, 2006). It examines actions and situations to develop an understanding of the meaning of ideas by drawing on qualitative research methods and techniques (Creswell, 2003; Johnson & Onwuegbuzie, 2004). There is an emphasis on developing an understanding of what works and examining solutions to problems to derive knowledge about the problems (Patton, 2002). It follows that the meaning of an idea or a proposition is developed by observing its application in real-world practice (Creswell, 2003). Therefore, a pragmatist approach dictates that research methods are matched to the aim and purpose of the research and the specific questions of an inquiry (Boaz & Ashby, 2003). Although a pragmatist paradigm is usually applied to scientific investigations (Mackenzie & Knipe, 2006), it is appropriate for this inquiry as it allows for the design of this inquiry to be shaped around investigating directly the use of *TinkerPlots* by students.

The theoretical underpinnings of the thesis are rooted in constructivist views of teaching and learning, which recognise the complex and multifaceted relations between educational contexts, content, and participants (Salomon & Ben Zvi, 2006). Many constructivists avoid ideas about individual cognition and instead focus on the social construction of knowledge (Hyslop-Margison & Strobel, 2007). In contrast, the theoretical position taken in this thesis not only recognises that all knowledge is socially constructed (Hyslop-Margison & Strobel, 2007), it also asserts that the teacher plays a pivotal role in creating activities that direct students toward the mastery of cultural tools (Vygotsky, 1978), and in the process of learning students construct a personalised form of understanding that is influenced by active interaction between the learner and the learning environment (Lave, 1988). Hence, learning is conceptualised as the interplay among the students' behaviours, the learning environment, the learning artefacts, the learning activities, which in turn influence the students' reasoning and cognitive development.

Research Perspective

The research methodology for this inquiry embeds within it the principles of scientific inquiry proposed by the NRC (2002). Evidence-based research is valued highly for its potential to inform government policy (Erickson & Gutierrez, 2002), guide the organisation of schools (Feuer, 2005; Olson, 2004; Slavin, 2002), and transform educational practices (Slavin, 2002). Slavin goes further and stresses the importance of adopting scientifically-based experimental design methodologies when conducting such research. With this in mind, an emphasis is put on selecting methodologies that make this inquiry align with the NRC (2002) principles.

The application of experimental design processes that exploit qualitative research methods such as case study and ethnography to explore the relationship between contextual factors within complex learning environments has much support (Cobb, Confrey, diSessa, Lehrer, & Shauble, 2003; Erickson & Gutierrez, 2002; Shavelson et al., 2003). The focus of this type of research methodology tends to be from a practice-based perspective that takes into account the context of classrooms and learning environments. It explores the realities of classroom practice (Creswell, 2003) and facilitates the development of systematic studies (Erickson & Gutierrez, 2002), with an aim to encapsulate "the specifics of practice and the potential advantages from iteratively adapting and sharpening theory in its context" (Shavelson et al., 2003, p. 25).

Shavelson et al. (2003) identified correlational and descriptive research as *design studies* to distinguish them more readily from research that utilised the more traditional experimental design research advocated by Slavin (2002). Design studies were built on the work of Brown (1992) and Collins (1992, as cited in Hoadley, 2004) and furthered by other researchers who include the Design-Based Research Collective (DBRC) (2003), Kelly (2003), Cobb, Confrey et al. (2003), Hoadley (2004), as well as Seeto and Herrington (2006).

Design studies are context bound in nature and employ a series of approaches to investigate learning and teaching in naturalistic settings (Barab & Squire, 2004) in order to develop theories and advance knowledge about a learning construct rather than fine tuning "what works" (Cobb, Confrey et al., 2003, p. 9).

Although Cobb, Confrey et al. (2003) agreed with the view expressed by the supporters of design studies, they refer to their research methodology as *design experiments*, as did Brown (1992), and describe the potential of the research methodology as "addressing the complexity that is a hallmark of educational settings" (p. 9). They contend that design experiments can be utilised in a variety of settings that range from one-on-one experiments with small groups of children through to multi-school, district-wide investigations.

More recently, Akker et al. (2006) were more specific and referred to design studies as "educational design studies" in order to be explicit about the context of the studies and to differentiate them from studies conducted in other contexts. Educational design studies do, however, encapsulate the notions of both design experiments (Brown, 1992; Cobb, Confrey et al., 2003) and design studies (Kelly, 2003; Seeto & Herrington, 2006; Slavin, 2002). The methodology selected for this inquiry is educational design research. For the purposes of this inquiry, the term educational design research is used when referring to methodologies that align with design studies and design experiments.

Educational Design Research Methodology

One of the objectives of this inquiry is to examine the way in which students use the exploratory data analysis [EDA] software package, *TinkerPlots*, to develop an understanding of covariation. This requires close examination of students working at computers whilst using *TinkerPlots*. To accommodate working closely with students at the computer, the research methodology most appropriate needed to be qualitative and interrogative (Merriam, 1998) in order to explore student learning within the complex technological learning environment afforded by *TinkerPlots*. As Creswell (2003) points out, adopting a qualitative research approach allows the researcher to interrogate the process that is occurring as well as the product or outcome to understand how things occur, which aligns with the objectives of this inquiry.

Qualitative research in education is about understanding what teachers and students do in educational settings. The goal is to study the participants' view of the situation (Creswell, 2003). Educational design research methodologies provide the opportunity to interrogate educational settings from this perspective. The methods utilised are selected on their ability "to fit the specifics of the problem and situation, and may consist of any one or a combination of explanatory, interpretive, experimental, computational, mathematical or exploratory methods" (Sloane, 2006, p. 44). To ensure educational design studies are considered credible, rigorous research propositions are needed to address the criticism that qualitative education research often fails to bridge the gap between theory and practice (Cobb, Confrey et al., 2003; DBRC, 2003; Sloane, 2006). The methodological objective of this thesis is to create a learning environment "to investigate the possibilities for educational improvement by bringing about new forms of learning in order to study them" (Kelly, 2006, p. 108) with the aim of linking theory and practice in situ to provide evidence that has the potential to inform teaching practices, curriculum development, and knowledge of student learning (McKenney, Nieveen & Akker, 2006).

Educational design studies are "a series of approaches, with the intent of producing new theories, artifacts, and practices that account for and potentially impact learning and teaching in naturalistic settings" (Barab & Squire, 2004, p. 2). Educational design studies are cyclical in nature and provide the opportunity to develop theories in conjunction with the implementation of instructional sequences. They are based on design-analysis-redesign cycles that move toward an understanding of learning and activity or artefact improvement. Importantly, they are theory driven as they test and advance theories through interrogation of educational problems and are particularly suited to the exploration of technology-based learning environments (Seeto & Herrington, 2006). As noted by Seeto and Herrington (2006) when implementing an educational design research project,

> The value of this approach is that it is focussed on designs and processes that respond to the local context; it is grounded in theory and yields knowledge or guidelines that can be shared and used by others to improve educational practice – demonstrating a commitment to theory constructions and explanations while solving real-world problems. (p. 744)

The interventionist nature of educational design research allows the implementation of a research process that is flexible in terms of facilitating the incorporation of lessons learnt throughout the inquiry into the research design, for the development of both theoretical understandings and the promotion of improved educational objectives (Sandoval, 2004). Educational design research promotes analysis of data in the early stages of a study with the intention that insights gleaned are used in the later stages (Bannan-Ritland, 2003). Analysis of data throughout a research project allows for ongoing theory development and refinement of the research process. The evaluative and reflective nature of the cyclic process allows early results from research to guide the development of interventions and learning sequences. It enables the interventions to be responsive not only to the needs of the researcher in terms of shaping a study and gathering vital evidence but also to the needs of practitioners and students to ensure learning outcomes are realised.

A general model of educational design research includes four phases – development of the research questions, selection of data and data collection methods, design of artefacts and processes, and analyses and evaluation. It "is a research approach that is particularly suited to the exploration of significant education problems and technology-based solutions" (Seeto & Herrington, 2006, p. 742). Seeto and Herrington aligned their research methodology with the integrative learning design (ILD) framework developed by Bannan-Ritland (2003). The phases of the ILD framework are (a) Informed Exploration, (b) Enactment, (c) Evaluation: Local Impact, and (d) Evaluation: Broader Impact. The Informed Exploration phase may include activities such as conducting literature reviews, carrying out needs analyses, and determining the form of teaching intervention to be developed. The Enactment phase is characterised by the development, implementation, and refinement of the intervention over a number of cycles. Refinement of theories and suggestions for redesign arise from the Evaluation: Local Impact phase, while dissemination of data, evaluation of the impact, and consideration of the consequences of the intervention for the long term occur in the Evaluation: Broader impact phase. The ILD framework is utilised by large projects that are expected to be delivered over a long period of time. The extended research period provides the opportunity for the implementation and evaluation of interventions to be iterative (Bannan-Ritland, 2003).

This inquiry utilised the ILD framework developed by Bannan-Ritland (2003) as a guide for designing the inquiry. In recognition that the present inquiry was short term as well as to accommodate the context of the inquiry, the titles of the phases of the ILD framework were modified. Consideration was also given to the educational design based research model used by Seeto and Herrington (2006), who implemented a study that explored the development of a web site for teacher education students. Although their study was relatively short term – less than one year – they applied the phases of educational design research successfully. Table 0.1 details the phases of the ILD framework, the Seeto and Herrington model, and their relationship to the phases developed for the present inquiry.

Table 0.1
Phases of Educational Design Research in Relation to the Present Inquiry

	Phase 1	Phase 2	Phase 3	Phase 4
ILD framework (Bannan-Ritland, 2003)	Informed Exploration	Enactment	Evaluation: Local Impact	Evaluation: Broader Impact
Seeto and Herrington model (2006)	Analysis of practical problems by researchers and practitioners	Development of solutions with a theoretical framework	Evaluation and testing of solutions in practice	Documentation and reflection to produce "design principles"
Reasoning about Covariation with *TinkerPlots* (Present inquiry)	**Analysis of practical problems**	**Development of solutions with a theoretical framework**	**Evaluation of solutions**	**Application of solutions and reflection on implementation**

Educational design research has been used in many studies that interrogated and informed the design of technological learning environments (Reeves, 2006). In the case of Seeto and Herrington (2006), educational design research based on the ILD framework guided the development of a web site for accessing online learning. The purpose of their study was to create a collaborative research environment where software designers worked with education researchers to develop a set of design principles for the development of web pages that facilitated the delivery of teacher education courses.

Although the present inquiry utilises the precepts of educational design research to investigate the way in which students use the software environment to conduct data analysis, it is not the intention of the inquiry to make a contribution to the design principles of

TinkerPlots. The intention is to take advantage of the iterative nature of educational design research to use the outcomes from each stage of the inquiry to inform the next stage of the inquiry. This strategy aligns with and is the enactment of Phase 4 of the ILD framework for this inquiry.

Inquiry Design

The inquiry design utilised a sequential exploratory strategy (Creswell, 2003) to explore the phenomenon of students' use of *TinkerPlots* to develop an understanding of covariation. The strategy facilitated the gathering of evidence to inform the emerging theoretical frameworks and assessment instruments as they were developed throughout the research process. The four phases of educational design research (Bannan-Ritland, 2003; Seeto & Herrrington, 2006) outlined in Table 0.1 provided a structure for the design of the present inquiry. The phases were used to construct a research process for the inquiry that evolved through seven stages, with each stage utilising the phases of the ILD framework (Bannan-Ritland) relevant to that stage. As each stage was developed, implemented and evaluated, a structure was established where the progress of the inquiry was dependent on the relationship between the outcomes of the stages and their application in the proceeding inquiry phases. The phases of the inquiry were derived from the integration of the ILD framework (Bannan-Ritland) and educational design principles adopted by Seeto and Herrington. To accommodate the activities necessary to fulfil the inquiry's objectives, seven stages of inquiry were utilised. The overall inquiry drew together the theoretical underpinnings of learning statistical concepts with the practical application of EDA software in a learning intervention.

Central to the inquiry is the *Model of Learning Behaviour in EDA Graphing Environments* developed in Stage 1 in response to Research Question 1. The *Model of Learning Behaviour* was developed from an interrogation of the research literature on student learning about graph creation and graph interpretation. It provided a focal point for the inquiry and is utilised in Stages 2, 3, and 4 of the inquiry. Its development is also revisited in Stage 6.

In recognition of the major contribution that the contexts of the technological learning environments, student prior knowledge, and instructional design have on student learning

(Angeli & Valanides, 2009; Moore, 1997), separate stages of the inquiry address the issues associated with these three main contributing factors. In Stage 2 the *TinkerPlots* learning environment was explored, in Stage 3 a Student Survey was developed and implemented to determine the students' prior knowledge of graphs, graph interpretation and covariation; and in Stage 4 the learning intervention was developed and implemented. In Stage 5 the results for the three research question were interrogated and in Stage 6 the inquiry design and the *Model of Learning Behaviour in EDA Graphing Environments* were revisited as part of an evaluation of the inquiry process followed by the discussion of the implications and limitations of the inquiry, and recommendations for future research.

The seven stages of the inquiry are outlined in Figure 0.2. The figure maps chronologically the inquiry process through the four phases of the inquiry noted in Table 0.1 to each of the seven stages of the inquiry. The alpha-numerical codes for each phase of the inquiry are used throughout the thesis to signpost which part of the inquiry is being addressed in that stage of the thesis. A graphic is provided at the beginning of each stage to indicate the phases included in that stage of the thesis.

Inquiry Phases / Inquiry Stages	**Analysis of practical problems**	**Development of solutions with a theoretical framework**	**Evaluation of solutions**	**Application of solutions and reflection on implementation**
Stage 0 Inquiry Commencement	**L0.** Literature reviewed on educational design based research	**D0.** Development of Inquiry Design	**E0.** Inquiry Design discussed in Stage 6	**A0.** Guides Inquiry Implementation and Thesis Structure
Stage 1 Development of a Model of Learning Behaviour - Research Question 1	**L1.** Literature reviewed on graph creation and interpretation, using technology, EDA, and models of graphing	**D1.** Development of *Model of Learning Behaviour in EDA Graphing Environments*	**E1.** Evaluation of *Model of Learning Behaviour in EDA Graphing Environments* conducted in Stage 6	**A1.** *Model of Learning Behaviour in EDA Graphing Environments* applied in Stages 2, 3, & 4
Stage 2 Evaluation of *TinkerPlots* – Research Question 2	**L2.** Literature reviewed on evaluating EDA software packages	**D2.** Development of *Criteria for Evaluating EDA Software Environments*	**E2.** Evaluation of *TinkerPlots* using *Criteria for Evaluating EDA Software Environments*	**A2.** Informs other inquiry activities – development of Student Survey and Learning Sequence
Stage 3 Establishment of Student Prior Learning- Research Questions 2 & 3	**L3.** Literature reviewed on assessment instruments for evaluating student learning of covariation and graphing	**D3.** Development of Student Survey to determine student prior learning in graph creation and graph interpretation.	**E3.** Trial and evaluation of Student Survey (n=71)	**A3.** Selection of Participants (n=12) for Stage 4.
Stage 4 Sequence of Learning and Outcomes – Research Questions 2 & 3	**L4.** Literature reviewed on student understanding of covariation and graphing	**D4.** Development of Learning Sequence - Covariation	**E4a.** Implementation of Learning Sequence (n=12) **E4b.**Administration of Student Interviews (n=12)	**A4.** Analysis of Student Interviews (n=12)
Stage 5 The Findings – Research Question 1, 2 & 3	**L5.** Literature revisited and used to support findings	**D5.** Results of Research Questions	**E5.** Discussion of the Research Questions.	**A5.** Recommendations for future research
Stage 6 Inquiry Conclusion	**L6.** Literature revisited and used to support inquiry implications	**D6.** Discussion of Inquiry Implications	**E6.** Evaluation of Inquiry Design using NRC (2002) principles of scientific inquiry. (E0 & E1))	**A6.** Recommendations for future research

Figure 0.2. Stages and phases of the inquiry.

Inquiry Stages

The inquiry design has seven stages, which are outlined and mapped to the ILD phases (Table 0.1) in Figure 0.2. A brief overview of the purpose of each stage is provided in this section. More detailed descriptions about the development and enactment of each of the stages are presented as each stage of the inquiry is reported in this thesis. Providing the detail about each stage of the inquiry as it arises in the thesis addresses the call from Collins, Joseph, and Bielaczyc (2004) to characterise the elements of an inquiry design and state the reasons for including the elements in the inquiry process.

Stage 0 – Inquiry Commencement. The purpose of this stage of the inquiry was to design the inquiry process. To design the inquiry it was necessary first to develop the aim and objectives of inquiry, determine the research questions, select the overarching theoretical perspective that guide the inquiry, and explore the broad methodological issues that best suited the aim of the inquiry.

Stage 1 – Development of Model of Learning Behaviour. The purpose of this stage of the inquiry was to develop a theoretical framework, *Model of Learning Behaviour in EDA Graphing Environments*, which exemplified the critical behaviours of working in EDA graphing environments from an interrogation of the literature on student learning about graphing and development of data analysis skills. The model was used repeatedly throughout other stages of the inquiry to inform the research process, evaluate research instruments, design criteria for the evaluation of *TinkerPlots*, and analyse student interviews. The *Model of Learning Behaviour* was developed in response to Research Question 1 and is revisited in Stage 6 to determine in what ways it contributed to the inquiry meeting its objectives.

Stage 2 – Evaluation of TinkerPlots. This stage was used to establish a clear understanding of the features of the software package and the different graph types it produced in order to answer Research Question 2. To do this it was necessary to develop criteria for evaluating *TinkerPlots* and then apply them to evaluate the potential for *TinkerPlots* to be used as a learning tool. The *Model of Learning Behaviour in EDA Graphing Environments* developed in Stage 1, in conjunction with the literature on evaluating software and technological learning environments, was used to establish the criteria. The literature on previous research about *TinkerPlots*, its application as a learning tool, and its application as a teaching tool were also reviewed in this stage.

Stage 3 – Establishment of Prior Learning. The purpose of this stage of the inquiry was to establish the prior learning of students in relation to their understanding of graphs, graph-sense-making, and covariation. In order to do this, an assessment instrument to be used as a student survey was developed. Research validated assessment items from previous research that evaluated students' development of statistical and graphing concepts were used to construct the student survey. The results from the administration of the student survey informed the design of the sequence of learning experiences developed and implemented in Stage 4 – Sequence of Learning. They also informed the selection of the participants for Stage 4.

Stage 4 – Sequence of Learning and Outcomes. The purpose of this stage of the inquiry was to develop and implement a sequence of learning experiences that would provide the opportunity for novice learners to use *TinkerPlots* to develop an understanding of covariation. As part of the implementation of the sequence of learning experiences, the final session was used to administer an interview protocol to gather evidence of the students' understanding of covariation and determine the way in which they interacted with *TinkerPlots* to create graphs and interpret data. Student profiles that characterised their statistical thinking and reasoning according to the dimensions of the *Model of Learning Behaviour in EDA Graphing Environments* were developed for the students who participated in this stage of the inquiry. The student profiles contributed to answering Research Question 1 in Stage 5.

Stage 5 – The Findings. In Stage 5 the student profiles built from the results in Stage 4 were analysed and used to answer Research Question 1. The student profiles were then analysed another two times to answer Research Questions 2 and 3, respectively.

Stage 6 – Inquiry Conclusion. This stage revisited the research design adopted for the inquiry and the *Model of Learning Behaviour in EDA Graphing Environments* developed in Stage 1 as part of an evaluation of the inquiry design using the principles of scientific inquiry developed by the NRC (2002). The implications and limitations of the inquiry were developed in this stage of the inquiry as were recommendations for future research.

Inquiry Overview

Figure 0.3 maps the overall inquiry process. It provides a representation of the relationship between the phases and stages of the inquiry and indicates which research question is the main focus of each stage. The main connections are indicated by the bold lines and arrows. It is important to note that there are many other more subtle connections, some of which are not indicated on the diagram whereas others are represented by dashed lines. Although not directly responsible for the outcome of a preceding or following stage, information from a particular phase may impact on other stages or phases of the inquiry. For example, evaluation of the students' prior learning in Stage 3 was used to inform the development of the sequence of learning experiences in Stage 4. The map also places the *Model of Learning Behaviour in EDA Graphing Environments* developed in Stage 1 in the centre and indicates its relationship with other stages of the inquiry.

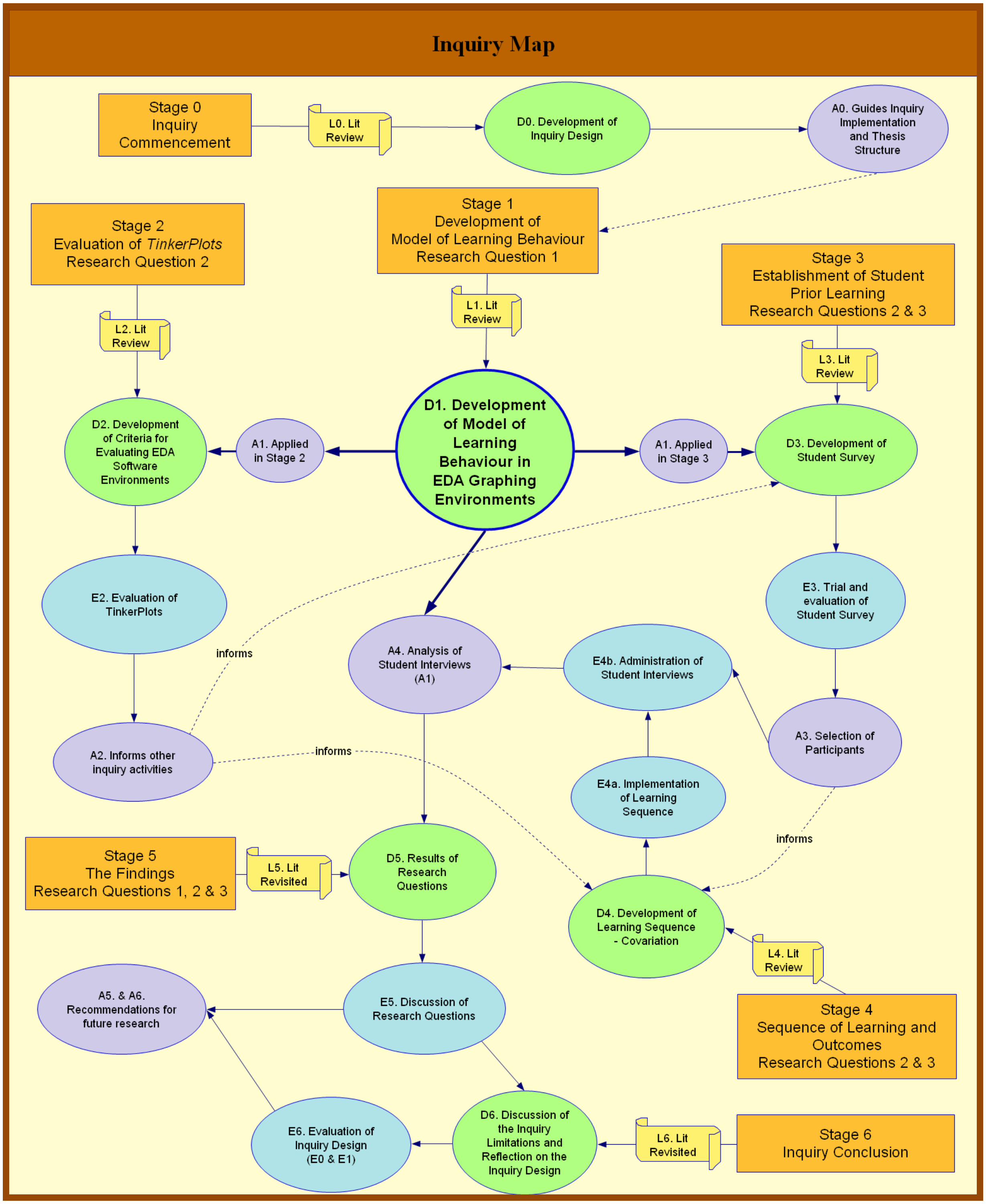


Figure 0.3. Inquiry map.

Sampling Design

The sampling design for this inquiry was based on four criteria: applicability to the inquiry design and ability to meet the inquiry's aim and objectives, ability to provide the data necessary for analyses, capacity to reflect the real situation, and feasibility in terms of available resources (Wiersma, 1995). The overall aim of the sampling design was to accommodate the research problem adequately. Sampling carefully and thoughtfully was important as the sample determines the conclusions that can be made (Patton, 2002). To this end, a purposeful sampling strategy was implemented (Miles & Huberman, 1994). Purposeful sampling was selected to provide information-rich cases to be studied in depth. The cases selected to study in this inquiry were considered the ones from which the most could be learned.

It is also acknowledged that the selection of participants in this inquiry involved an element of convenience sampling (Creswell, 2005). As this inquiry was associated with the MARBLE project, the participants were selected from schools in the project. Walford (2001) suggests that although convenience should not be the sole reason for selecting a sample "it is understandable that academics and research students should include convenience in their consideration of which sites to approach to try and gain access" (p. 14). Even though the participants were a sample of convenience, they were selected on their ability to provide a typical sample that displayed variation of achievement (Patton, 2002).

Internal Validity

Creswell (2003) asserts that the internal validity of a study is strengthened by the rigor of the inquiry design. In this inquiry, internal validity is enhanced by utilising multiple features of internal validity. These features include the development of evidence based theoretical frameworks (Akker et al., 2006; Creswell, 2003; Seeto & Herrington, 2006), the integration of the work of other researchers (Merriam, 1998; Watson, 2002), application of established models of cognitive development (Shaughnessy, 2007), validation of theoretical frameworks and evaluation instruments (Creswell, 2003; Lincoln & Guba; 2000), and possibly most importantly, utilisation of peer review processes (Miles & Huberman, 1994; Merriam; 1998). These notions align with the ideals of the educational design research methodology adopted for this inquiry (Akker et al., 2006) and similarly inform the integration of ILD and educational design research (Table 0.1) strategies that underpin the development of the inquiry design.

Qualitative education design research by its very nature provides rich descriptive data that reflect the real experiences of the participants (Akker et al., 2006). When implemented in full, education design research promotes theory building through the application of and development of theoretical frameworks as a matter of course. Throughout the stages of this inquiry the literature on previous research is used to inform the inquiry process. Where possible established theoretical frameworks from the research literature are utilised and where necessary established frameworks are incorporated into new frameworks more relevant to the present inquiry. The researcher has also engaged in peer review of the inquiry through publication of various components of the inquiry in research conference proceedings. These publications are listed in Stage 6: Evaluation of the inquiry. The strategies for strengthening the internal validity adopted for this inquiry are also revisited in Stage 6 of the thesis.

Inquiry Significance

This inquiry recognises that the statistical thinking and reasoning used when developing an understanding of covariation in the *TinkerPlots* learning environment is important for several reasons. First, understanding the development of understanding of covariation through exploring the relationship between two attributes, particularly when using *TinkerPlots*, can inform teachers how to structure learning so that the opportunity for students to progress through all the levels of development is incorporated into learning activities. Second, understanding the influence *TinkerPlots* has on students' understanding of covariation can contribute to teachers' knowledge of how to use *TinkerPlots* purposefully to target particular learning outcomes. Third, understanding the stage at which students can develop an appreciation of covariation and how it is conceptualised by students can inform future curriculum development and contribute to improved classroom practices.

The importance of this inquiry extends beyond the implications for classroom practice. It makes a significant contribution to calls from the literature for future research directions to include "(a) research on conceptual issues in statistics, (b) research on teaching issues in statistics, and (c) some methodological issues for research in statistics" (Shaughnessy, 2007, p. 999). Ben-Zvi (2004) identified that there is the need to conduct research that may inform pedagogical practice about the ways in which students' reasoning about data analysis develops. In relation to the concept of covariation, Ross and Cousins (1993), Cobb, McClain et al. (2003), and Moritz (2004) are indicative of the continuing call from the literature to research and interrogate further students' development of understanding of covariation. Shaughnessy (2007) also identified that "too little research has been conducted on the effects of statistical software packages on students' conceptual growth and thinking statistics" (p. 1000). As part of the research agenda, assessment instruments and materials that evaluate statistical thinking and reasoning in technology environments need to be developed (Garfield & Ben-Zvi, 2004; Shaughnessy, 2007). This inquiry attends to all these major research issues. It investigates students' development of covariation directly within the technological environment afforded by *TinkerPlots*; it develops, applies, and evaluates assessment instruments; it develops theoretical frameworks from reviews of the literature; it applies the theoretical frameworks throughout the inquiry; and it evaluates the inquiry design to determine the soundness of the methodological approach taken.

Ethical Approval

Ethical approval for the research was gained from the *Southern Tasmania Social Sciences Human Research Ethics Committee* at the University of Tasmania in 2006 – Ethics reference number H8778. The committee adheres to the guidelines outlined in the *National Statement on Ethical Conduct in Human Research* (National Health and Medical Research Council, 2007). The research also had permission and approval from the Department of Education, Tasmania, and satisfied department criteria for *Conducting Research in Tasmanian Government Schools and Colleges* (2006).

Student participation in the inquiry was voluntary and all the protocols necessary to gain informed consent were adhered to (Miriam, Edwards, & Alldred, 2001). The students' parents/guardians were given an informed option, via a withdrawal of participation form. Included on the form was a request that the researcher be able to use the students' first names in all publications that arose from the inquiry. Although it would be possible to identify the students from their first names the associated risks from being identified were considered minimal. The benefits, however, were considerable. These included: contributing to the authenticity of the inquiry, allowing the researcher to make meaningful connections between the data and the participants, negating the need to obscure students' names in audio recordings, and most importantly, honouring the contribution each of the students made to the inquiry. The work students completed was only used by the researcher and did not contribute to the students' end of year results.

Structure of the Thesis

This PhD thesis is unique in that it uses the structure of the inquiry design as a structure for this thesis document. It is set out to follow the design process through Stages 0-6 as chapters of the thesis. The decision for the structure of the thesis to follow the inquiry design was made so that the thesis reflected the evolving exploratory nature of the inquiry in chronological order as mapped in Figure 0.2. Setting out the thesis chronologically allowed the thesis to demonstrate clearly how each stage of the inquiry was developed from the literature and how the literature was used to inform each stage as the inquiry progressed. It allowed the factors that influenced the enactment of each stage to be made explicit when they were relevant. It also made it possible to document the way in which outcomes from various stages were integrated in the proceeding stages.

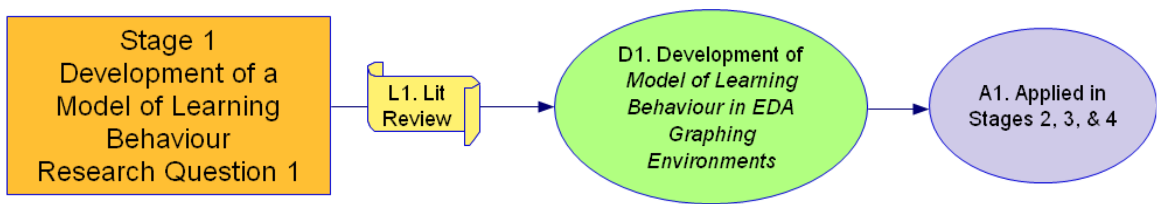
All of the conventions for a regular thesis such as literature review, methodology, and results are included in the thesis; however, they are not presented as individual chapters. The thesis is organised so that literature reviews and methodological considerations for each of the stages are presented in the stage for which they are associated.

The structure of this thesis is based on the inquiry map detailed in Figure 0.2. It is divided into seven sections with each section of the thesis for the first six stages of the inquiry opening with a literature review, as noted in the first phase of each stage in Figure 0.2. This is followed by information about the implementation of the other phases for that stage together with information about the methodological considerations relevant to that stage.

Stage 0 – Inquiry Commencement sets the scene for the inquiry and includes the broad methodology that underpins the inquiry and explains how the methodology fits the purpose of the inquiry. The aim and objectives, the research questions, the origins of the inquiry, the inquiry design, the sampling design, the significance of the inquiry, and a consideration of the ethical issues associated with the inquiry are also included. Following Stage 0, the next four stages detail the enactment of the inquiry. Stage 5 – The Findings presents the results for the three research questions, with the evaluation of the inquiry design and the implications and limitations of the inquiry presented in Stages 6 – Inquiry Conclusion. Stage 6 links back to the other stages of the inquiry through an evaluation of the inquiry design. The principles of scientific inquiry developed by the NRC (2002) are used to critique the inquiry. Linking the Stage 6 – Inquiry Conclusion back to other stages of the inquiry completes a cycle of inquiry that fulfils the requirements for conducting scientifically-based educational design research (Phillips, 2006).

Stage 1

Development of a Model of Learning Behaviour



The purpose of Stage 1 of the inquiry is to develop the *Model of Learning Behaviour in EDA Graphing Environments*. The *Model of Learning Behaviour* characterises the way students work when using EDA strategies to construct and interpret graphs. The development of the model is informed by models of graphing sourced from statistics education research literature and takes into consideration research on student understanding of covariation and the increasing propensity to use technology to create and interpret graphs. Constructing the *Model of Learning Behaviour* from various statistics education research sources increases the internal validity of this inquiry (Akker et al., 2006; Creswell, 2003; Richardson, 2003).

In this stage of the inquiry the background to graphs and graphing is presented, followed by a summary of the historical developments in graphing. The historical developments map an increasing complexity of graph types and describe the changing use and purpose of graphs since the 1600s. This leads to the way in which graphs combined with new technologies contribute to the application of EDA strategies in modern times. The historical developments in graphing foreshadow the way in which the introduction of different graphs types is sequenced in curriculum documents, which is presented in Stage 4 – Sequence of Learning and Outcomes. A summary of EDA strategies is presented in the next section, which is followed by an exploration of the research literature on covariation. The development of the *Model of Learning Behaviour in EDA Graphing Environments* contributes to answering Research Question 1. The *Model of Learning Behaviour* is presented after exploring the models of graphing sourced from statistics education research.

Graphs and Graphing

Tufte's three books, *The Visual Display of Quantitative Information* (1983), *Envisioning Information* (1990), and *Visual Explanations: Images and Quantities, Evidence and Narrative* (1997), provide a commentary on the different types of visual images developed since the 1600s. They show a variety of representations that are often complex and multi-dimensional, requiring well developed analytical and critical thinking skills to interpret. They also illustrate the way in which powerful imagery has been used to communicate with audiences.

A recurring theme throughout Tufte's three books is the need for graphing practices to be committed to finding, telling, and showing the truth about the data. He emphasises the importance of visual representations in being able to sum up and convey information as well as stimulate ideas. He also recognises that the construction of data displays is as much about reasoning about statistical evidence as it is about displaying statistical information effectively. The effectiveness, however, depends heavily on the ability of the user to understand the imagery used and the conventions applied.

In 1805, Playfair recognised the potential complexity of graphs and cautioned:

> Opposite to each Chart are descriptions and explanations. The reader will find, five minutes attention to the principle on which they are constructed, a saving of much labour and time; but, without that trifling attention, he may as well look at a blank sheet of paper as at one of the Charts. (p. xvi)

With this in mind, it is appropriate to develop an understanding of the way graphs are structured, in order to appreciate the way in which they communicate information. Although there are many types of graphs, they can be broken down into the same basic-level constituents (Kosslyn, 1989). Kosslyn suggests a schema for the analysis of graphs that can be used to communicate information clearly and concisely. The elements include the “background,” the “framework,” the “specifier,” and the “labels.” Figure 1.1 illustrates the basic-level constituents of a typical graph.

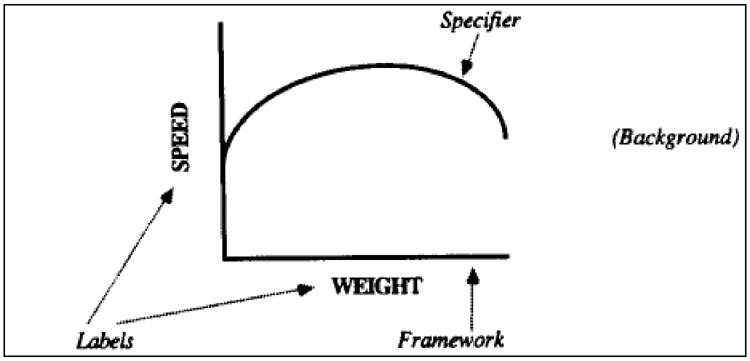


Figure 1.1. The basic-level constituent parts of a graph (Reproduced from Kosslyn, 1989, p. 188).

According to Kosslyn (1989), the *Background* is the pattern over which the other component parts of a graph are presented. In most instances the background is blank as it is not necessary to include a pattern or picture. The pattern of a background such as a photograph can assist in conveying the information of the graph, but when too detailed may interfere with the ability to read the graph. The *Framework* extends to the edges of the graph and its function is to organise the graph as a meaningful whole. Some graphs may have an inner framework that is nested within the outer framework. An inner framework is a structure (e.g., a grid) that maps points on the outer framework to other parts of the display. The *Specifier* conveys specific information about the framework by mapping parts of the framework to other parts of the framework. The specifier may be a point, line, or bar and is often based on a pair of values. The *Labels* of a graph are an interpretation of a line or region. They may be letters, words, or pictures that provide information about the framework or the specifier.

To analyse graphs it is necessary to understand the interrelated connections among the constituents of a graph as the connections foster the interpretation of graphs on three levels (Kosslyn, 1989). First the individual elements and their organisation can be described. Second, understanding of the display can be determined by looking at the relations among the elements of the graph. Third, the analysis can extend to interpretation of the symbols and lines that goes beyond the literal reading of the information. The interpretation of the meaning of the graph is conveyed by the way in which the information is organised and the connections that can be made among the elements.

A graph also has many characteristics that can be used when thinking about the data. The characteristics, such as the mode, scale of an axis, or the variation in the spread of the data can be extracted directly from the graph (Kosslyn, 1989; Roth, Pozzer-Ardenghi, & Han, 2005) or from calculations performed by graphing software, such as *TinkerPlots*. In these instances, the software is a tool that is used to gain access to the characteristics of the graph. Another factor that influences conclusions that are made about graphs is the context of the data. An understanding of the context and the nature of the variables of interest may be gained from personal experiences of the context, from information about the data, or from the details embedded in the scales and frameworks of the graphs (Kosslyn, 1989; Roth et al., 2005; Fitzallen & Watson, 2011). Collectively, the elements of graphs are resources that provide a link between the two-dimensional representations and the real world measurement situations.

Historical Developments in Graphing

Quantitative graphics have been used since ancient times to represent information. They have their origins in mapping but have evolved over time to become important data analysis tools. The Egyptians used a coordinate system to show the location of points in real space c3200BC and used graphical representations to show the area of shapes, including squares, trapeziums, triangles and circles c1500BC (Beniger & Robyn, 1978). From that time until the 1600s, for the most part, graphical representations were used for mapping and recording the orbits of planets over time. It was not until the late 18th and early 19th centuries that the use of graphs and charts for data displays became accepted practice (Fienberg, 1979). Since that time the development of graphic methods has depended on advances in technology, data collection and statistical theory (Friendly, 2007).

In a brief overview of the history of quantitative graphics in statistics, Beniger and Robyn (1978) describe four stages that correspond to successive historical periods that began in the early 1600s. During these periods developments were made that came about as a result of major graphical problems that preoccupied scientists and data analysts at the time. The four stages are: spatial organization for data analysis, discrete quantitative comparisons, continuous distributions and multivariate distributions and correlation. Beniger and Robyn also include an additional section that introduces the innovations developed in the 20th century. Collectively, the four stages and the 20th century innovations period describe progressive developments that have influenced the way in which data are represented and analysed today. Although the developments were introduced in successive historical periods, the ideas introduced in earlier periods were not superseded by successive developments. Instead, elements of each stage are carried over and incorporated into the next.

Spatial Organisation for Data Analysis

Technological innovations in the form of automatic measuring devices invented in the 17th and 18th centuries made it possible to collect and record large sets of data. Measuring devices such as the air and water thermometer, weather-clock, pendulum clock, and mercury thermometer were used to develop scientific instruments that were capable of making multiple measurements. As a result, new ways of organising and analysing data were necessary to handle the large collections of data. Generally, the automatic recording devices produced moving line graphs that represented data collected over a period of time using a coordinate system. At the time, it was common for the data to be translated from this graphical form into tabular form for analysis. The graphical form was considered a means of recording data and the potential for it to be used for analysis purposes did not occur until later. The coordinate system of Cartesian plots reintroduced in mathematics by Descartes in 1637 did not become an important tool for data analysis until the 1830s (Beniger & Robyn, 1978).

Discrete Quantitative Comparisons

The combination of visual imagery and statistical data to create graphical representations of information other than scientific data was instigated by Playfair in 1786 (Tufte, 1983). Playfair replaced tables of numbers with visual representations, creating the opportunity to use pictures and graphics to reason about quantitative information. The graphical representations he produced were very complex, often displaying multivariate data on the same graphic. Although others preceded Playfair in using graphics to display data, he extended their use to the areas of economics and finance, making the use of statistical graphics popular for general interest information. (Wainer & Velleman, 2001).

Continuous Distributions

Furthering the development of demographic statistics, in 1821, J. B. J. Fourier was the first to apply graphical analysis to population statistics (Beniger & Robyn, 1978). Fourier developed the cumulative frequency distribution in order to show the number of inhabitants of Paris who were of a given age or over per 10,000 people. He began with a bar chart representing the age groupings, and then placed the bars one atop each other for a particular age range. This was repeated for other age ranges at regular intervals to produce a graph. Fourier analysed the cumulative frequency distribution to determine geometrically "the mean duration and probable duration of life, the mean age of a population and the stability of life" (Funkhouser, 1937, p. 296). The cumulative frequency distribution was named an "ogive" by Galton in 1875 (Beniger & Robyn, 1978). Cumulative frequency is used to construct a box-and-whisker plot, a semigraphical innovation designed by Tukey in 1977.

Another innovation developed from the bar chart was the histogram. In 1833, A. M. Guerry produced histograms by arranging ordered categories for continuous data (Beniger & Robyn, 1978). He used columns of equal width to represent the frequency for each class at equal intervals of the data. A frequency polygon was obtained by joining the midpoints of the class intervals. The line formed in this way begins and ends on the horizontal axis, resulting in an irregular polygon. The frequency curve was the smoothed curve derived from the frequency polygon (Funkhouser, 1937). The word "histogram" was first used in 1895 by Karl Pearson in *Contributions to the Mathematical Theory of Evolution – II*. Pearson recorded data collected in a histogram and compared the graph with a corresponding theoretical skewed frequency curve. Adolphe Quetelet also furthered the development of the graphics of continuous distributions by applying the theory of probabilities to graphical methods (Funkhouser, 1937). In 1846, Quetelet recorded the results of sampling goods from urns as symmetrical histograms, and then showed the limiting "curve of possibility." This was developed further and later called the normal curve (Friendly, 2007).

Multivariate Distributions and Correlation

During the mid-19th century the data related to vital statistics and demographics became more complex and involved interrelationships among more than two variables. Contour maps and stereograms were developed at that time to accommodate the increase in the complexity of data as as they provided two dimensional representations of multivariate distributions and correlations.

The simplest case of multivariate distributions is the bivariate distribution. Bivariate data provide information about two variables that are not necessarily dependent on each other. A scatterplot is a graphical technique used to display paired measurements of two quantitative variables. Scatterplots are useful for exploring and identifying clusters of points and outliers in a distribution of bivariate data and assist in the identification of the relationship and dependence between the variables, and variation from those (Cleveland, 1993). Hence, scatterplots are used to display covariation.

Advances made in the development of powerful computer technologies have led to further developments in the field of multivariate distributions. This has extended the ability to apply statistical tests to multivariate distributions to determine covariance or correlation (Johnson & Kotz, 1972) and construct interactive and animated three-dimensional graphical representations (Friendly, 2007). Correlation is used to identify statistical relationships between two or more variables and, where appropriate, to seek causal explanations (Moritz, 2004).

20th Century and Future Developments

After a period of time in the early 1900s, when formal statistical analysis of data was favoured over graphical analysis, the importance of the visualisation of data for graphical analysis regained prominence (Friendly, 2009). This can be attributed to three main developments. First Tukey (1977) introduced the concept of exploratory data analysis (EDA), second Bertin (1981) developed a theory of graphics, and third technological innovations allowed for the computer processing of statistical data (Friendly, 2009).

EDA is both a philosophical stance and a collection of data analysis tools and techniques. It is a way of searching data to reveal patterns that might otherwise go unnoticed

(Smith & Prentice, 1993). The emphasis is upon using visual displays to reveal vital information about the data being examined. The aim is to use alternative techniques to investigate the same set of data (Hartwig & Dearing, 1979).

The second development of the 20th century is attributed to the work of J. Bertin. In 1967 he published his theory of information visualisation in French, *Semiologie Graphique,* which was translated into English in 1983. Bertin's interest in this area started when he identified that graphical representations produced in scientific publications were not understood. The theory he developed focused on the interpretation of the visual and perceptual elements of graphics. His work made a distinction between how the qualitative and quantitative elements imparted meaning (Card, Mackinlay, & Shneiderman, 1999). Tufte (1983) also developed a theory of data graphics that emphasised the maximisation of the density of information and the minimisation of extraneous information he termed "chart junk."

Advances in computer technologies have had and continue to have the potential to have a significant impact on the analysis of data and the visual displays used to represent the data. Graphing software and associated technologies provide an alternative to hand drawn graphics, embellishment of older graphical types, analysis of large multivariate data sets, and representation of multivariate data sets in two or three dimensions (Beniger & Robyn, 1978). Graphing software such as *TinkerPlots* is evidence of the way in which interactive and digital technologies are intersecting with the theories of graphics and the philosophy of EDA. In the future, the application of computer technologies to data analysis has the potential to produce new and innovative data representations. These may be used to "support teachers' and students' deeper understandings of data by enabling them to create different representations that can lead to different questions and different arguments" (Rubin & Hammerman, 2006, p. 253).

Exploratory Data Analysis

Exploratory data analysis (EDA) is an informal yet robust graphical approach to data analysis that focuses on the appearance of graphs to provide insights about the data rather than making formal inferences from statistical calculations. It is a paradigm that is flexible and allows data analysis to be repetitive, iterative, and creative (Tukey, 1977). The aim is to reveal patterns in data sets, seek connections between data sets, explain and describe variation, identify outliers, get to know the data, and look for associations between variables (Ben-Zvi, 2004; Biehler, 1997).

EDA is commended for its ability to reveal structure in data and facilitate the interpretation of the structure in language that makes direct connections to the context of the data (Diaconis, 1983). Biehler (1997) notes, however, that opponents of EDA warn “people tend to see structure where there is really only chance, and that may mislead science” (p. 5). He also argues that learning EDA ways of thinking can contribute to an understanding of the application of more formal statistical analysis strategies.

Data analysis is not just about getting the right answer to a question (Tukey, 1980). It is also about what influences the answer, what questions are asked, and the way in which they are asked. These notions are in opposition with confirmatory data analysis (CDA). CDA employs principles and procedures that look at a sample and considers what the sample reveals about the larger population. The inference made is then assessed to determine the confidence with which the inference from sample to population is made. Tukey advocates that EDA should be taught along with the techniques of CDA. He states: “We need to teach exploratory as an attitude, as well as some helpful techniques, and we probably need to teach it before confirmatory” (p. 25). Tukey (1980) refers to EDA as:

> It is an attitude, AND
> A flexibility, AND
> Some graph paper (or transparencies, or both).
> The graph paper – and transparencies are there, not as a technique, but rather as a recognition that the picture-examining eye is the best finder we have of the wholly unanticipated. (p. 24)

EDA is based on graphical representations of data with a few added quantitative techniques. The analytical strategies applied preserve the information contained in the data and present it in useful and meaning ways (Smith & Prentice, 1993). The semigraphical designed displays developed by Tukey provide a visual representation of the data as well as statistical information. These include the stem-and-leaf display and the box-and-whisker plot.

The stem-and-leaf display is an alternative to tallying values into frequency distributions. It organises a batch of numbers graphically and directs attention to various features of the data. It displays a distribution of a variable with numbers themselves. In overall appearance the display resembles a horizontal histogram with interval widths often displayed on a contracted number line (Emerson & Hoaglin, 1983). An example is provided in Figure 1.2a. The distribution of two data sets can be compared when displayed as a back-to-back stem-and-leaf display (Figure 1.2b). As well as being useful representation for analysing data, the stem-and-leaf display has been used successfully to assist in the development of students' understanding of place value (Dunkels, 1988).

pulse rate

6	8 8 8 9
7	0 1 1 4 6 6 8
8	2 6 8 8
9	0 6
10	4
11	0

Figure 1.2a. Example of stem-and-leaf display (Australian Curriculum, Assessment and Reporting Authority [ACARA], 2011, p. 39).

pulse rate

before		after
9 8 8 8	6	
8 6 6 4 1 1 0	7	
8 8 6 2	8	6 7 8 8
6 0	9	0 2 2 4 5 8 9 9
4	10	0 4 4
0	11	8
	12	4 4
	13	
	14	6

Figure 1.2b. Example of back-to-back stem-and-leaf display (ACARA, 2011, p. 72).

The box-and-whisker plot is another useful display that Tukey developed for comparing multiple data sets (Emerson & Strenio, 1983; Fienberg, 1979). The box-and-whisker plot can be determined from cumulative frequency and is directly related to the ogive representation described by Galton (1875) (cited in Beniger & Robyn, 1978). The box-and-whisker plot summarises a data set, locates the median, displays the spread and skewness of the data, as well as identifies the outliers, but does not display the overall distribution of the data (Friel, Curcio, & Bright; 2001). Direct comparison of several data sets or subsets of data can be conducted by displaying box-and-whisker plots in parallel. The graph in Figure 1.3, for example, shows the shift in the median body weight of students and the changes in the spread of the data in the inter-quartile range across the four grades.

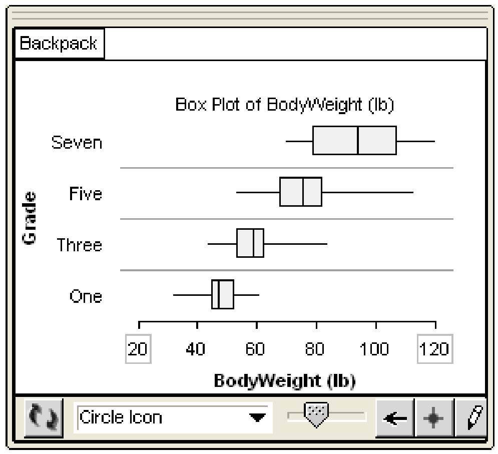


Figure 1.3. Parallel box-and-whisker plots produced using a data set in *TinkerPlots*.

A statistical strategy employed by supporters of EDA, is the transformation of data (Emerson & Stoto, 1983). This involves making changes to data displays such as, changing the scale of a graph, splitting the data into subsets, or eliminating outliers. This may result in the identification of trends or facilitate comparison of data sets. Other useful EDA strategies, such as the examination of residuals, analysis of two-way tables by medians, and construction of linear regression lines to summarise *x-y* data have been developed (Velleman & Hoaglin, 1981), but are not expanded upon in this thesis as they are not relevant to the context of this inquiry.

When first introduced, EDA strategies and techniques added a new dimension to the way that data were analysed. After using the strategies for 10 years and being “continually impressed by how easily they have enabled us, our colleagues, and our students to uncover features concealed among masses of numbers” (p. xiii), Velleman and Hoaglin (1981) identified that computer programs that act as exploratory tools would promote the spread of exploratory methods. Velleman and Hoaglin worked with Tukey to develop the programming required to translate the EDA strategies into computer software. Although EDA strategies were developed before personal computers became ubiquitous, Tukey suggested that computer software be developed for EDA. His collaboration with Velleman and Hoaglin (1981) paved the way for the development of innovative software, of which, *TinkerPlots* is an example.

Covariation

Covariation is recognised as an important aspect of statistical reasoning and is used to explore the relationship between two variables. Statistical covariation, as described by Moritz (2004), is "the correspondence of variation of two statistical variables that vary along numerical scales" (p. 228). It is determined by examining graphs of bivariate data that are represented in Cartesian graphs, such as scatterplots. Scatterplots are characterised by data points that correspond to the measures of two variables designated at the same time (O'Keefe, 1997). Each data point on a scatterplot corresponds to one unit of analysis between the two variables (Roth et al., 2005) and the values of the two variables may be said to involve some form of relationship, association, function, dependency, or correspondence (Moritz, 2004; Zieffler & Garfield, 2009).

Scatterplots are an economical way of organising large amounts of information and depicting covariation of two sets of measurements that vary along numerical scales in a two-dimensional space. They provide information about two variables that are not necessarily dependent on each other and show the correspondence of the ordination of each variable (Moritz, 2004). Scatterplots assist in the identification of the relationship between two variables and variation from that relationship (Cleveland, 1993). They are particularly useful for identifying clusters of points and outliers in distributions of bivariate data (Beniger & Robyn, 1978; O'Keefe, 1997).

The overall image of a scatterplot provides one representation that can be used to make sense of the data. It is a picture that is viewed from a global perspective and is often used to identify a trend in the data. A trend is the functional relationship between two sets of corresponding measurements. The process of identifying a trend involves imagining continuity in the data that is not necessarily accessed directly from the data. It may be expressed verbally using the characteristics of a graph to describe and explain the relationship, be represented on a graph by the inclusion of a hand drawn line-of-best fit Figure 1.4a, or be expressed by "curves rendered by means of hand gestures during face-to-face conversations" as indicated in Figure 1.4b (Roth et al., 2005, p. 26).

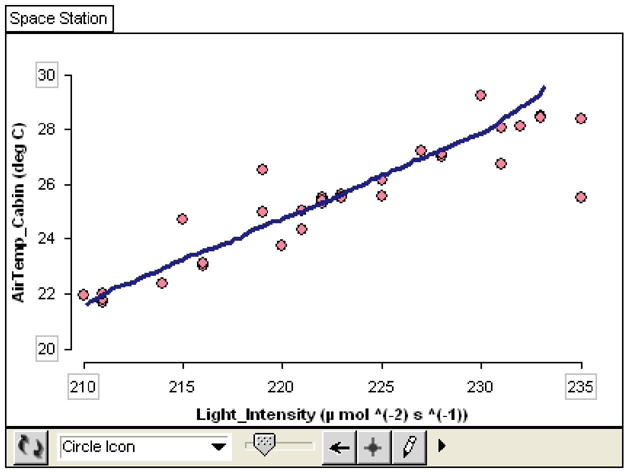


Figure 1.4a. Scatterplot with line-of-best fit inserted using the drawing tool in *TinkerPlots.*

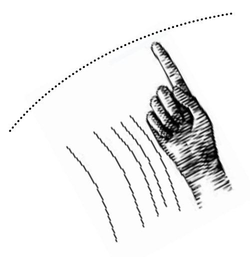


Figure 1.4b. Trend indicated by means of a hand gesture (Reproduced from Roth et al., 2005, p. 26).

The definition of covariation is, however, often included in the definition for association (Batanero, Estepa, Godino, & Green, 1996; Moritz, 2004; Zieffler & Garfield, 2009). Covariation is described by Zieffler & Garfield (2009) as "Reasoning about *association* (or *relationship*) between two variables, also referred to as *covariational reasoning*, or reasoning about *bivariate data*" (p. 7). Similarly, Moritz (2004) suggests:

> Covariation concerns association of variables; that is, correspondence of variation. Reasoning about covariation commonly involves translation processes among raw numerical data, graphical representations, and verbal statements about statistical covariation and causal association." (p. 227)

Moritz then distinguishes statistical covariation from statistical association stating:

> The more general term *statistical association* may refer to associations between two categorical variables, commonly represented in two-way frequency tables, and between one categorical and one interval variable, often formulated as the comparison of group." (p. 228)

Batanero et al. (1996) also describe covariation as a form of association. They note that association is "the analysis of contingency tables, the determination of correlation between quantitative variables, and the comparison of a numerical variable in two or more samples" (p. 151).

In this inquiry, the emphasis for covariation was on the relationship between two numerical attributes, which was interpreted as general trends that show the variation of the two attributes due to the ordination of the values along each axis of a graph with numerical scales (Moritz, 2004). This inquiry did not include formal interpretation of the trend in algebraic form, nor did it involve formal assessment of the correlation between the attributes. Hand drawn lines-of-best fit were used in some instances but the interpretation of those was limited to qualitative descriptions and were often expressed as a hand gesture as shown in Figure 1.4(b). Examples of covariation and possible trends are shown in Figure 1.5a-e. The graphs were created from data sets available in *TinkerPlots*.

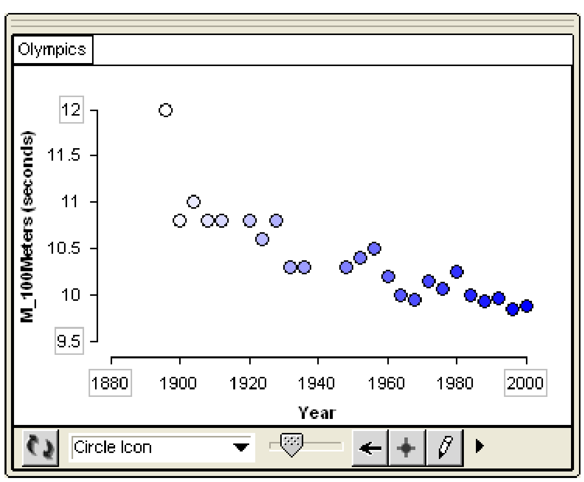


Figure 1.5a. Scatterplot displaying covariation as a downward trend.

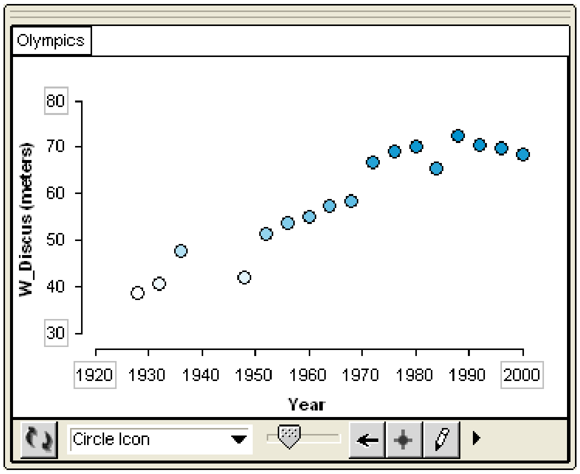


Figure 1.5b. Scatterplot displaying covariation as an upward trend.

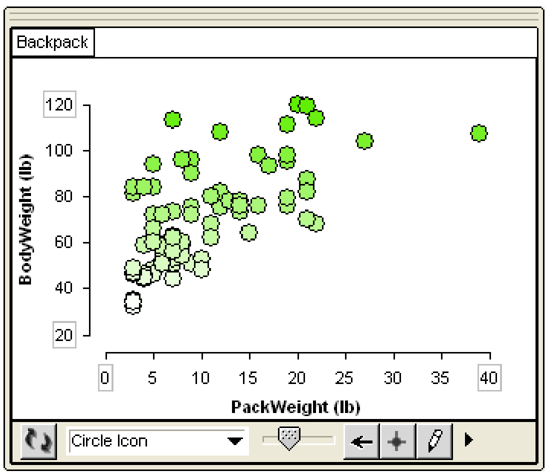


Figure 1.5c. Scatterplot displaying covariation shows a possible upward trend but there is less evidence of covariation than in the graph in Figure 1.5b as the data are more spread out.

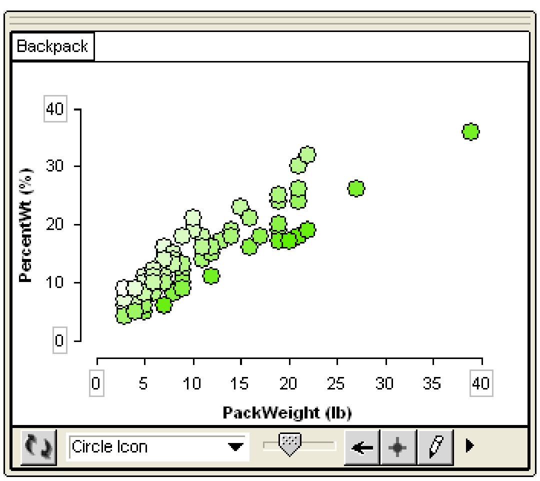


Figure 1.5d. Scatterplot displaying covariation shows less variation on the upward trend than the graph in Figure 1.5c, which indicates stronger evidence of covariation.

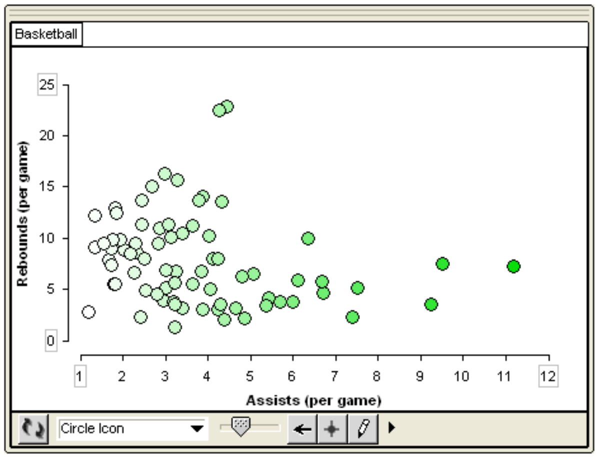


Figure 1.5e. Scatterplot displaying shows non-covariation as the data are spread out broadly and there is no trend evident.

Learning about covariation is possible across a broad spectrum of ages from pre-kinder to the tertiary level. A study of 71 pre-schoolers between the ages of four and six that used images of objects to explore the relationship between two variables determined that young children could reason about covariation (Koerber, Sodian, Thoermer, & Nett, 2005). It showed that when data patterns are simple and do not require complex statistical inference even pre-schoolers can get a grasp on the notion of covariation. The Koerber et al. study also showed that 5-year-olds are able to understand that the lack of pattern between two variables support the conclusion that the variables are not related.

The use of scatterplots to determine a trend in data has also been explored at the middle school level. An activity developed by O'Keefe (1997) was designed to provide a physical model of a scatterplot. It involved putting a scale across a classroom wall to represent an axis of a scatterplot for the height measurements and another scale on a perpendicular wall to represent the other axis of a scatterplot for the arm span measurements. In this situation, the corner of the room represented the origin of a coordinate plane. Students then positioned themselves in the classroom according to their height and arm span measurements. Each student represented an individual data point that constituted the designation of two variables at the same time. The activity gave students the opportunity to explore the trend in the data and identify clusters of data from the way the students were grouped and spread out across the room.

In statistics education research in recent times, reasoning about covariation has had little attention, particularly for younger students. The research is mostly qualitative in nature and involves conducting interviews, teaching experiments, and student surveys. Research into the use of technology by students when working statistically is starting to gain attention but needs to be expanded further (Bakker, 2001; Biehler, 1997; Shaughnessy, 2007).

Exploration of Year 8 students' understanding of covariation using *Minitools* software was the focus of a study conducted by Cobb, McClain et al., (2003). The *Minitools* software is described as route-type software because one graph type is accessed at a time to facilitate student learning for a particular concept, thereby supporting a route to targeted learning outcomes (Bakker, 2002). The *Minitools* software has five established graph types to choose from. The teaching experiment implemented in the Cobb, McClain et al. study followed a sequence of learning experiences that involved: exploring univariate data sets as distributions, developing ways of inscribing bivariate data, and thinking about stacked data as bivariate distributions. The distributions of data within successive stacks of data or slices of distributions were compared rather than searching for patterns and trends from a relational perspective across the full range of the data (Cobb, McClain et al.). These activities were designed to be initial steps towards using bivariate distributions displayed in scatterplots to understand covariation. They found that the students experienced difficulty transitioning from stacked data (univariate plots) to unstacked data (scatterplots). Cobb, McClain et al. suggested making the connections between the data and familiar contexts would support students to focus on regions of scatterplots (slices) before moving towards discussing the global trends in the plots. The part of the investigation that explored the way in which the students used scatterplots to analyse data found that focusing on the density and the shape of data was crucial for the students to interpret scatterplots comprehensively. They also noted that students "typically reduced scatter plots to lines that signified fixed relationships of covariation rather than conjectured relationships about which the data were distributed" (p. 75).

In a study of students in Years 3, 5, 7, and 9, Moritz (2004) looked specifically at speculative data generation, which involved students translating a scatterplot into a verbal statement, and at numerical graph interpretation that involved students reading values and interpolating data. He found that students tended to focus on individual data points rather than look globally for the trend in the data and often only considered one of the variables in isolation. Ben-Zvi and Arcavi (2001) also found that Year 7 students focused on particular points of a graph. They noticed, however, that this did not always constrain the students and "served as a basis upon which the students started to see globally" (p. 62).

A time series graph displaying the winning times in the men's 100 metre foot race during the modern Olympic Games was used in a study that focused on middle school students' development of reasoning about data analysis using EDA strategies (Ben-Zvi, 2004). Of particular interest was the students' shift between local and global observations when making sense of the graph. Ben-Zvi noted that at the beginning of the study:

> Instead of looking at the graph as a way to discern patterns in the data, students' response[s] focused first on the nature and language of the graph as a representation — how it displays discrete data, rather than as a tool to display a generality, a trend. (p. 130)

Ben-Zvi also reported that "this understanding of pointwise information served later on as the basis for developing a global view" (p. 129) of the data. By the end of the study the students were able to move interchangeably between pointwise observations and global considerations.

The importance of building up an understanding of covariation upon an initial understanding of variation and distribution is advocated by Bakker and Gravemeijer (2004), Ben-Zvi (2004), Cobb, McClain et al. (2003), Konold (2002), and Konold and Higgins (2003). Very few studies, however, have examined this notion in depth. A study conducted by Zieffler and Garfield (2009) supported the suggestions from the literature but also found that students' reasoning about bivariate data was closely tied to but not dependent on their reasoning about univariate distribution. Although the Zieffler and Garfield study focused on tertiary students' patterns of development of reasoning about quantitative bivariate data, the results are still relevant to this inquiry because almost all of the tertiary students had not studied statistics at the high school or undergraduate level previously. The students did, however, have established intuitions about covariation that served as a starting point for extending their understanding further.

Like Cobb, McClain et al. (2003), Moritz (2004) considered the context of investigations to be important when exploring students understanding of covariation and based his activities on contexts that were familiar to the students. He found, however, that familiar contexts were useful but sometimes contributed to limiting the students' thinking and reasoning about the data. The students in the Moritz study had difficulty determining the appropriate trend in a graph when the graph displayed a negative covariation that was counter-intuitive. These students relied on their experience with the context to make conclusions rather than letting the data tell the story. In a study that focused on students' use of contingency tables to determine association between variables, Batanero, Estepa, and Godino (1997) also found that students had problems reasoning about relationships that were negative.

The following main themes emerge from research that focuses on middle years students' understanding of covariation. The list condenses the literature reviewed on students' development of understanding of covariation and its presentation as a thematic summary is an indication of the lack of research on covariation.

1. Students often focus on individual data points rather than look more broadly at the global trend in the data.
2. Students' prior knowledge about the context can enhance or inhibit students' reasoning about data.
3. Students' intuitions about covariation provide a springboard for extending their thinking further.
4. Students' understanding of covariation should be built upon an initial understanding of variation and distribution.

Models of Graphing

To inform the inquiry, a comprehensive consideration of the relevant literature was undertaken to identify theoretical models of statistical thinking and reasoning that were directly related to data analysis and in particular, graphing. Models developed by Friel et al. (2001); Jones, Thorton, Langrall, Mooney, Perry, and Putt (2000); Mooney (2002); and Moritz (2004) were considered. In addition, suggestions by Shaughnessy (2007) for an additional level to be added to Friel et al.'s levels of thinking and the notion of transnumeration presented by Pfannkuch and Wild (2004) were reviewed to provide an extensive view of the development of graphing and graph sense-making.

The model developed by Friel et al. (2001) was conceptualised around a focus on understanding students' development of graph comprehension. The term "graph sense" was introduced and described as what "develops gradually as a result of one's creating graphs and using already designed graphs in a variety of problem contexts that require making a sense of data" (p. 145). This description placed the emphasis for graph comprehension on the use of graphs as tools for making sense of information. In line with this view, they identified six behaviours associated with graph sense: recognising components of graphs; speaking the language of graphs; understanding relationships among tables, graphs, and data; making sense of a graph; interpreting a graph and answering questions about it; and recognizing appropriate graphs for a given data set and its context.

In 2007, Shaughnessy extended the work of Friel et al (2001) further and suggested two additional behaviours in recognition of the influence of context on the data analysis process and to provide a wider view of data handling. The additional behaviours were: look for possible causes of variation, and look for relationships among variables in the data. When considered together, the additional behaviours suggested by Shaughnessy gives the Friel et al. model greater depth in terms of the way students may develop graph sense. The behaviours do not, however, recognise explicitly the impact that constructing graphs may have on the development of statistical thinking and reasoning.

Prior to the work presented by Friel et al. (2001) and influenced by the work of Kosslyn (1989), Curcio (1989) considered school students' interpretation of graphs from three perspectives, *Read the data, Read within the data, Read beyond the data* (cited in Shaughnessy, 2007). Shaughnessy suggested that each of the six behaviors identified by Friel et al. (2001) fits with one of Curcio's three levels of graph reading. Shaughnessy went on to suggest extending the categories to include the two behaviours he identified under the level of *Reading behind the data.*

The levels of graph interpretation suggested by Curcio (1989) (as cited in Shaughnessy, Garfield, & Greer, 1996) have been used as a foundation for exploring primary students' interpretation of graphs (Jones et al., 1997, 2000; Mooney, 2002). Although applied successfully by Jones et al. (1997, 2000) to describe and assess students' levels of statistical reasoning and by Mooney to develop a cognitive model of graph interpretation, these studies did not extend the levels of graph interpretation further. The cognitive model put forward by Mooney can be seen as a hierarchy but like Friel et al. (2001) did not include consideration of the thinking processes associated with constructing graphs.

An extensive four-dimensional model proposed by Wild and Pfannkuch (1999) and elaborated on by Pfannkuch and Wild in 2004 includes behaviours associated with constructing graphs. The model relates to the way statisticians work and think statistically and applies to the way in which students engage in statistical investigations. It includes four dimensions:

Dimension 1: The investigative cycle

Dimension 2: Types of thinking

Dimension 3: The interrogative cycle

Dimension 4: Dispositions

Dimension 1 is related to the thinking processes employed when working through a statistical investigation. This involves planning an investigation, collecting data, analysing data, and drawing conclusions. Dimension 2 is related to the types of problem solving strategies applied when working through a statistical problem. Wild and Pfannkuch (1999) posit that the types of thinking in this dimension are "the foundations on which *statistical* thinking rests" (p. 227). Dimension 3 adopts a cyclical process of data interrogation that involves thinking critically about the data in order to distil and encapsulate ideas and information. Dimension 4 includes the personal qualities, dispositions, and habits of mind employed when working with data.

Dimension 2: Types of Thinking of the Statistical Thinking Model (Pfannkuch & Wild, 2004; Wild & Pfannkuch, 1999) is particularly useful when considering the way in which students work with data when creating graphs. The section most relevant to this inquiry is transnumeration. Pfannkuch and Wild (2004) describe transnumeration as changing data representations to engender understanding, capturing the characteristics of a real situation, and communicating messages in data. The notion of transnumeration is extremely important as new technologies that incorporate interactive and dynamic commands as a way of working within software environments foster the manipulation of both data and data representations – an element not included in other models of graphing sourced in this inquiry. Transnumeration also encompasses the application of data summary EDA strategies, such as the box-and-whisker plot.

The translation processes when reasoning about covariation detailed in the Moritz model (2004, p. 523) (Figure 1.6) are *graph production*, *graph interpretation*, and *speculative data generation*. The arrows on the model in Figure 1.6 indicate processes of translating among numerical data, graphical representations, and verbal statements of covariation. Unlike the models developed by Mooney (2002), Jones et al. (1997, 2000) and Friel et al. (2001), Moritz did not construct his model as a hierarchy. He proposed that students could enter the graph interpretation process from multiple entry points. The starting point could be constructing a graph from raw data or a verbal statement. It could also be extracting data from a graph or making informal inferences based on the trend in a graph. The model does not, however, incorporate elements associated with summarising data other than displaying the data in graphs. Of all the models accessed for this inquiry, the Moritz model is the only one to address covariation specifically.

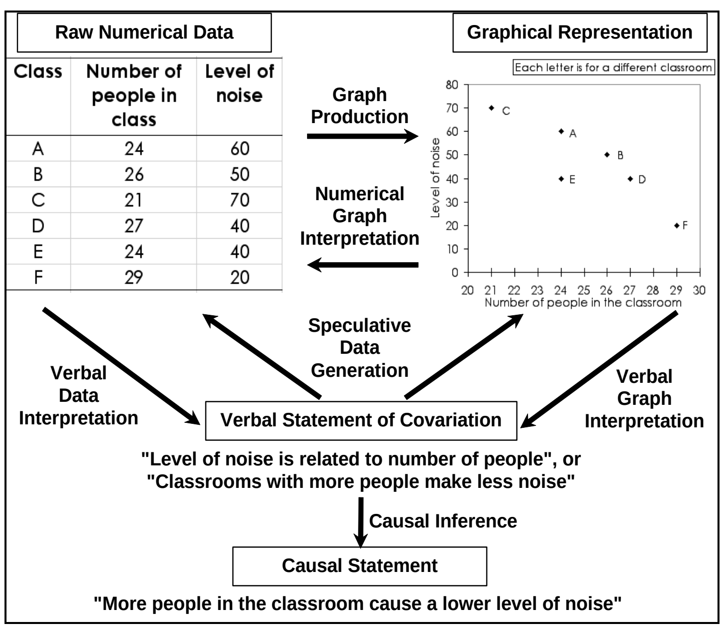


Figure 1.6. Translating processes involved with reasoning about covariation.
(Reproduced from Moritz, 2004, p. 523)

Summary of Identified Limitations of Existing Models of Graphing

Although preferring to build on previous research without modification, none of the models presented in this section took into consideration the way in which students learn in technological environments and how these environments may impact on the construction of mathematical knowledge. Nor did they focus on the construction of graphs. As such, none are completely suitable when conducting research into the ways in which students develop statistical thinking and reasoning within EDA software environments. For the purposes of this inquiry, there is the need to develop a new model of learning behaviour that takes into account the potential application of graphing software in the EDA process. Relevant elements from each of the existing theoretical frameworks cited are incorporated into the new model. To address the lack of attention in the literature to the affordances of educational technologies in relation to graphing, the work of Alessi and Trollip (2001) is integrated into the new model as well. Alessi and Trollip developed a checklist for evaluating online resources, which is somewhat applicable to this inquiry.

Development of a Model of Learning Behaviour in EDA Graphing Environments

The theoretical framework, *Model of Learning Behaviour in EDA Graphing Environments*, developed for this inquiry draws on the work of the following commentators. Friel et al. (2001), Moritz (2004), Pfannkuch and Wild (2004), and Shaughnessy (2007) in relation to graphing and graph sense-making, Kosslyn (1989) in relation to the characteristics of graphs, and Alessi and Trollip (2000) in relation to the features of technological learning environments. The framework, presented in Figure 1.7, incorporates the key behaviours of graph creation and interpretation extracted from each of the models of graphing, takes into account the characteristics of graphs and affordances of technologies from the literature into four interconnected dimensions: *Being creative with data*, *Understanding data*, *Thinking about data, and Generic knowledge*. The dimensions characterise the behaviours employed when using EDA strategies to analyse data.

<table>
<tr><th colspan="2">Dimensions</th><th>Key Behaviours</th></tr>
<tr><td rowspan="3">Generic knowledge
Speaking the language of data and graphs.
Recognising the characteristics of data and graphs.
Understanding how to use the features of software and technology environments.</td><td>Being creative with data</td><td>Reducing data to graphical representations.
Summarising data.
Constructing different forms of graphs.
Describing data from graphs.</td></tr>
<tr><td>Understanding data</td><td>Making sense of data and graphs.
Understanding the relationship among tables, graphs, and data.
Identifying the messages from the data.
Answering questions about the data.
Recognising appropriate use of different forms of graphs.</td></tr>
<tr><td>Thinking about data</td><td>Asking questions about the data.
Recognising the limitations of the data.
Interpreting data.
Making causal inferences based on the data.
Looking for possible causes of variation.
Looking for relationships among attributes in the data.</td></tr>
</table>

Figure 1.7. Model of Learning Behaviour in EDA Graphing Environments.
(Adapted from Fitzallen, 2006; Fitzallen & Brown, 2006b, 2007)

The intention of the model introduction is to recognise that each of the dimensions is an independent, yet functionally related, set of behaviours that can be accessed in isolation, in any combination, or collectively to answer questions about the data. How the dimensions of the model come together when working with software packages depends on the complexity of the statistical question asked and what needs to be done to answer the question. For example, a question that requires the time a particular event occurs to be determined from a time series graph only employs behaviours from the *Generic knowledge* dimension to answer the question. As another example, an answer to a more complex question that requires collecting the data, representing the data using EDA strategies, summarising the data, and making inferences about the data based on the representations and the variation evident to answer a question, will incorporate all the dimensions of the model in an interrelated fashion.

The *Model of Learning Behaviour in EDA Graphing Environments* (Figure 1.7) is not offered as a hierarchical construct. Like the Moritz (2004) model (Figure 1.6), it recognises that reasoning about covariation can occur from different entry points. At times, questions about data are directly related to a given graph and there is no need to create new representations, yet summarising data in another way may be useful. In this instance, entry into the data analysis process may utilise the behaviours in the *Being creative with data* dimension. On other occasions, questions about the data may be directly related to determining the trend evident in graphs. In these circumstances, access to the data analysis process may utilise the behaviours in the *Thinking about data* dimension only.

An important feature of the framework is that it recognises that there are some generic understandings that are inherent in all aspects of data analysis, graphing, and graph sense-making. These are included in the *Generic knowledge* dimension. The *Generic knowledge* dimension is situated in the model so that it stands as an independent set of behaviours, which then supports the other three dimensions that encompass behaviours associated with creating and interpreting graphs. As well as the models of graphing discussed in the previous section, this *Generic knowledge* dimension was informed by Kosslyn's (1989) schema for describing the constituent parts of a typical graph. It is important not only to understand how to read data values from a graph but also to understand how to read the constituent parts that make up the structure of a graph, such as the scale on an axis.

The *Generic knowledge* dimension also considers the technical skills required to use technology and software environments. Bakker (2004) contends that "the software itself needs to be learned before it can effectively mediate between the learner and what is to be learned" (Bakker, 2004, p. 279). This infers that influencing the effective use of learning activities that use technology is reliant on the student's knowledge of the features of the technology. Bakker's comment is important and should be taken into consideration when thinking about how students use graphing software packages to construct and interpret graphs.

To address the issues raised by Bakker (2004), the work of Alessi and Trollip (2001) in relation to the behaviours associated with using technology were incorporated into the *Generic knowledge* dimension. Alessi and Trollip developed a checklist for evaluating online learning resources for the teaching of mathematical concepts. The check list includes the usability of the learning resource interface, ease of navigation, access to invisible features via drop-down menus, and access to supplementary materials, such as the “Help” functions. Alessi and Trollip contend that the interface of technologies impacts on how students access the features of the technologies and the way in which they navigate around the learning environment. Including these features of educational technologies into the *Model of Learning Behaviour in EDA Graphing Environments* (Figure 1.7) recognises the potential for graphing software to influence student behaviour and development of understanding of statistical concepts.

The *Model of Learning Behaviour in EDA Graphing Environments* (Figure 1.7) developed for this inquiry is presented again in Table 1.1. The right-hand column of the table details the contributors from the literature that informed the development of the dimensions of the model. In this thesis, when referring to the *Model of Learning Behaviour*, the words “model” and “framework” are used interchangeably and no particular connotation is intended by the use of either word.

Table 1.1

Contributors to the Model of Learning Behaviour in EDA Graphing Environments *Dimensions*

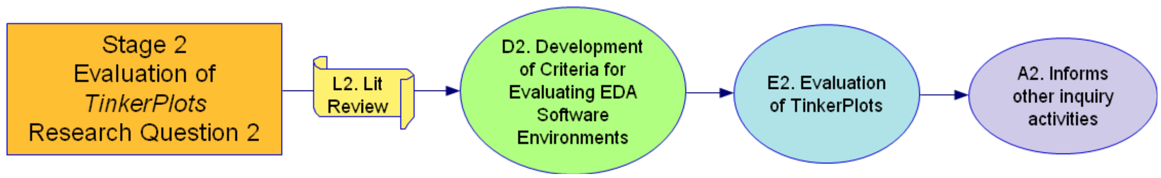
Dimensions	**Key Behaviours**	**Contributors**
Generic knowledge	Speaking the language of data and graphs.	Friel et al., (2001), *Reading the data* Moritz (2004), *Verbal graph interpretation* Kosslyn (1989), *Constituent parts of graphs*
	Understanding how to use the features of software and technology environments.	Alessi & Trollip (2001), *Navigating the interface*, *Accessing the features*, *Supplementary materials*
	Recognising the characteristics of data and graphs.	Friel et al., (2001), *Reading the data* Moritz (2004), *Numerical* and *verbal graph interpretation*
Being creative with data	Reducing data to graphical representations.	Pfannkuch & Wild (2004), *Transnumeration*
	Summarising data.	Pfannkuch & Wild (2004), *Transnumeration*
	Constructing different forms of graphs.	Pfannkuch & Wild (2004), *Transnumeration* Moritz (2004), *Graph production*
	Describing data from graphs.	Moritz (2004), *Verbal graph interpretation*
Understanding data	Making sense of data and graphs.	Friel et al., (2001), *Reading within the data*
	Understanding the relationship among tables, graphs, and data.	Friel et al., (2001), *Reading within the data*
	Identifying the messages from the data.	Pfannkuch & Wild (2004), *Transnumeration*
	Answering questions about the data.	Friel et al., (2001), *Reading beyond the data*
	Recognising appropriate use of different forms of graphs.	Pfannkuch & Wild (2004), *Transnumeration* Friel et al., (2001), *Reading beyond the data*
Thinking about data	Asking questions about the data.	Moritz (2004), *Verbal graph interpretation*
	Recognising the limitations of the data.	Shaughnessy (2007), *Reading behind the data*
	Interpreting data.	Friel et al., (2001), *Reading beyond the data*
	Making causal inferences based on the data.	Shaughnessy (2007), *Reading behind the data* Moritz (2004), *Causal inference*
	Looking for possible causes of variation.	Shaughnessy (2007), *Reading behind the data*
	Looking for relationships among attributes in the data.	Shaughnessy (2007), *Reading behind the data*

Concluding Remarks

The theoretical *Model of Learning Behaviour in EDA Graphing Environments* (Figure 1.7) developed in this stage of the inquiry provides a comprehensive view of the behaviours associated with graph creation and reasoning about graphs and takes into account the potential for the features of technologies to influence those processes. It is used in Stage 2 – Evaluation of *TinkerPlots* to develop criteria for interrogating the affordances of the *TinkerPlots* learning environment. In Stage 4 - Sequence of Learning and Outcomes, the *Model of Learning Behaviour* is used as a framework for analysing the data collected from student interviews conducted after the development and implementation of the sequence of learning experiences developed in Stage 4. The data from the student interviews are presented as student profiles, which are then utilised as data for answering Research Questions 1, 2, and 3 in Stage 5 – The Findings. The Model of Learning Behaviour is revisited again in Stage 6 – Inquiry Conclusions when an evaluation of the inquiry design is conducted to reflect the iterative nature of the education design research methodology employed for this inquiry.

Stage 2

Evaluation of TinkerPlots



The purpose of Stage 2 of the inquiry is to develop an understanding of the features of *TinkerPlots*, the graphical representations it constructs, and the learning environment it offers. As *TinkerPlots* was developed it was trialled and tested extensively; direct evaluation of the learning environment afforded by *TinkerPlots*, however, had not been conducted prior to this inquiry. Early indications from research (Shaughnessy, 2007; Konold, 2002; Konold & Higgins, 2003; Hammerman & Rubin, 2002) and anecdotal evidence (Konold & Miller, 2005) support the conviction that *TinkerPlots* is useful in developing students' understanding of statistical concepts. It is, however, important to determine the ways in which *TinkerPlots* provides the opportunities for students to engage actively in exploratory data analysis.

In this stage of the inquiry a review of the literature on *TinkerPlots'* background and development, research conducted using *TinkerPlots*, and applications of *TinkerPlots* in classrooms and professional development programs is presented. In addition an evaluation of the functionality of *TinkerPlots* was conducted using the framework, *Criteria for Evaluating Exemplary EDA Software*, developed in this stage of the inquiry. Relevant frameworks for evaluating educational software packages and visual representations, in general, were combined with the *Model of Learning Behaviour in EDA Graphing Environments* (Figure 1.7) to compile the criteria. The results from the evaluation of *TinkerPlots* are presented at the beginning of this stage in order to develop an understanding of the features referred to in the applications, research and historical development of *TinkerPlots*. The evaluation of *TinkerPlots* contributes to answering Research Question 2. It assists with understanding the learning environment afforded by the software package as well as understanding how *TinkerPlots* supports student learning.

Evaluating EDA Graphing Software

Critical examination of software packages is necessary to determine how well they meet teaching and learning needs. Evaluation of software needs to encompass not only the technological aspects – how user friendly is the package? – but also the ways the software facilitates the learning of mathematical concepts and supports thinking. It has been recognised that there is a need to be able to evaluate educational software critically in terms of the software's ability to address mathematical learning needs (delMas, 1997; Shaughnessy, 2007; Squires & Preece, 1996) and provide visualisation representations (Handal, Handal, & Herrington, 2006; Kidman & Nason, 2000). As Biehler (1997) suggested, the aim was to develop "a perspective, a guideline, an ideal system of requirements for critically evaluating existing software" in order to develop "a pedagogy of educational use of software" (p. 169).

According to Squires and Preece (1996) evaluation of software must consider the interaction between the learning process and the usability factors afforded by the software. They suggest that tools developed for this purpose "need to identify the way learning is supported by the software, the ease or otherwise with which learners are able to operate the software and the integration of usability features with learning intentions" (p. 16). Squires and Preece also suggest that any evaluation should determine how well "the software designer understood the learners' needs" (p. 19). Goyne, McDonough, and Padgett (2000) add that evaluation of software should take into consideration the interactive nature of innovative software developments.

Guidelines to evaluate existing software critically and guide the development of better software in the future have been advocated by Biehler (1997) and Shaughnessy (2007). In addition, evaluation of software design that draw on relevant learning theories has been recommended by delMas (1997) and Goos and Cretchley (2003), particularly with respect to developing higher order thinking, problem solving, and inductive reasoning. In the specific field of statistics education, Shaughnessy has similarly recommended that research be conducted to improve understanding of the features of software that elicit and transform students' statistical thinking and reasoning. These viewpoints indicate a continuing need not only to evaluate software from a technological perspective but also to ascertain how it may contribute to developing statistical concepts for students. To address these concerns, it was necessary to develop a framework that could be applied within the context of this study.

Efforts to provide teachers with the tools to determine which products may be useful in the classroom have focused on multimedia learning activities known as computer-based Integrated Learning Systems (ILS). These include stand-alone software simulations, e-learning objects, and applets that use visual representations to create learning environments for students. They often purport to facilitate the construction of students' knowledge, provide feedback, and assist in the development of higher order thinking, such as problem solving skills (Freebody & Muspratt, 2007). In 1997, delMas developed a framework for evaluating software that focused on the evaluation of simulations. The purpose of his framework was to provide a tool that could be used to guide the development of software. delMas applied his framework to a Sampling Distribution simulation and determined that using the framework helped him identify features that were missing and identify features that could be modified to enhance the effectiveness of the simulation. Although applied successfully in that context, the delMas framework is written as a set of guidelines for the development of software in general, and does not promote the examination of particular elements of software.

More recently, Handal et al. (2006) used a checklist for evaluating ILS on 500 mathematics education websites. The evaluation items checklist, developed by Alessi and Trollip (2001) included nine items: Subject, Auxillary information, Affective considerations, Interface, Navigation, Pedagogy, Invisible features, Robustness, and Supplementary materials. The checklist proved to be a useful tool for identifying the positive and negative features of online resources. Handal et al. reported that the checklist was easy to use for the evaluation of such a large collection of resources. The application of the checklist highlighted the differences in design and usability among ILS resources but, like the delMas framework (1997), did not focus the evaluation on specific mathematics content. As Squires and Preece (1996) warned, items on check lists are open to interpretation, focus on technical rather than educational issues, and do not take into consideration the interactive nature of innovative un-populated software packages. Squires and McDougall (1996) also suggested that checklists failed to accommodate the evaluation of different subject areas as each subject area required different sets of selection criteria.

The Development of Criteria for Evaluating EDA Graphing Software

Guidelines to evaluate critically existing software and guide the development of better statistical software in the future have been advocated by Biehler (1997), Shaughnessy (2007) and Camden (2009). In addition, evaluation of software design drawing on relevant learning theories has been also recommended by Goos and Cretchley (2003), who emphasise a need to evaluate the capacity of software to enhance students' higher order thinking. In the specific field of statistics education, Shaughnessy similarly recommends that research be conducted to improve understanding of the features of software that elicit and transform students' statistical thinking and reasoning. This inquiry attended to these continuing calls from the literature by evaluating *TinkerPlots.* It was, however, necessary to develop a set of criteria for evaluating EDA software packages to apply as a review of the literature revealed that there were no suitable established frameworks available.

In order to develop criteria for evaluating *TinkerPlots*, it was necessary to use the relevant models of statistical thinking and reasoning about data and graphs encapsulated in the *Model of Learning Behaviour in EDA Graphing Environments* (Figure 1.7) as well as draw on generic frameworks that evaluate educational software packages, in terms of the visual representations generated and the learning environment offered. The three frameworks relevant to the context of this inquiry are:

- *Practical Guidelines for Evaluating Educational Software* (Goyne et al., 2000);

- *Principles for Analysing Visual Representations* (Kidman & Nason, 2000); and
- *Model of Learning Behaviour in EDA Graphing Environments* (Figure 1.7).

In this section, the first two frameworks are interrogated and discussed in relation to the applicability of the frameworks to the learning environment afforded by EDA software. The relevant features of these frameworks are then melded with the third framework that focuses on thinking and reasoning about data and graphs in EDA graphing environments (Figure 1.7). This culminates with the development of the *Criteria for Evaluating Exemplary EDA Software*. The checklist developed by Alessi and Trollip (2001) and utilised by Handal et al. (2006) is not included specifically at this stage as it was incorporated into the *Model of Learning Behaviour in EDA Graphing Environments* (Figure 1.7). The features of the checklist that contributed to the model are noted in Table 1.1.

Guidelines for Evaluating Educational Software

To assist teachers and administrators to select appropriate educational software from the myriad of options available, Goyne et al. (2000) developed a set of practical guidelines for evaluating software. The guidelines are a series of twelve questions that focus on the important elements of interactive software.

1. Is the software consistent with the curriculum and learning outcomes?
2. Does the software offer the learner choices and control?
3. Does the software provide a variety of appropriate media and activities?
4. Does the software provide positive, formative feedback and achievement measures?
5. Is the software appropriately challenging?
6. Does the software have high-quality technical components?
7. Does the software provide opportunities for practice and review?
8. Does the software present material in an enjoyable, interesting manner?
9. Is the software interactive?
10. Will the software allow for social interaction?
11. Will the software foster learning in an authentic, relevant context?
12. Will the software be accessible to students?

The twelve questions posed by Goyne et al. (2000) are useful as they recognise the elements of interactive software that the frameworks designed for evaluating ILS resources lacked. Of the twelve questions, only question 10 does not at first appear to be directly relevant to EDA software. Technologies today, however, allow for the social interaction of users through social media applications, such as Facebook or Twitter. This is not the primary purpose of EDA software, but loosely applied social interaction could include the use of EDA software as a medium for sharing of information and the publication of a group project. With this interpretation in mind, question 10 is relevant and is considered in this vein. The other 11 questions are considered self-explanatory and do not need further clarification.

Evaluating Visual Representations

In order to identify the way in which visual representations used in ILS facilitate students' construction of mathematical knowledge, Kidman and Nason (2000) developed a set of seven principles (Table 2.1) that are based on "the research literature from the fields of mathematics education, cognitive science, computer-aided learning, computer graphic design, and semiotics" (p. 179). These principles were applied successfully by Kidman and Nason as a diagnostic tool to determine the effectiveness of dynamic mathematical representations employed within ILS. The principles were used to evaluate 500 ILS, which, for the most part, were found to only partially fulfil the requirements for interactive construction of mathematical knowledge. The principles developed by Kidman and Nason focus on the evaluation of general mathematical content and need to be interrogated further to determine if they are suitable for evaluating EDA software.

Table 2.1
Principles for Analysing Visual Representations (Kidman & Nason, 2000, p. 181)

No.	Visual representations should:
1	be clearly displayed and explicitly understood by the student. This facilitates the process of stimulating relationships among the problem data and may also help students to recall knowledge and skills by making connections between prior internal representations and new situations.
2	enable the student to focus on the deep structural rather than surface structural aspects of the problems being investigated.
3	provide physical/iconic environments for learners to abstract and understand mathematical concepts or a relationship of information within problems.
4	provide students with an external memory to display information temporarily during the process of problem solving. By doing this, the visual representation can reduce the working memory demands of the problem solving process.
5	facilitate the exploration and construction of understandings about aspects of mathematical ideas and concepts that cannot be adequately represented in the semantics of natural language.
6	facilitate the process of translating between mathematical expressions and natural language.
7	be used for both interpretative and expressive learning activities.

Principle 1. This principle focuses on the clarity of visual representations and assumes there is a causal link between providing a clear display and the development of thinking and understanding. It also maintains that if the visual representation is explicitly understood, then students may be able to draw on prior learning and translate the learning gained to new situations. The interactive nature of some ILS may facilitate this but it cannot be assumed that the display itself is a main contributor. With this in mind, it appears that Principle 1 as it stands is insufficient. When considering EDA software it is necessary to consider if students can access and use the features of the software as well as understand the display.

Principle 2. Mathematical content and the development of deep learning are the focus of Principle 2. The development of deep learning should be an expectation of any mathematical learning environment, therefore, essential to include in an evaluation. In the case of ILS, such as e-learning objects, the evaluation of mathematical content and the extent to which it is developed is necessary, as the information is usually displayed without any input from the user. In the case of un-populated software (without any data entered), the display of information is in the control of the user and the software provides a different learning context to those used in ILS resources. With EDA software in particular, deep learning may be facilitated by interacting with the software, engaging with learning activities, and accessing support offered by the teacher. It is, therefore, important to go beyond evaluating what mathematical content is accessed and developed to including what opportunities are provided for students to develop their mathematical understanding further within the context of the learning environment.

Principle 3. This principle takes into account the learning environment and the role it has to play in allowing students to make connections across information within problems. It is, however, quite broad and does not allude to the specific aspects of the ILS environment that may be attributed to the development of understanding. Additionally, it does not consider how students may engage in the learning sequence. Determining if ILS allows learners to have some control over the learning sequence is important. This is particularly relevant to statistics education as students make sense of data through activities, such as constructing and deconstructing graphs, comparing data sets, and representing data in different ways (Watson, 2006). For this to be successful, the learning environment needs to be flexible enough for students to engage with the learning sequence from different entry points.

Principle 4. This is an important principle that recognises the potential of ILS to provide an external memory to display information temporarily. It supports the notion that the ILS environment can display some information thereby reducing the cognitive load when problem solving. This principle also recognises that computer environments should provide students with the tools to reorganise information and support their thinking. Those tools and the external memory may be in various forms. The Help function, for example, provides external memory as do drop-down menus and pop-up annotations. Tools for reorganising information may be graphing functions or calculators as well as text boxes for inserting information.

Principle 5. Principle 5 recognises the value of visual representations for conveying mathematical ideas and concepts that are not easily expressed using natural language. In statistics education graphs are used for this purpose. Often, trends or patterns in data are more easily expressed in a graph than explained from raw data. This principle does maintain that visual representations should facilitate the exploration and construction of mathematical ideas and is related to Principle 3, as the learning environment determines the way in which students engage in the learning process.

Principle 6. This principle makes the assumption that ILS display both natural language and mathematical expressions. As this is not always the case, this principle is only partially appropriate. ILS, such as learning objects, display mathematical representations and explanatory notes in both symbols and written text, whereas EDA software is un-populated with data, and requires input from the user to construct visual representations that are not easily represented by natural language. The value of EDA software lies in the user being able to construct visual representations that are not easily represented in natural language.

Principle 7. This is the only principle in the Kidman and Nason (2000) framework that recognises that students can contribute actively to the learning process. They note the important role visual representations play in the construction of knowledge and describe that they do this through engagement with interpretive and expressive learning activities. Expressive learning activities provide the opportunity for students to construct and/or modify visual representations, such as manipulate an axis of a covariation graph to change the scale. Interpretive learning activities come into play when students engage in activities that require them to reason, think logically, and make decisions. Such activities include comparing data sets with box plots, making inferences about larger populations from the results gleaned from a smaller sample, and making judgements about the inclusion of outliers.

Summary of Evaluating Visual Representations

Education software evaluation frameworks must consider the opportunities provided for students when engaging with interactive software. They must accommodate the two-way nature of software, where the user plays a large role in the development of the mathematics, has autonomy and control of the learning that the software provides, contributes creatively to ideas and solutions, and responds to feedback from the software.

Although the *Principles for Analysing Visual Representations* (Kidman & Nason, 2000) were designed for applying to ILS resources, they provide an acceptable framework for the analysis of general mathematical software. They do not, however, take into account the interactive nature of EDA software nor do they consider pedagogical issues related to the learning of statistical concepts highlighted in the *Model of Learning Behaviour in EDA Graphing Environments* (Figure 1.7). Neither do they acknowledge that students play an active role in the learning process. In order to be considered for use with EDA software the Kidman and Nason set of principles need to be more specific.

The guidelines developed by Goyne et al. (2000) are useful as they recognise the role of the learner when using software; however, like the Kidman and Nason set of principles, the guidelines do not look closely at any specific content area or students' application of software. As a consequence, there is the need to develop a new framework that considers the useful elements of the Kidman and Nason' set of principles, the Goyne et al. guidelines, as well as the way in which students develop an understanding of EDA skills in relation to graphing detailed in the *Model of Learning Behaviour in EDA Graphing Environments* (Figure 1.7).

Criteria for Evaluating Exemplary Statistical Software

In this section, the thesis continues to address the call from Shaughnessy (2007) and others to develop criteria for evaluating EDA software. The approach taken recognises the advantages and limitations of the frameworks developed by Kidman and Nason (2000), Goyne et al. (2000), and the *Model of Learning Behaviour EDA Graphing Environments* (Figure 1.7) developed for this inquiry.

In the previous section the *Principles for Analysing Visual Representations* (Kidman & Nason, 2000) were interrogated specifically for their application to statistical software. The set of principles has some relevance to the context of the inquiry but lacks the student-centric specificity needed in order to be applied effectively to EDA software. The principles do not account for the input of the user and the way in which statistical concepts are developed through the use of software. They are better suited to evaluating ILS that require very little input from the user and display openly the information about the mathematical ideas and concepts to be developed.

To accommodate for the way in which students engage with statistical analysis activities the *Guidelines for Evaluating Educational Software* developed by Goyne et al. (2000) are considered useful for evaluating statistical software because the guidelines address the gap identified in the Kidman and Nason (2000) set of principles. They too, however, have shortcomings in relation to identifying the features and elements that facilitate the development of specific mathematical content.

To be relevant for evaluating EDA software the Kidman and Nason (2000) set of principles and the Goyne et al. (2000) guidelines are combined and incorporated with the *Model of Learning Behaviour in EDA Graphing Environments* (Figure 1.7), which looks broadly at statistical thinking and reasoning about data and graphs without reference to the impact of visual representations. To determine the criteria the main themes from Goyne et al. and Kidman and Nason were extracted and combined with the associated dimensions of the *Model of Learning Behaviour in EDA Software Environments* (Figure 1.7). The resulting framework, *Criteria for Evaluating Exemplary EDA Software*, is presented in Table 2.2. The relationships between the three frameworks used and each criterion are noted in the right-hand column. No attempt is made to modify the work of others, but instead to capture the advantages of each framework and ameliorate the limitations before the application of the resulting criteria in this inquiry.

Table 2.2.
Criteria for Evaluating Exemplary EDA Software and Contributors to Each Criterion

No.	Criterion	Contributors
1	Ability to provide an accessible interface and software features that are easy to use.	Kidman & Nason (2000), *Principles 1 & 3* Goyne et al. (2000), *Questions 2, 6, 8, & 12* Figure 1.7, *Generic knowledge*
2	Ability to assist with the recall of knowledge and represent data in different and multiple forms.	Kidman & Nason (2000), *Principles 1 & 5* Goyne et al. (2000), *Questions 3 & 7* Figure 1.7, *Being creative with data, Understanding data*
3	Ability to facilitate the process of translating between mathematical expression and natural language.	Kidman & Nason (2000), *Principles 6 & 7* Goyne et al. (2000), *Question 1* Figure 1.7, *Being creative with data*
4	Ability to provide extended memory when organising and reorganising data.	Kidman & Nason (2000), *Principle 4* Goyne et al. (2000), *Question 5* Figure 1.7, *Being creative with data, Understanding data*
5	Ability to provide an interactive environment that allows multiple entry points for abstraction of concepts.	Kidman & Nason (2000), *Principles 2, 3 & 5* Goyne et al. (2000), *Questions 1, 7, 10 & 11* Figure 1.7, *Being creative with data, Understanding data*
6	Ability to construct and display representations used for both interpretative and expressive learning activities.	Kidman & Nason (2000), *Principles 2 & 7,* Goyne et al. (2000), *Questions 4 & 9* Figure 1.7, *Generic knowledge, Thinking about data*

A version of the *Criteria for Evaluating Exemplary EDA Software* was published previously (Fitzallen & Brown, 2006; Fitzallen, 2007). That version only included consideration of the *Model of Learning Behaviour in EDA Graphing Environments* and the Kidman and Nason (2000) principles. The version in Table 2.2 has the same criteria as the 2009 version but is strengthened by the addition of the contribution from Goyne et al. (2000).

The *Criteria for Evaluating Exemplary EDA Software* presented in Table 2.2 progress the work of Kidman and Nason (2000) and Goyne et al. (2000) by specifically considering the way in which students develop an understanding of statistical concepts and of the interactive nature of EDA software. The criteria are used in the next section to evaluate the EDA software, *TinkerPlots*, used by students in this inquiry.

Evaluation of *TinkerPlots* Using Criteria Developed

In this section, the *Criteria for Evaluating Exemplary EDA Software* developed for this inquiry are used to evaluate the learning environment afforded by *TinkerPlots*. This evaluation only interrogates the basic functions of *TinkerPlots*, which includes the Plot, Hat, Dividers, and Measures of Centre functions. Functions such as the Slider, the Filter, the Dividers, and the Add Case Tool are not considered as they do not apply to the context of the inquiry and the tasks undertaken by the students.

Criterion 1 – Ability to provide an accessible interface and software features that are easy to use

TinkerPlots is an unpopulated software application. Like suites of software that are used for word processing and publishing, *TinkerPlots* provides an interface for entry of information by the end-user. When initialised, *TinkerPlots* presents a clean and uncluttered interface, satisfying Criterion 1 of the *Criteria for Evaluating Exemplary EDA Software* (Table 2.2). The screen consists of a blank white page with two menu bars at the top of the screen (Figure 2.1). The file is a blank file with no data included. From this starting point the user can either enter data to create a new file or browse and open saved files. *TinkerPlots* comes with 40 multivariate data sets.

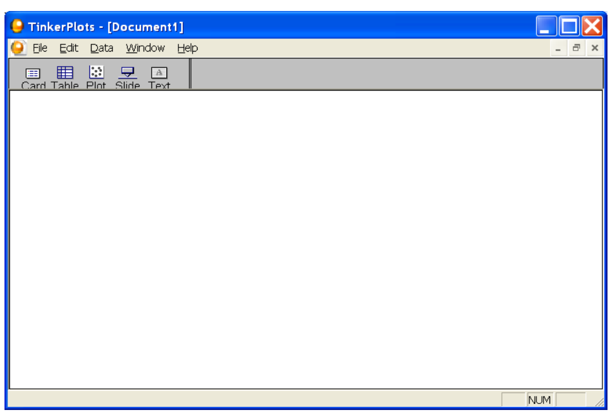


Figure 2.1. Blank page displayed when *TinkerPlots* is initialised.

The menu bar at the top of the page has the usual selection of file management options found in most computer software applications (Figure 2.2a). The File-Edit-Data-Window-Help items have drop-down menus for making selections and have most of the functions expected in software today (Figure 2.2b). The items in the second menu bar are icons that represent the choice of data displays available in *TinkerPlots* – Card, Table, Plot, Slider, and Text (Figure 2.2a). The functionality of these displays is described as they apply to other criteria. To access these data displays, *TinkerPlots* uses a drag-and-drop function to place the selected data display onto the blank page. Figure 2.2c shows the initial and final stages of dragging a Card icon onto the screen. Any, all, or multiples of the data displays can be introduced onto the screen as required.

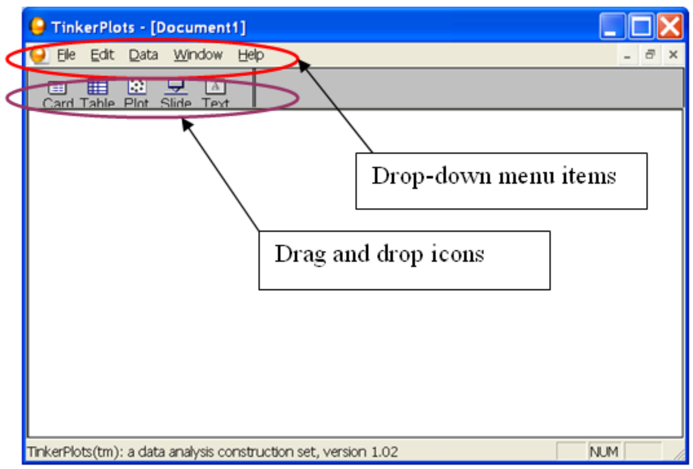


Figure 2.2a. Menu displays in *TinkerPlots*.

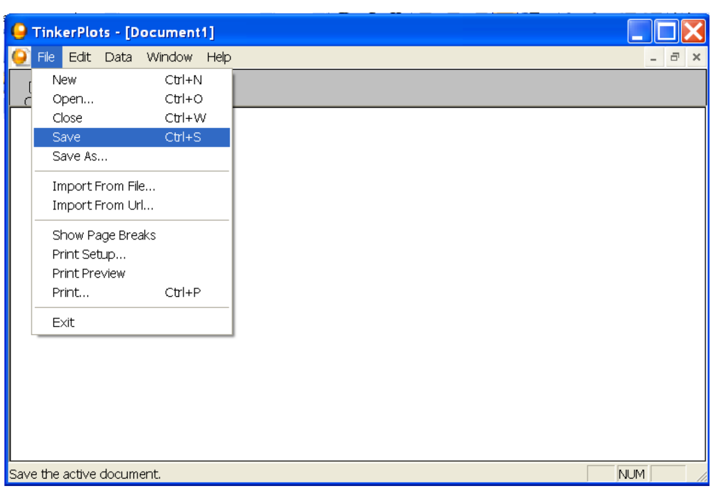


Figure 2.2b. Drop-down menu accessed from the File item.

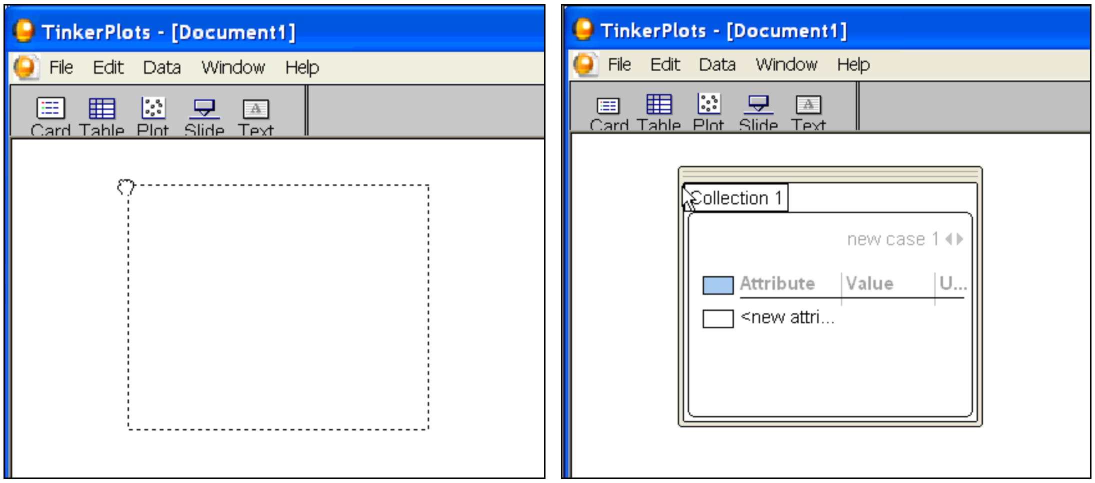


Figure 2.2c. Result of dragging the Card icon onto the screen.

To assist the user, *TinkerPlots* has an extensive Help section. This details the features of *TinkerPlots* and describes clearly how to use the program. The Help section, however, is written as a teacher resource and is not student-friendly. Although the software was developed for students in Years 5-8, the instructions in the Help section appear not to be written with these students in mind. However, this is not unique to *TinkerPlots*.

In the Help section, *TinkerPlots* has five video tutorials that demonstrate and explain the basic features of the software. These are in the form of jpeg movies that run only for a few minutes each. The movies are more useful than the Help section for younger students. The availability of the video tutorials within the program enables the user to access the tutorials at any time to refresh the user's understanding of how to utilise the features of the software. The video tutorials are also available online at www.keypress.com.au/tinkerplots.

Criterion 2 – Ability to assist with the recall of knowledge and represent data in different and multiple forms

TinkerPlots satisfies Criterion 2 as data can be represented in multiple forms - data cards, tables, and/or displayed in graphs. All three representations can be displayed at the same time, as can multiple copies of each representation type. The interface allows the user to move seamlessly back and forth between different data displays to make changes to the data displays and add data whilst the software maintains the connections among the various data displays. This has the potential to facilitate the recall of information.

TinkerPlots includes a stacked data card system for the organisation of case-based data. The stacked data cards store information about the attributes of an individual case on a single card and there is one card for each case in a data set. Once the first card in the data set is organised with the nominated attributes and units of measure, the following cards will display the same information. For subsequent cases it is only necessary to enter the actual values of the attributes into the cards. Clicking the cursor on the direction arrows at the top of the data cards provides access to other cases in the data set.

When data are entered into the data cards after opening a Card window by dragging the Card icon onto the screen, the information is recorded simultaneously into a table. This can be observed if a Table widow is open on the screen at the same time. It is, however, not necessary for a table window to be open for the data to be entered into a table. The table function has a spreadsheet format, similar to that used in other commonly used spreadsheet software.

As well as entering the data into the cards, there is the option of entering the data directly into the table. This facility reverses the process and displays the data in the stacked data cards automatically. Additionally, data can be imported directly into *TinkerPlots* from spreadsheets constructed in other software, provided the data are organised with the variables horizontally and the cases vertically. Again, as the data are imported they are organised automatically into both the data card and the table formats. It is also possible to import data directly from a web page, such as DASL (http://lib.stst.cmu.edu/DASL/) by dragging the URL into an empty *TinkerPlots* document. *TinkerPlots* puts the data automatically into the data cards.

The Plot icon in the second menu bar is used to open a window on the screen to open a window where graphs are created. When a Plot window is open at the same time as data are entered into the cards or a table, a circle icon representing each case is added to the Plot window. At this initial stage, the icons are placed in a random manner in the window without showing any information about the cases, nor are they contributing to any graphical representation. An additional menu is added to the *TinkerPlots* interface as the Plot window is opened (Figure 2.3). This provides functions that can be used to interrogate the data, including sorting and stacking the data, adding reference lines to a graph, and using various measures of centrality.

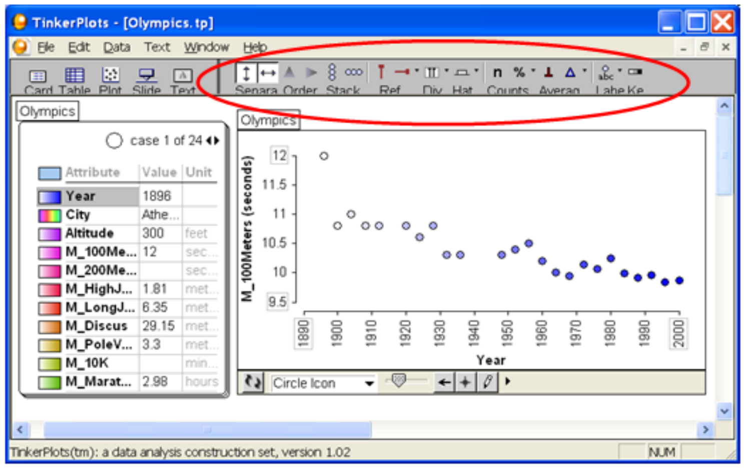


Figure 2.3. Plot window menu.

Figure 2.4a shows the on-screen display with Card, Table, and Plot windows open before any data are entered. Figure 2.4b shows the on-screen display after one case has been entered. In this example, information about only one attribute has been recorded. The circle icon in the Plot window shows that there is one case in the data set. An example with multiple cases and multiple attributes is shown in Figure 2.4c. One icon in the Plot window in Figure 2.4c is highlighted red. This happens when an icon in the Plot window is clicked with the cursor. When this occurs the data for the case are highlighted in the table and the corresponding data card is brought to the front of the stacked data cards.

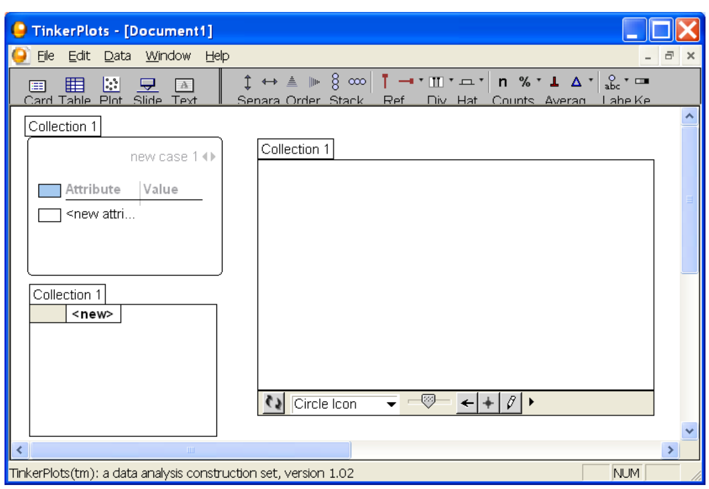


Figure 2.4a. On-screen display with no data.

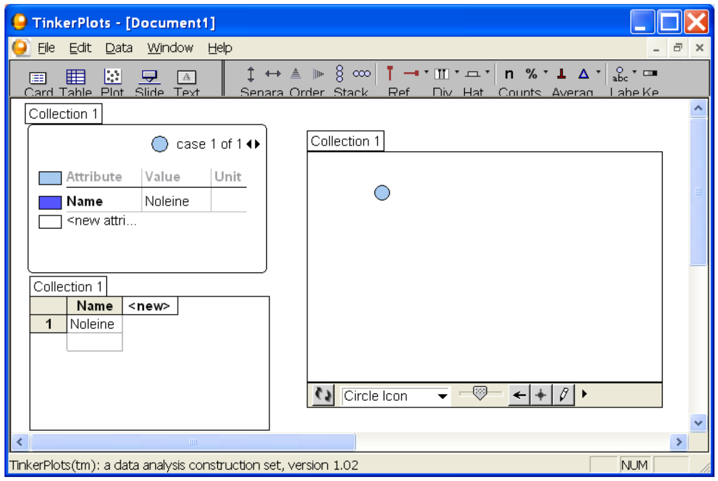


Figure 2.4b. On-screen display with data for one case entered.

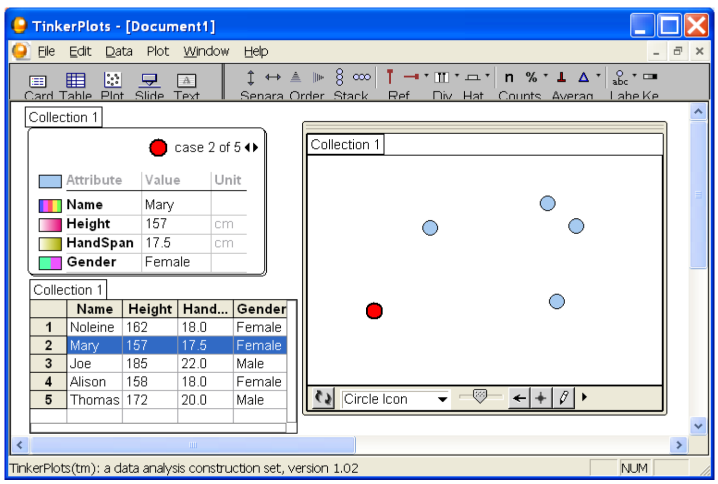


Figure 2.4c. On-screen display with data for multiple cases entered.

Criterion 3 – Ability to facilitate the process of translating between mathematical expression and natural language

TinkerPlots satisfies Criterion 3 in a two specific ways. The first is the overall learning environment afforded by *TinkerPlots* and the other is the innovative Hat Plot representation that can be overlaid on graphs.

As the actions in *TinkerPlots* are icon-based, the software does not provide information or explanations about the statistical concepts in natural language at the user interface. It does, however, have that information in the Help section. *TinkerPlots* provides the opportunity for users to convey their understanding of the representations they have created through adding text to a Text box to explain their thinking or using the drawing tool to write user-generated commentaries directly onto the Plot window. Of particular importance is the flexibility with which the user can move back and forth between the different functions.

TinkerPlots provides a learning environment where students can engage actively in the data analysis process by entering data, creating graphs, summarising data, and comparing data sets. It is also possible to import digital images, display tables of data, and add written commentaries to the *TinkerPlots* interface, as a statistical investigation is conducted. The capacity to display various forms of information on the screen allows for the context of an investigation to be transferred easily to the software interface (Figure 2.5). The opportunity to add written comments into Text windows placed on the screen as statistical investigations are undertaken, allows students to demonstrate their understanding of the context and the data in natural language that is meaningful to them.

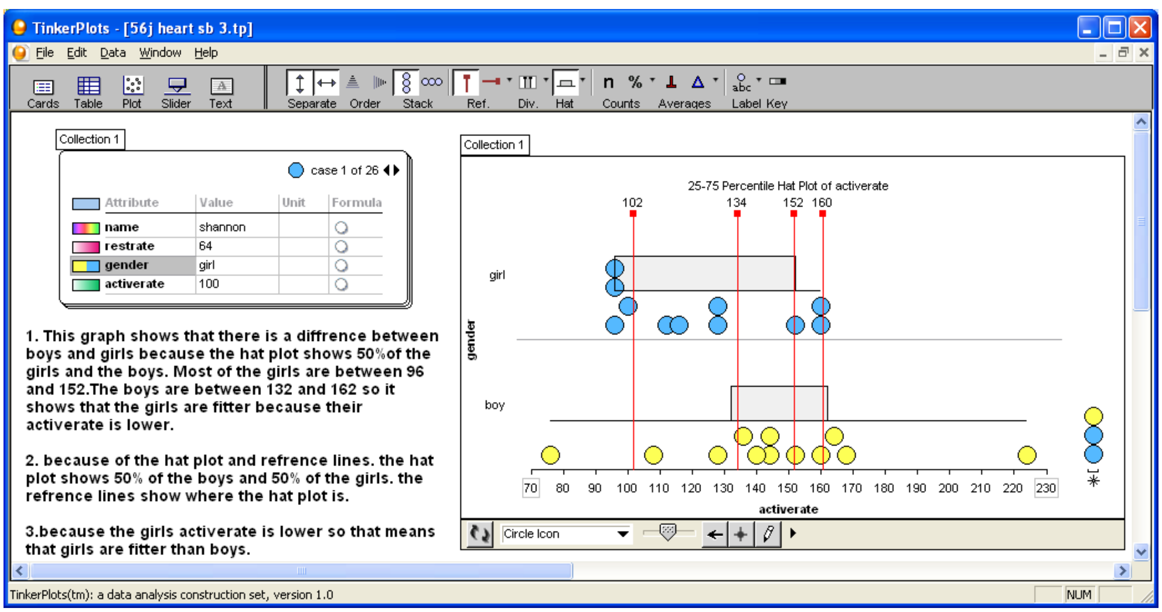


Figure 2.5. Report produced in *TinkerPlots* by a student in the MARBLE project (Watson & Donne, 2009).

TinkerPlots utilises a unique summary tool called a “hat plot.” The hat plot resembles a hat and is made up of two main components – the crown and the brim. When first initialised, the crown of the hat is a rectangle that shows the middle 50% of the data and the brim of the hat is a line that extends across the full range of the data set. When accessed from the Plot menu, the hat plot is superimposed onto a graph (Figure 2.6). It can only be applied when one or more of the axes have a continuous scale. When superimposed onto a covariation graph, the hat plot is representative of the data associated with the horizontal axis.

The hat plot is a generalised version of the box plot and as such provides a representation that can be described using natural language rather than the statistical language used to describe box plots. For example, students can use the term crown instead of inter-quartile range when describing the middle 50% of a data set. It is advantageous for students to use the analogy of the crown for identifying the inter-quartile range as it provides them with a tangible construct that they can relate to immediately for discussing their developing understanding of inter-quartile range in natural language (Watson, Fitzallen, Wilson & Creed, 2008). The crown analogy has the potential to support students to make the transition from making meaning from the physical representation of a hat to understanding the formal notions of inter-quartile range as a statistical measure, when represented in box-and-whisker plots.

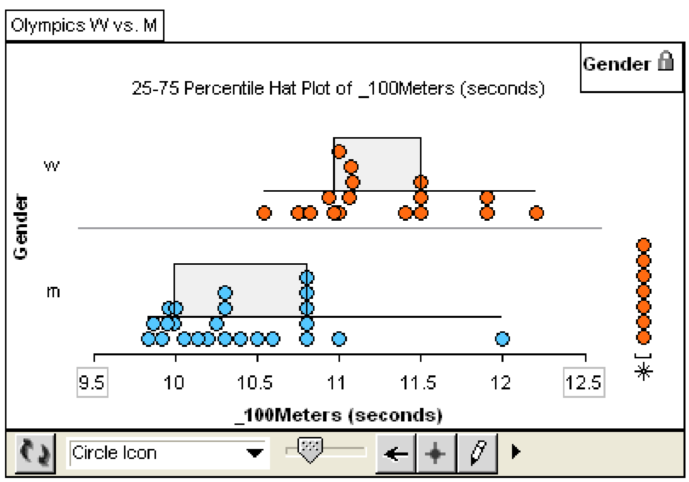


Figure 2.6. Typical hat plot representation on a split stacked dot plot.

Criterion 4 – Ability to provide extended memory when organising and reorganising data

TinkerPlots satisfies Criterion 4 as many of the features of *TinkerPlots* provide extended memory for organising and reorganising data. This is demonstrated when data points in a graph in *TinkerPlots* are manipulated by the drag-and-drop function in order to change the scale of a graph. Changes are quick and fluid, and are viewed as animations. Actions such as these change the scale and the form of the graph quickly. Figure 2.7a shows a graph that could be used to explore the relationship between the weight and height of basketball players. The graph has a continuous scale and is a typical covariation graph. Dragging a case icon to the left in the graph and adding the mean to the graph from an icon in the Plot menu results in the display changing to the example in Figure 2.7b. The scale is now separated into sections called bins. The mean weight for each bin is represented by a triangular icon. The numerical value of the mean can also be added to the screen. In Figure 2.7c the graph has been changed by swapping the axes, changing the scale of the weight axis to a continuous scale, and the height axis to bins. The means now apply to the height of the players. It is also possible to construct multiple graphs to be viewed at the same time from the one data set.

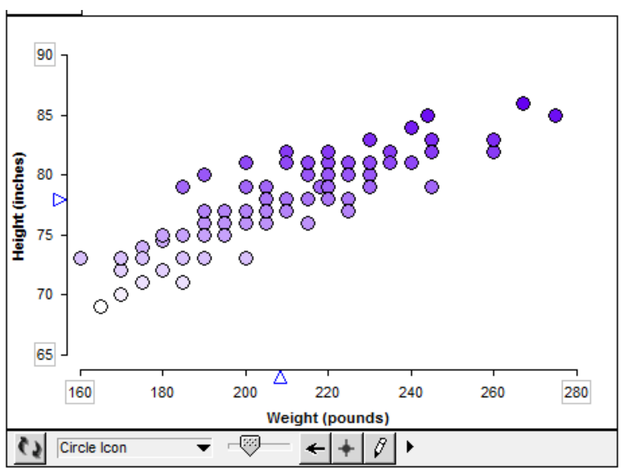


Figure 2.7a. Graph with a continuous scale on each axis showing the relationship between weight and height of basketball players.

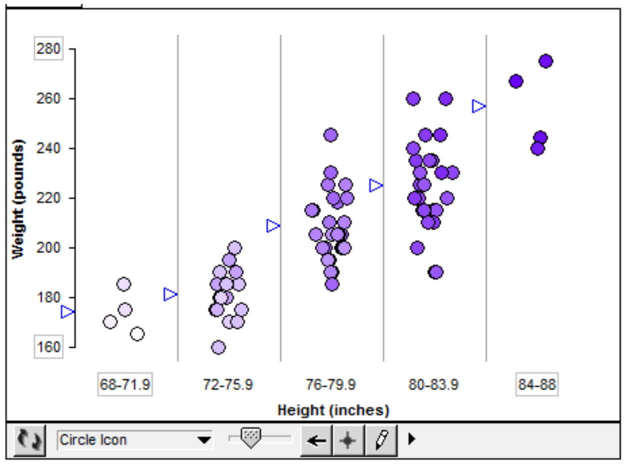


Figure 2.7b. Graph with the horizontal scale divided into bins showing the relationship between weight and height of basketball players.

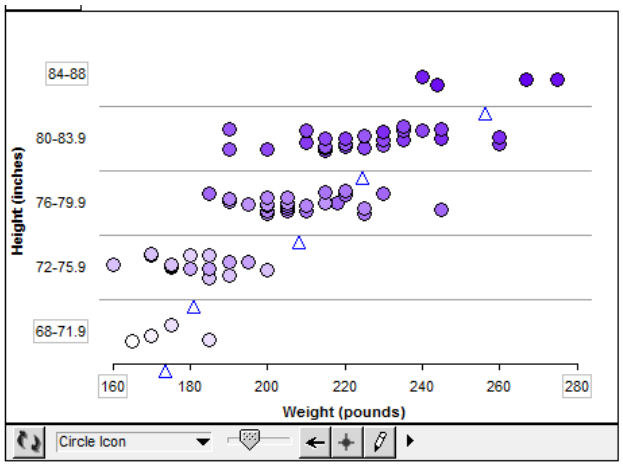


Figure 2.7c. Graph showing the relationship between height and weight of basketball players after the axes were swapped from the graph in Figure 2.7b.

Extended memory is made available when computer software enables multiple tasks to be initialised at the same time, some of which may not be immediately obvious to the user. These include background printing or automatic saving of documents. The extended memory in *TinkerPlots* has the usual "undo" and "redo" functionality common in many software packages. It also has a refresh function that mixes up the data when the user wants to restart the graph creation process. The Mix-up button is found at the bottom left hand corner of the Plot window and is circled in Figure 2.8a. When initialised the icons in the graph swirl in the Plot window as an animation and the axes and the bin separators are removed from the Plot window (Figure 2.8b).

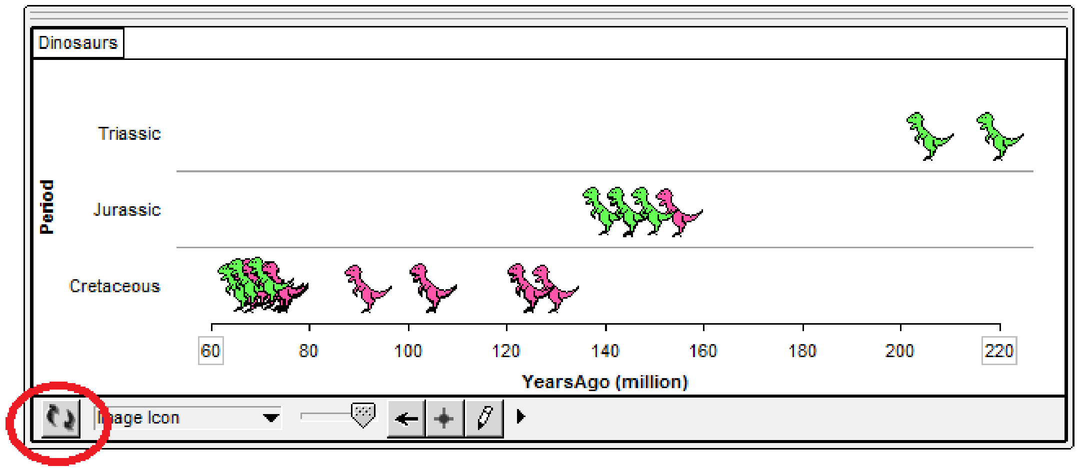


Figure 2.8a. Mix-up tool circled on a plot window.

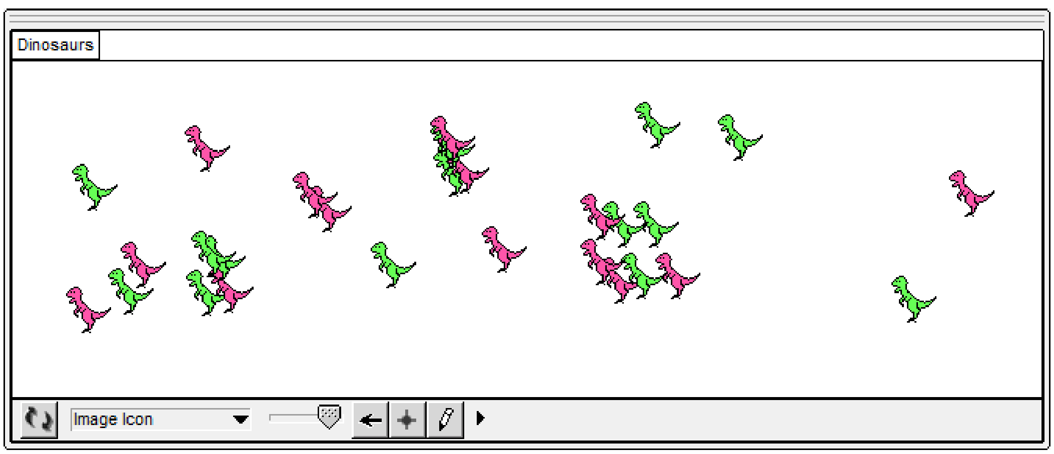


Figure 2.8b. The graph from Figure 2.7a after using the Mix-up tool.

Criterion 5 – Ability to provide an interactive environment that allows multiple entry points for abstraction of concepts

An excellent feature of *TinkerPlots* is the way a set of data can be explored and accessed from multiple entry points thereby satisfying Criterion 5. *TinkerPlots* has three main forms of data display: case-based data cards, tables, and graphs. As noted in Criterion 2, it is possible to display all three representations within *TinkerPlots* at the same time. Entry into the data analysis process can be achieved from any or all of these representations. Graphs can be constructed by dragging and dropping the name of the attribute from either the Case Cards or the Table into a Plot window to construct either the horizontal or vertical axis of a graph. Graphs can then be changed to represent a different attribute by dragging the new variable name onto the axis of the graph in the plot window. Scatterplots that display covariation are constructed by dragging a second attribute onto the axis not already occupied.

Statistical investigations with *TinkerPlots* can start with the entry of raw data and then proceed to graph creation and graph interpretation activities. Alternatively, *TinkerPlots* can be set up as a student activity with the graphs already created and questions to explore written in text boxes. The data can be added and graphs created from the data by a teacher with the intention that students need only answer questions about the graphical representation already created. The entry point into the data analysis process is different in both of these cases as is the thinking evoked through the different access points. The abstraction of concepts, however, is facilitated by the ability to manipulate and change the graphical representations until a representation is constructed that is meaningful to the user and answers to his/her satisfaction the questions about the data being explored. The sorting and stacking functions contribute to this purpose by organising data into clusters and arranging data in a particular order.

Criterion 6 – Ability to construct and display representations used for both interpretative and expressive learning activities

TinkerPlots provides the opportunity for the user to engage in both interpretative and expressive learning activities thereby satisfying Criterion 6. The user has access to many functions that assist in the interpretation of graphs. These include: summarising the data by adding the hat plot, the mean, the median, and the mode to a graph (e.g., Figure 2.7b); transforming the data by changing the scale of a graph from a continuous scale to a scale with bins (Figures 2.7a & 2.7b), or vice versa; comparing data by constructing a split stacked plot (Figure 2.6); and using the drawing tool to indicate a trend (Figure 2.9). Another useful tool is the reference line. A reference line can be inserted into a Plot window either horizontally or vertically and multiple reference lines can inserted on the one graph. Reference lines can be used to mark a cut-off point in a graph or be used to denote the value of a particular data point (e.g., Figure 2.10).

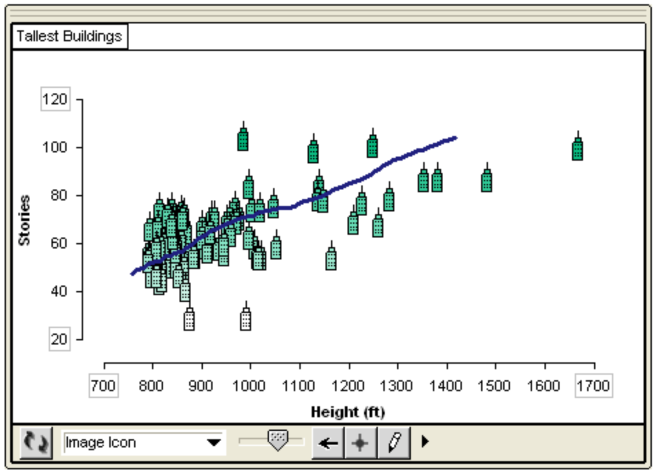


Figure 2.9. A trend line drawn using the drawing tool on a scatterplot displaying covariation.

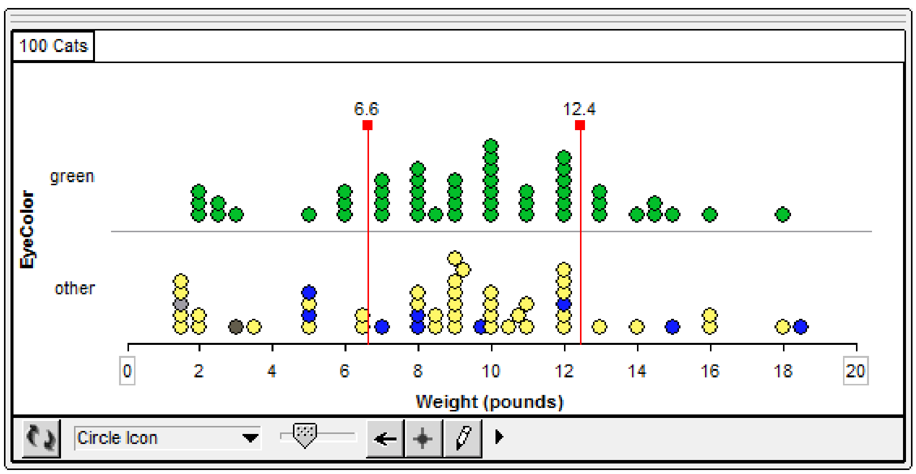


Figure 2.10. Reference lines used to section off the middle clusters of data.

Expressive learning activities involve creating visual representations to convey meaning. *TinkerPlots* facilitates this by allowing users to be autonomous and to have full control of the graph creation process. *TinkerPlots* does not have any pre-established graph types built in, hence users are free to construct graphical representations that assist them in answering questions about the data. Users can also personalise the graphs and the information by selecting the font type, size, and colour of text in a Text window, adding photos, selecting different icon shapes including picture icons (Figures 2.8a, 2.8b & 2.9), and putting the information together in ways that are meaningful to them to create engaging and individualised displays.

Summary of Evaluation of TinkerPlots

The evaluation of TinkerPlots identified may strengths and positive features inherent in *TinkerPlots*. No apparent features were identified as possibly constraining or inhibiting students working flexibly with *TinkerPlots*. The evaluation, however, identified that the Help function was written for teachers and not user-friendly for students.

The evaluation suggests that the graph creation process in *TinkerPlots* makes available visual representations that provide immediate feedback to students as the animations of changes instigated are fluid and quick. This begins with the input of the data into a table or data card system, moves to the creation and manipulation of graphical representations, and extends to the interpretation of graphs through using box plots and other means of summarising data. Throughout the process the user is in control of the functions and is provided with the opportunity to maintain the connections between the data and the graphical representations.

When using *TinkerPlots*, the creative process is not constrained by the application of pre-established definitions or graph types. This allows the user to focus on the interpretation of the data as part of the creation of graphical representation process. As *TinkerPlots* takes on the work of rescaling axes and the placement of icons on the graph, the burden of graph creation is reduced for users, which allows for the emphasis to be placed on the interpretation of graphs rather than on how to construct and create the graphs. As little effort is required to change the graphs, the user can construct, modify, and reconstruct a number of graphs quickly in order to decide which graph provides the best representation for answering the question being explored.

Development of *TinkerPlots:* A History

TinkerPlots: Dynamic Data Exploration (Konold & Miller, 2005) was designed to promote the development of understanding of data analysis processes and statistical concepts (Konold, 2002). Its development was funded by the National Science Foundation grant ESI-9818946. *TinkerPlots* was designed specifically for middle school students and addressed suggestions by Biehler (1997) that educational software tools should become more sophisticated as the user gains more expertise. It was designed as a construction set to enable students to build intuitive ideas about distributions and explore the potential of different graphical representations to answer questions about data. The purpose of *TinkerPlots* was to provide a means of exploring data in the spirit of exploratory data analysis, as intended by Tukey (1977). A fundamental underpinning idea was that *TinkerPlots* should build on what students know and are inclined to do when organising data displays and reasoning about data (Konold, 2007). Cliff Konold was the principle developer of *TinkerPlots* with expertise in curriculum and educational software development. Craig Miller was the senior software engineer with expertise in user interface design who “worked with researchers to create a finished product from rough ideas” (http://www.keypress.com/x2909.xml).

Prior to the development of *TinkerPlots,* Konold co-designed *DataScope* (Konold & Miller, 1994). *DataScope* was a data analysis program with accompanying data sets and instructional activities for Years 7-12. It was designed to support students using statistical techniques when analysing data about issues of interest to them. A spreadsheet format was used to enter data: columns represented attributes and rows represented cases or values of the attributes. *DataScope* was designed to be computationally powerful yet simple to use (http://srri.umass.edu/data_scope). This view is supported by researchers of issues related to data handling (Garman, 1996; Shaughnessy et al., 1996). In a review of *DataScope,* Garman (1996) commented,

> I assessed DataScope by trying to do as much as I could without reading the user guide. I was able to discover nearly all the features this way. After completely reading the user guide, I picked up several neat ideas and shortcuts. (p. 359)

Garman felt that the program was not extensive enough to be used for many college courses and commented on the limited printing capabilities of the program. He recommended the program be used for basic concepts of statistics and dynamic data analysis for middle and high school projects.

The displays in *DataScope* are limited to histograms, box plots, scatterplots, one and two-way tables of frequencies, and tables of descriptive statistics. These displays are akin to the EDA methods and techniques developed by Tukey (1977). Konold (2007) contended that by limiting the students’ choices to a set number of data displays more time can be spent on learning underlying concepts and data inquiry skills. To some extent this was realised but Konold found that students selected from the choice of data displays without understanding what the displays showed in terms of the questions explored. He also found that the students used a trial-and-error strategy in the hope one of the displays proved to be useful.

Konold's observations about how students used pre-established graph types when using *DataScope* (Konold, 2007) are echoed in research into how students reason about data (Konold & Higgins, 2003). The research indicates that students should begin with graphs in which they can retrace each individual data value and should also have the opportunity to construct graphs from data using "bottom-up software" in preference to using "top-down" software that produces pre-established graphical representations, such as histograms, at the click of a button. Konold then used his understanding of how the *DataScope* learning environment influenced the way students worked with data to guide the development of *TinkerPlots*.

Konold (2007) credits the development of *TinkerPlots* to three main motivating factors: first, the expansion of statistical practice to apply "data analysis" as a means of doing "statistics;" second, the introduction of data analysis into the curriculum from the early years of schooling; and finally, the growing availability of computer tools for the visualisation of data. Accompanying these factors, Konold's belief is that educational tools should address the way in which students learn in those environments.

As well as the lessons learnt from the development of *DataScope*, Konold's thinking about how to design an effective data analysis tool for students was influenced heavily by Biehler (1997), who identified that data analysis tools available at the time were stripped down versions of professional statistics software (Konold, 2007). These software programs included collections of conventional basic graph types, such as histograms, line graphs, scatterplots, and pie charts. When initiated, the programs performed the construction of the nominated graph type selected and generated the representation for the user. Konold wanted students to maintain control of the construction process and designed *TinkerPlots* to allow students to build the representations for themselves.

Biehler (1997) suggested that educational software tools should become more sophisticated as the user gains expertise, in much the same way as complexity is incorporated into computer games. He also made suggestions about how to make the complexity of a data analysis tool manageable, that is, to build the tool around a conceptual structure and include elements in the software that could be drawn on piece by piece when needed. Konold embraced these ideas and built *TinkerPlots* around a conceptual structure that allowed "students [to] manipulate case objects in a plot window using operations similar to those they would use if they were organizing physical objects on a flat surface: separating them into groups, ordering them, stacking them" (Harradine & Konold, 2006, p. 1). As well as the basic functions of stack, order, and separate, Konold and Miller (2005) included features such as the reference line, dividers, measures of centre, and the hat plot. *TinkerPlots* can be distinguished from professional graphing tools as the software package has many operators and features that can be combined collectively but can be accessed individually to build graphical representations from the bottom up. "It is this ability to combine operators in *TinkerPlots* that makes it complex, and powerful" (Konold, 2007, p. 279).

Worthy of particular attention is the hat plot representation. The hat plot is a new type of display introduced in *TinkerPlots*. It was developed from the box-and-whisker plot created by Tukey (1977). Konold (2007) realised that the box-and-whisker plot does not reflect the intuitive ideas students have about distributions and as Watson, et al. (2008, p. 6) noted:

> Although a box-and-whisker plot is often used to represent data, researchers have found that students do not find it easy to interpret because of the inverse relationship between the size of the four sections of the plot and the spread or density of the values represented.

Konold (2007) struggled to find a way of building up to a box-and-whisker plot from the representations possible in *TinkerPlots* but used his knowledge of how students intuitively look at clumps of data to design the hat plot.

Konold and Higgins (2003) determined that students have a tendency to summarise univariate data using centre or modal clumps of data. He used this notion and insights gained from research he and others conducted (see Konold et al., 2002), to design the hat plot so that the crown of the hat covers the middle 50% of the data and the brim of the hat extends across the range of the data (Konold, 2007). The brim on either side of the crown represents 25% of the data. Konold increased the complexity and flexibility of the hat plot by allowing the range of the crown to be changed by choosing settings other than the default settings or by dragging the crown edges once the plot has appeared, thereby changing the proportion of the data covered by the brim on either side of the crown as well.

In addition to *TinkerPlots* and *DataScope*, Konold and Miller (1994) designed *ProbSim: A probability simulation program. ProbSim* was designed for teaching probability via simulations in Years 7-12. To model a probabilistic situation the program constructed a "mixer" that contained the elementary events of interest from which random samples are taken after selecting replacement options, sample size, and number of repetitions to search for or count specified events of interest in that and subsequent samples (http://srri.umass.edu/prob_sim). Ernie (1996) trialed the program with high school and college students and was impressed by the way the open-ended modelling system supported students when learning about sampling. Although other simulators restrict probability modelling to one type of artifact, such as coins or dice, *ProbSim* offered a generic modelling system where the collection was defined by the user. In addition, Ernie was impressed with the program's capability of handling large samples.

Watson (2007) used *ProbSim* with Year 6 and 9 students to investigate their understanding of sampling and probability. She found that the program was successful in challenging students' thinking and assisted in "providing repeated 'experimentation' that would be impossibly tedious to complete in the classroom" (p. 73). Watson contended that the ability to experiment repeatedly with *ProbSim* by "increasing sample size and viewing distributions drawn on the same scale [was] likely to consolidate understanding" (p. 73). She recommended that experiences with simulators, such as *ProbSim*, be embedded in classroom discussion to help students make connections to mathematical thinking in order to promote long lasting understanding.

The fundamental ideas of *ProbSim* are incorporated in the second version of *TinkerPlots*, with the inclusion of a probability simulator – the "Sampler" (Konold & Miller, 2011). The inclusion of a probability simulator in *TinkerPlots* adds an extra level of functionality to the software. Probability simulations are constructed by selecting basic features and then combining them with other features to increase the capability and the complexity of the simulation, similar to the way in which graphical representations are built upon in *TinkerPlots*. The developers applied the overarching conceptual structure of *TinkerPlots* to the new probability simulator.

Application of *TinkerPlots*

According to Bakker's (2002) definition, *TinkerPlots* is designed as landscape-type software. In contrast to route-type software, landscape-type software does not assume a particular route of learning. Bakker suggests that *TinkerPlots* "offers students many possibilities for making their own, often unconventional graphs … and does not assume a particular learning trajectory" (p. 2) for the development of statistical concepts. In this way, *TinkerPlots* provides many different avenues for developing particular learning outcomes. Bakker also draws attention to the possibility that teachers may find it difficult to know *TinkerPlots* well enough to support students to answer statistical questions in a variety of ways. This is a valid concern but reports on the application of *TinkerPlots* indicate that teachers are designing learning activities that exploit the openness of *TinkerPlots* to engage students in meaningful data explorations (MacKinnon, 2009; P. Tabart, personal communication, April 19, 2011; Watson, 2008a; Watson et al., 2008).

The MARBLE project, with which this inquiry was associated provided professional learning for teachers on the use of *TinkerPlots* (Brown, Watson, & Wright, 2011; Watson & Beswick, 2009; Watson et al., 2008; Watson, Skalicky, Fitzallen, & Wright, 2009; Watson & Wright, 2008). One of the aims of the project was to build teachers' knowledge about data analysis in order for them to gain the confidence necessary to incorporate data analysis activities in their classrooms. To facilitate this, the researchers in the project designed learning activities for students and modelled how they could be used in the classroom with the teachers. From the project many teachers translated the activities into their teaching programs, some of which are reported in Brown et al. (2011), Watson et al. (2009), and Watson et al. (2008). The following examples of these activities and others developed by classroom teachers outside of the MARBLE project illustrate the usefulness and scope of *TinkerPlots*.

The development of data analysis skills and knowledge, more specifically variation, was the focus of activities designed by Watson and Wright (2008) for the MARBLE project.

In the activities teachers measured arm spans and used the data to explore notions of: measuring accurately, measuring arm spans of a group, comparing measurements of two groups, and comparing measurements on two variables. Variation in the data was explored using stacked dot plots, whilst scatterplots and split stacked dot plots were used for comparing measurements and the data collected from groups. Watson and Wright used the capacity within *TinkerPlots* to apply formulas and make calculations to create new variables. The teachers calculated and plotted differences between height and arm span measures and the ratio of height to arm span measures to explore the variation in the data beyond basic comparisons. The activities reflected the content of the Measurement Strand of the Australian Curriculum – Mathematics (Australian Curriculum, Assessment and Reporting Authority [ACARA], 2011).

The production of licorice was the context for an activity that focused on within and between group variation (Watson et al., 2009). In this activity the production of licorice was modelled using children's *Play-Doh Fun Factory* devices (www.hasbro.com/en_AU/). The weight of the licorice pieces made with the *Play-Doh* factories was compared with the weight of hand-made *Play-Doh* pieces and real licorice produced commercially. Stacked dot plots and hat plots were used to make comparisons and display variation within and among data sets. In addition, the mean was used to identify trends in the data.

Data reported in a newspaper article on the percentage decrease of salt content in margarine and breakfast cereals was used by Watson and Beswick (2009) to design an activity that provided the opportunity to develop quantitative literacy skills and knowledge. The newspaper article set the context for exploring further the salt levels of a variety of products. *TinkerPlots* was used to investigate the variation of salt levels from regular and salt-reduced products sourced from supermarkets. A particular focus in the activity was the investigation of the effect of removing outliers from a data set.

The connection between the mathematics and its application in real contexts was the focus of a series of activities that started with approximating Pi from measurements of round objects (Brown et al., 2011). This was complemented with investigating approximations of Pi using Count Buffon's strategy and the Monte Carlo method. The activity was extended to the application of Pi in an investigation that explored the relationship between the physical characteristics of elephants. Pi was useful in determining the size of an elephant's foot, which was then used for other calculations. *TinkerPlots* played a role in looking at the three approximations of Pi. The reliability of the approximations was determined by comparing data sets with only ten measurements and then with 30 measurements for each of the three approximation methods.

The merits of hat plots, their relationship to box-and-whisker plots, the links to percentages, and their application for reasoning about data were the focus of an article titled *The Representational Value of Hats* (Watson et al., 2008). In the article a number of data sets were used to show how hat plots could be used to compare data sets and describe modal clusters in data. Hat plots were also used extensively in an activity that questioned the claims made in a newspaper article that reported people with brown eyes had quicker reaction times than blue-eyed people (Watson, 2008a). As well as eye colour, gender difference was explored. This article highlights the possibilities for data analysis to be used in curriculum areas other than mathematics. The context of the eye colour/reaction time investigation is relevant to both Science and Physical Education subjects and provides the opportunity for cross-curricula explorations.

The use of survey data to make statements about a larger population was the focus of activities developed by Rubin (2005). Her emphasis was on the importance of phrasing questions to stimulate students to explore the data from a general perspective and not just state the results for the people the data were collected from. Rubin recommended young students use *TinkerPlots* to explore survey data by using the dividers in *TinkerPlots* to highlight main clumps in the data displayed in split stacked dot plots. She suggested that this assisted students to shift their thinking from a local to a global perspective. Friel, O'Connor and Mamer (2006) also divided data displays into sections but used the reference lines to "slice" (section off) the stacked dot plot distribution into bins. Their emphasis was on the use of *TinkerPlots* to facilitate meaningful, challenging investigations to extend students' thinking about variation, measures of centre, and distribution.

An unusual application of *TinkerPlots* was designed by MacKinnon (2009). He used the calculation facility in *TinkerPlots* to generate data that represented Pascal's triangle. Subsequent plotting of the data produced a triangle, reminiscent of more traditional Pascal's triangle displays. Other activities involving number patterns in the 100 chart, factors, and number properties are suggested in the *TinkerPlots Workshop Guide* (Key Curriculum Press, 2007).

Data collected from the Australian War Memorial database that includes the number of casualties, causes of injury, and details about imprisonment of Australian Infantry Forces during World War 1 provided an engaging context for Year 8 and 10 students. The timing of the investigation to coincide with Anzac Day celebrations made the investigation relevant for the students and made the connections between the context and the mathematics more meaningful (P. Tabart, personal communication, April 19, 2011). Figures 2.11a and 11b show the type of data collected and saved in *TinkerPlots*, some of the questions that the students considered, and three of the plots that were created for or by the classes. The series of questions explored by students led to making informal inferences about the influence of the weather on the number of casualties at particular times over the duration of the military campaign. Informal inference involves making generalisations about populations that go beyond the data (Makar & Rubin, 2009).

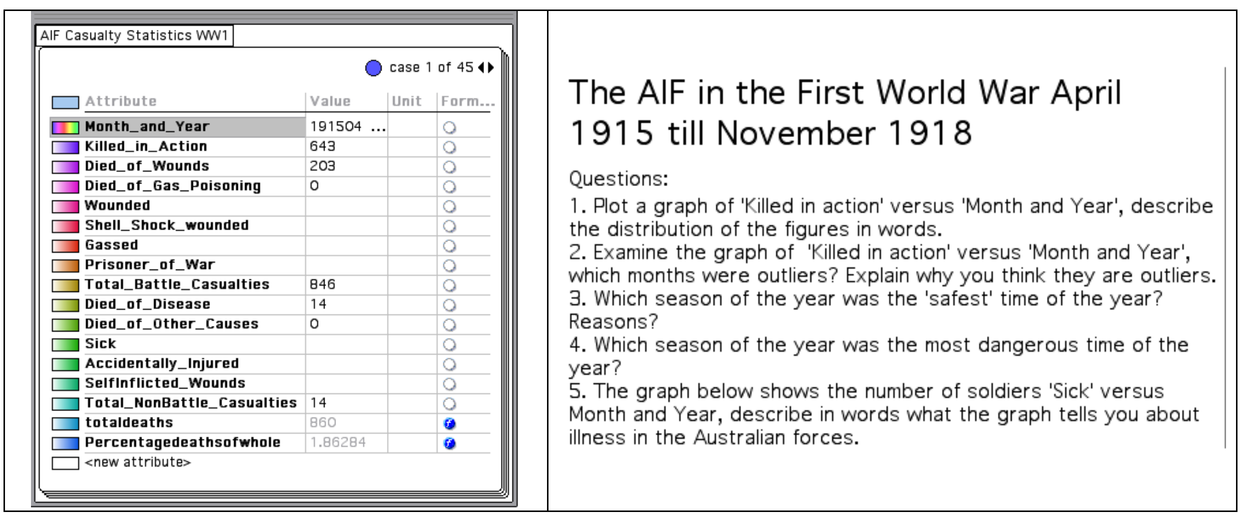
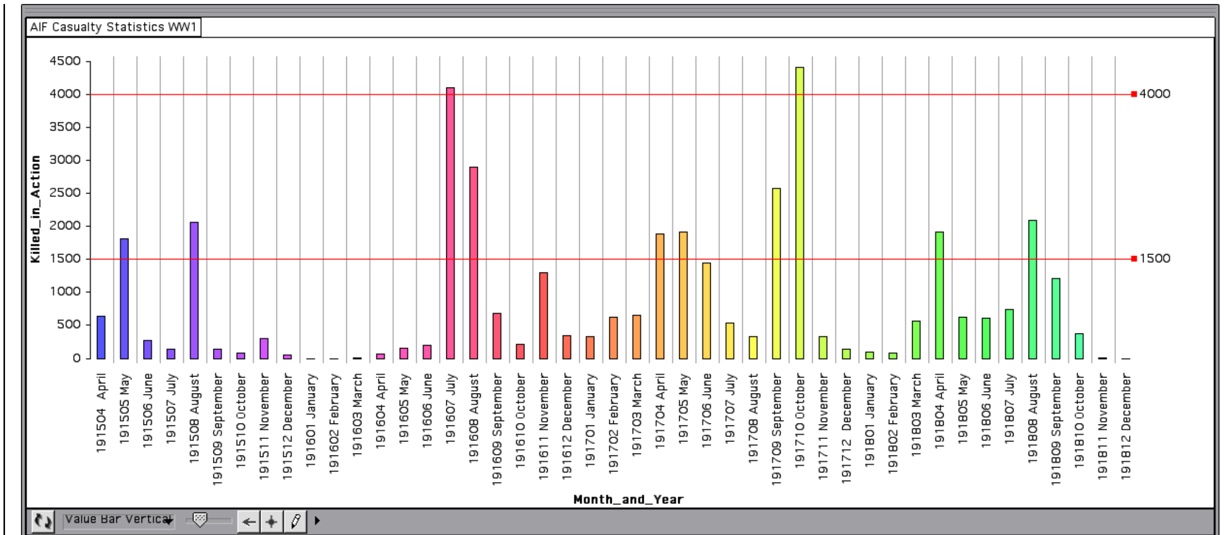
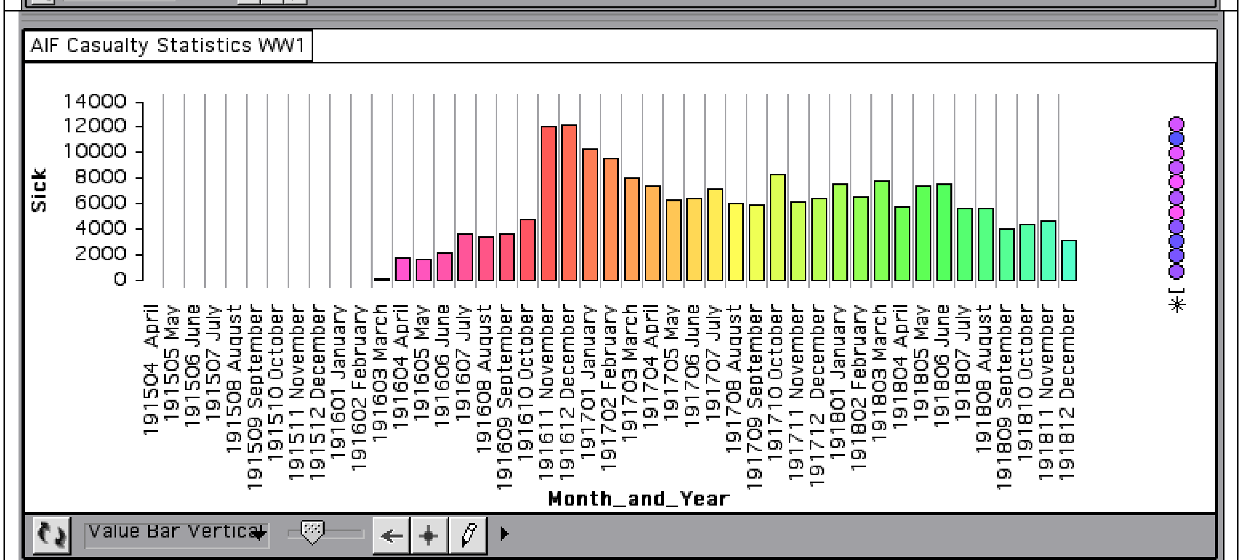
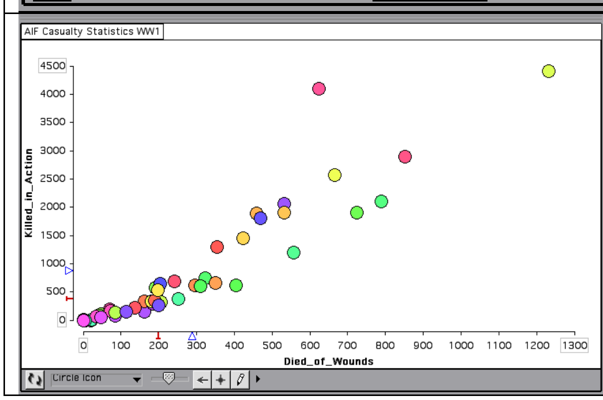


Figure 2.11a. Data analysis activities associated with Anzac Day using *TinkerPlots*.







Problems to answer

1. Create a plot of Killed in Action. Use this to determine the following:
a) The mean number of soldiers killed in action per month.
b) The median number of soldiers killed in action per month.

2. Create a plot of Gassed against Month and Year. What conjectures can you make by looking at this graph? Give reasons.

3. Create two plots; one of self inflicted wounds versus month and year and the second plot of totaldeaths versus month and year. Examine both graphs side by side. What conjectures can you make in relation to the occurrance of self-inflicted wounds? Give reasons.

Figure 2.11b. Data analysis activities associated with Anzac Day using *TinkerPlots*.

Summary of Application of TinkerPlots

The sample of learning activities presented in this section demonstrates the flexibility of *TinkerPlots.* The activities draw on a variety of contexts to enhance the learning of fundamental statistical concepts, such as variation. They also take advantage of the functionality of *TinkerPlots* to provide representations most appropriate for the learning of the statistical concepts targeted. The contexts of the activities also provide the opportunity for cross-curricula explorations. Many data sets, lesson plans, and teaching ideas suitable for students in the middle years of schooling are published in *Digging into Australian Data with TinkerPlots* (Watson, Beswick, Brown, Callingham, Muir, & Wright, 2011). The introductory learning activities developed provide the opportunity for students to learn about the fundamentals of data analysis and the basic features of *TinkerPlots.* The other activities are sequenced to include learning outcomes across the Statistics strand of the mathematics curriculum (ACARA, 2011) and encompass applications of most of the features in *TinkerPlots*. A number of different contexts are used throughout the activities to engage students in meaningful statistical investigations, which have the potential to extend to cross-curricular learning (Fitzallen & Watson, 2011).

Research with *TinkerPlots*

Similar to the applications of *TinkerPlots,* the research with *TinkerPlots* utilises a variety of contexts to engage learners in statistical investigations and address the learning needs of students. In addition, many research projects focus on supporting teachers to develop an understanding of how to use *TinkerPlots* as well as increase their knowledge of statistical concepts.

The research conducted by Hall (2008) with teachers found that using *TinkerPlots* in conjunction with Canadian Census at School data (www.censusatschool.ca) provided an effective avenue for supporting teachers who thought statistics was scary and intimidating. The teachers commented that *TinkerPlots* was user-friendly and intuitive in structure and Hall reported that combining the use of *TinkerPlots* with the context-rich data from the Canadian Census at School website assisted the teachers to develop positive attitudes towards teaching statistics and learning with technology. Fitzallen and Watson (2011) also reported on the benefits of selecting data with meaningful contexts for students to explore and recommended that:

> …discussion of context needs to be developed from two standpoints: that of interpreting contextual information specifically available within the graph itself (e.g., trends observed) and that of the extra real-world understanding brought to the data set by the students' life experiences. (p. 259)

The Fitzallen and Watson study of 12 Year 5/6 students reveals that students' understanding of the context influences the conclusions they make about the data. Sometimes there are direct connections made to the characteristics of graphs. Other times, the inferences made are based on the students' understanding of the context only.

Hammerman and Rubin (2003) described how teachers were able to see and manipulate data in a way that was new to them when given the opportunity to explore data with *TinkerPlots*. The teachers identified particular points, noticed clumps, and described gaps in the data. In the Hammerman and Rubin study, *TinkerPlots* offered an environment where many different graphs were generated and compared quickly, far more than possible when using pen-and-paper. The particular advantage of using *TinkerPlots* was its ability to construct data in bins of different sizes and to display visual and numerical information simultaneously (Rubin, Hammerman, & Konold, 2006).

In their work with student teachers, Monteiro, Asseker, Carvalho, and Campos (2010) reported that the ability to make changes to representations in *TinkerPlots* facilitated the shift in the student teachers' thinking. They suggested that this allowed the student teachers' interpretation of graphs to move from local to global perspectives. Monteiro et al. applauded *TinkerPlots* for its potential as a pedagogical tool but also emphasised that more research was needed into how it facilitates learning.

Ben-Zvi (2006) reported on a study that explored the development of informal inference and argumentation of 75 Year 5 students using *TinkerPlots* in open-ended data exploration tasks. Although some of the students were challenged by the openness of the tasks and required additional assistance to work through what was required, all the students were able to use *TinkerPlots* fluently and approached the work with enthusiasm. Ben-Zvi found that *TinkerPlots* both supported and hindered students' reasoning about informal inference. He commented:

> For example, one clear advantage was students' use of the software as an argumentative tool in presenting their ideas to others instead of just a representation tool. On the other hand the emphasis on dots (representing cases) in *TinkerPlots* focused attention on individual views of data and hindered spontaneous progress in some cases to aggregate views of data.

The ability to focus on individual cases was found by Ben-Zvi to allow some students to develop a deep knowledge and understanding of how to interpret graphs and identify trends in the data. This facilitated some students to shift their thinking to move beyond looking at individual cases.

Fitzallen and Watson (2010) found that the ability to click on a data point and read the data values on the case card helped students to confirm their hypotheses. This supported the students to be able to move back and forth between hypothesis creation and graph creation to make sense of the data. In their study, Fitzallen and Watson interviewed 12 Year 5/6 students one month after they had participated in data analysis activities with the rest of the class (n=26) as part of their regular learning program. The results showed that although a limited number of graphical representations were explored during the classroom activities, in the interview the students were not constrained by that experience and extended their use of *TinkerPlots* to create and reason about a larger range of data representations.

Gaining insights into how students develop an understanding of informal inference was also the goal of Paparistodemou and Meletiou-Mavrotheris (2008) when they worked with 22 Year 2 students using *TinkerPlots*. They noted:

> Attributes of TinkerPlots like the ability to operate quickly and accurately, to dynamically link multiple representations, to provide immediate feedback, and to transform an entire representation into a manipulable object enhanced students' flexibility in using representations and provided the means for them to focus on statistical conceptual understanding. (p. 101)

Paparistodemou and Meletiou-Mavrotheris also contended that *TinkerPlots* enhanced the students' learning by making inferential reasoning accessible to the young students. They suggested that although "the children in the study were young, most of the time they tried to find relationships between two variables in the data in order to draw their conclusions" (p. 101).

In a study, Visualizing Statistical Relationships (VISOR), Hammerman and Rubin (2003) worked extensively with middle and high school teachers and students to study how people learn about data analysis and how computer visualisation tools enhance that learning. The data analysis tools *TinkerPlots* and *Fathom*: *Dynamic Data Software* (Key Curriculum Press, 2005), were used in the professional learning sessions offered in the project. From the sessions with *TinkerPlots*, Hammerman and Rubin found that teachers used several methods to view and manipulate data, which include: i) grouping data into bins; ii) using percents to compare groups; iii) focusing on the trend in the data to reduce variation; and iv) focusing on the variation only. Through the process of reducing data into bins or using cut off points to look at aggregates of data, the teachers in the VISOR study created many different data representations that focused on different aspects of the data. They also grounded their analysis about the data in what they knew about the context.

When the researchers in the VISOR project explored how distributional shape influenced how teachers analysed data, they gave the teachers access to both *TinkerPlots* and *Fathom* (Hammerman & Rubin, 2004). Hammerman and Rubin reported that the teachers used both software packages in much the same way but took advantage of the additional features of *TinkerPlots* in their analysis. Although it was possible to construct box-and-whisker plots with both software packages, the teachers did not use them when using *TinkerPlots*. They used the moveable dividers to indicate cut off points and slices of data to make comparisons between two groups. Hammerman and Rubin suggested that the embedded features of different software packages support the creation of different ways of looking at the data. They indicated that this requires further interrogation to elaborate on fully.

To support the development of informal inference for Year 7 students, Watson and Donne (2008) replaced the box-and-whisker plot with the hat plot representation, along with the full range of tools available in *TinkerPlots*. When given the freedom to choose, not all the students continued to use the hat plot when it was not explicitly included in the interview protocol. Like the teachers in the Hammerman and Rubin (2004) study, those students selected the features of *TinkerPlots* that they decided were useful, which included splitting data in bins and using reference lines to highlight the value of specific data points.

In another part of the VISOR study, the researchers explored the way in which teachers analyse a data set using pen-and-paper, and then with *TinkerPlots*, to explore the effect of interactive data visualization (Rubin, 2002). The teachers in the study identified outliers in the data in both contexts but used different strategies in each context to do so. The data set was related to lifespan of brands of batteries. With pen-and paper, the teachers calculated the numerical mean for each brand of battery and then compared the brands using the mean. Instead of considering if the difference between the means was significant, as expected, the teachers focused on the battery with the smallest lifespan and discussions proceeded to focus on what difference it would make to the mean of the data set if the outlier was removed. Following on from this the teachers used *TinkerPlots* to construct a univariate plot with the attribute "lifespan" split into six bins. From this representation, the teachers decided that there were four outliers as those data points were separated from the modal clump and argued that all four outliers should be deleted. Rubin suggested that *TinkerPlots* allowed the teachers to look at the data differently and base their conclusions on what they considered to be more convincing evidence. The research indicates that the ability to sort data into bins in *TinkerPlots* may have unexpected influences on users' thinking, in particular, their identification of outliers.

Students' development of informal inference was the focus of a study with 15 Year 7 students (Watson, 2008b). Watson found that *TinkerPlots* was a valuable resource for introducing beginning inference to middle school students. In particular, the hat plot representation proved very useful as it allowed students to explore if there were small or large differences between two data sets. The ease with which multiple representations were constructed in *TinkerPlots* also assisted in supporting the students' developing intuitions about informal inference.

In 2009, Watson and Donne compared the graphical displays produced by students from a pen-and-paper based study with the displays created by students using the same data sets in *TinkerPlots*. They determined that *TinkerPlots* assisted students in thinking about the data sets but also noted the affordances of *TinkerPlots* that assisted the exploration of student understanding. These affordances were put into three categories: flexibility of representation, speed of analysis, and exposure of levels of understanding. The findings from the Watson and Donne study confirm the other research that described the flexibility of *TinkerPlots* to allow students to create their own representations and the way in which it facilitates the construction of many graphical representations quickly to challenge and support student thinking. The findings also go further to describe how the different representations available in *TinkerPlots* can assist in determining the students' level of thinking and expose students' misconceptions.

Summary of Research Using TinkerPlots

Research projects that utilised *TinkerPlots* provide evidence of the versatility and flexibility of the software, both as a learning tool and as a research tool. At times, the reported outcomes provide a description of how *TinkerPlots* can be utilised to develop an understanding of particular statistical concepts, such as informal inference, distribution, or variation. Other reports provide a critique of how *TinkerPlots* influences users' thinking and consider the potential positive and negative aspects in relation to the learning objectives. Studies that employed *TinkerPlots* as a research tool looked specifically at how the data visualisation software enhances the engagement of the user in data analysis activities. In these studies, the learning objectives in terms of the statistical concepts explored were not the main goal but served as a vehicle for examining the way in which the software assists students to work statistically and reason about data. Other studies focused on student learning or used *TinkerPlots* in professional learning programs for teachers. In most instances, the teachers in the studies were novice users of *TinkerPlots* and very often beginners in the application of EDA strategies. Rubin and Hammerman (2006) report outcomes for teachers to be the similar to those for students when conducting an analysis of data *TinkerPlots*. Therefore, the insights gained from how teachers learn when using *TinkerPlots* are useful when considering the ways in which students learn and vice versa.

Concluding Remarks

This stage of the inquiry draws on a model of statistical thinking and reasoning developed for the inquiry in Stage 1 (Figure 1.7), a framework for evaluating ILS resources (Kidman & Nason, 2000), and a set of practical guidelines for evaluating educational software (Goyne et al., 2000) to develop a set of criteria for the evaluation of EDA software. The set of *Criteria for Evaluating Exemplary EDA Software* (Figure 2.2) developed was then used to evaluate *TinkerPlots*. It was determined that *TinkerPlots* provides the opportunity for students to be involved actively in the data analysis process. This was evidenced when a student participating in a research project investigating students' development of key statistical concepts – comparing distributions and comparing samples – explained *TinkerPlots* in the following way:

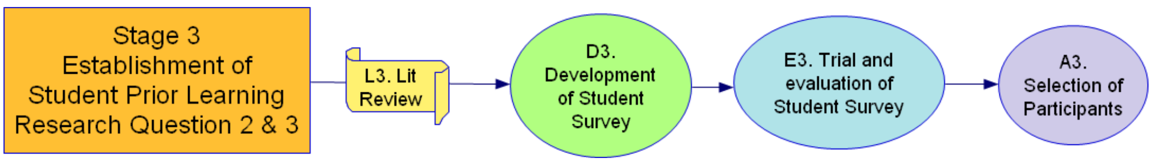
> It is sort of like a graphing sort of system that can be used to figure out that you can make graphs out of and actually, it is pretty fun, too. It is sort of like, … it's got the same sort of things as any other sort of graphing thing but you actually make the graph yourself instead of entering all the numbers in and the computer does it for you. (Bakker & Fredrickson, 2005, p. 88)

The applications of *TinkerPlots* demonstrate that the software offers a dynamic and effective learning environment for students. Its use in many different contexts and for many different purposes shows that *TinkerPlots* has obvious advantages as an educational tool as it allows students to create representations that are meaningful to them. It is, therefore, pertinent to explore the ways in which students engage with the software and use it to analyse, display, and interpret data. In addition, the value of *TinkerPlots* as a research tool is exemplified in the research studies reviewed. The studies confirm its applicability in this inquiry to explore students' thinking when conducting data analyses.

The sequence of learning experiences developed in Stage 4 of the inquiry provides the opportunity for students to use *TinkerPlots* to develop an understanding of covariation. The outcomes from that stage of the inquiry are used to answer the research questions in Stage 5. In the next stage, the students' prior learning is established in order to inform the development of the sequence of learning experiences.

Stage 3

Establishment of Student Prior Learning



The purpose of Stage 3 of the inquiry was to ascertain the prior learning of the students in the study in relation to their understanding of covariation and experience with creating and interpreting graphs. The aim was to develop a Student Survey that included research validated assessment items from previous statistics education research, where possible, in order to inform the development of a teaching intervention in Stage 4. The results from the Student Survey were also used to select the group of students who participated in the implementation of the teaching intervention.

The research process for educational design experiments that involve the implementation of a teaching intervention should include identification of the significant mathematical ideas that are the targeted learning outcomes, the anticipated starting points in terms of the students' prior learning, and the envisioned sequence of learning experiences to support the development of the targeted learning outcomes (Cobb, McClain et al., 2003). This stage of the inquiry focuses on the second point, the identification of the students' prior learning. The identification of the mathematical ideas and connections to the curriculum are included in Stage 4, in conjunction with the development of the Sequence of Learning Experiences.

In this stage of the thesis the development of an assessment instrument for eliciting students' prior knowledge of graphing and data analysis is described. The method of implementation of the assessment instrument with 71 Year 4/5/6 students is also explained. The results from the administration of the Student Survey are presented. The results from this stage of the inquiry contribute to Research Questions 2 and 3.

Student Prior Learning

Contemporary assessment practices include formative evaluation that is used by teachers to guide and inform instructional practices when planning for particular learning outcomes (Earl, 2003). Black and Wiliam (1998) assert that formative evaluation is at the heart of effective teaching and should focus on how students learn. Wiliam and Black (1996) credit the current meaning of the term "formative evaluation" to Bloom, Hastings and Madaus, in 1971. They later proposed that formative evaluation was useful in helping teachers, students, and curriculum writers to improve and make progress in what they do and emphasised that formative evaluation is an assessment process and is not limited to specific assessment tasks (Bloom, Madaus, & Hastings, 1981).

Often students are given a formative evaluation task prior to instruction as a pre-test to determine what the students already know. By establishing students' prior learning, teachers are in a position to adjust tasks and strategies to ensure that the teaching and learning programs are responsive to student learning needs and build on prior knowledge and skills (Woolfolk & Margetts, 2007; Van de Walle, 2007). Having an understanding of students' prior knowledge provides a leverage point from which to extend and challenge established intuitions and facilitate future learning (Schwartz, Sears, & Chang, 2007).

There is widespread agreement that prior knowledge influences learning, and that learners construct concepts from prior knowledge (Roschelle, 1995). Ben-Zvi (2004) also found that although students' background knowledge influenced the conclusions they made, it was a base from which to work to refocus students' attention. He stated that students' prior knowledge was multifaceted, “sometimes hindering progress and at other times advancing knowledge in interesting ways” (p. 139).

Evidence of prior learning also provides insights into the misconceptions of students and can lead to strategies for challenging students’ thinking or supporting the shift to new ways of thinking (Smith, deSessa, & Roschelle, 1993-1994). Of paramount importance is the need to establish the diversity of understanding within a group of students and as Roschelle (1995) claims, this cannot be done effectively through a review of the literature. He maintains that eliciting prior knowledge of students is best done by working closely with students and exploring the way they interact with assessment tasks.

Development of Student Survey

Determining the prior learning of students presented many challenges for this inquiry (Smith et al., 1993-1994), particularly when the evidence was to inform the development of learning experiences in new and unexplored contexts, such as that offered by *TinkerPlots*. In the case of statistics education, there is the need to develop assessment instruments and materials that evaluate statistical thinking and reasoning in technology environments (Garfield & Ben-Zvi, 2004; Shaughnessy, 2007). It is, therefore, pertinent to ensure the features of software available for use in statistics education and the way students access and engage in data analysis are considered when designing assessment instruments. In this inquiry it was also important to consider that prior student learning may have been gained from traditional pen-and-paper experiences as well as from the engagement with graphing software technologies.

Another challenge for this study arose as the assessment instrument was administered away from the computer environment. This was necessary as the aim of using the assessment instrument was to determine students’ learning in terms of their ability to reason about data. Should the assessment instrument be administered as an exercise on the computer, there was the concern that the emphasis would be on the students’ ability to access the features of computer programs instead of the mathematics. Also, administering the assessment instrument on the computer may have disadvantaged students who were inexperienced in using graphing software. Although the intention was to administer the assessment instrument as a pen-and-paper activity, particular attention was given to the selection of items with formats that remained authentic to the *TinkerPlots* environment.

In order to determine the students' prior knowledge of data analysis and graphing, a student survey was developed to serve as an assessment instrument. The selection of the items for the student survey was based on the theoretical framework, *Model of Learning Behaviour in EDA Graphing Environments*, developed in Stage 1 (Figure 1.7). In designing the Student Survey it was also necessary to:

a) provide opportunities for students to demonstrate an understanding of statistical concepts at varying levels;
b) consider the range of data representations possible when using pen-and-paper and/or technology-rich environments; and
c) select items for the instrument from previous research conducted in statistics education.

Additionally, it was important to ensure that each of the key elements in the theoretical framework organisers of the *Model of Learning Behaviour* be addressed in multiple items of the assessment instrument. As Moritz (2003b) had done, the Student Survey was organised so that the graph creation tasks "were placed before the interpretation tasks to ensure exposure to the printed graphs in the interpretation tasks did not suggest a graphing method" (p. 234).

Student Survey Analysis

In order to provide opportunities for students to demonstrate an understanding of statistical concepts at varying levels, the Student Survey items included closed and open-ended questions. The Structure of Observed Learning Outcomes (SOLO) model developed by Biggs and Collis (1982) was used to analyse the results from the Student Survey. Application of the SOLO model (Biggs & Collis, 1982) was justified as "the levels of the Biggs and Collis learning cycle have provided a powerful theoretical base for situating research on students' statistical reasoning from the elementary schools through college" (Jones, Langrall, Mooney, & Thorton, 2004, p. 100) and has been used extensively in studying students' mathematical thinking and reasoning in statistics education (Jones et al., 2000; Jones, Langrall, Thorton, & Mogill, 1997; Mooney, 2002; Moritz, 2004; Reading &

Reid, 2006; Watson & Callingham, 2003). Although accepted widely, Shaughnessy (2007) contends that further debate is required on the merits and demerits of the SOLO model but concedes it has been useful for describing students' reasoning about statistical concepts.

The SOLO model is a structural hierarchy, which focuses on students' responses rather than their level of thinking or stage of development (Pegg & Tall, 2005). Its strength as a research tool "is that it provides a framework to enable a consistent interpretation of the structure and quality of responses from large numbers of students across a variety of learning environments in a number of subject and topic areas" (p. 469). At the first level, termed pre-structural (P), students' responses include irrelevant or incomplete information. At the second level, termed uni-structural (U), responses employ single elements to describe the relationship seen in the data. These are isolated statements that are not linked together to draw conclusions. At the third level, termed multi-structural (M), responses employ multiple elements to describe the relationship seen in the data. These are offered as a string of ideas that are supporting evidence about the relationship seen but are not presented in an interrelated fashion. At the fourth level, relational (R), responses employ multiple elements to describe and justify the relationships seen in the data and make explicit how the elements are interrelated (Watson & Callingham, 2003).

The SOLO model was used to develop coding schemas for analysing the students' responses to the questions in the items that were open-ended in nature and required written explanations. Questions that elicited responses that did not display a developmental progression but did display varying levels of complexity were coded using hierarchical frameworks developed for those questions. Other questions that elicited responses that were either correct or incorrect were coded accordingly. The coding schemas for each item are described as the results for the items are presented in the following Student Survey Items and Results section.

A description for the coding applied to each item is described as the results are presented in the following Student Survey Items and Results section. To strengthen the validity of the Student Survey results, a double coding procedure was used that was in keeping with the procedure as described by Miles and Huberman (1994) to ensure inter-rater reliability. Before the coding commenced, the researchers discussed the application of the coding schema developed and then applied the coding schema to five randomly selected student surveys. When each of the researchers was confident with the ability to distinguish between the levels of the coding schema the full data set was coded.

Following consultation between the researchers and the trialling of the coding schema the data were then independently coded. Where different levels of response according to the P-U-M-R sequence were provided by a student for a particular question, only the highest level of response was recorded by each of the researchers. After the coding of the full set of data was completed, the two researchers compared the two sets of coding. Where discrepancies were identified, the researchers discussed and explained their reasoning in order to reconcile differences in order to reach a consensus. The one instance where this did not occur the highest rating was selected.

Student Survey Participants

The Student Survey was administered to three classes of students aged 10-12 years of age – one Year 4/5/6 class (n=22) at a district high school (Years K-10) and two Year 5/6 classes (n=50) at a primary school (K-6). The schools from which these classes were selected were involved in the MARBLE project, which this inquiry was associated. All of the students in the three classes completed the paper-based Student Survey, which took approximately 45 minutes. Of the 72 students, one Year 4 student from the Year 4/5/6 class withdrew participation in the research project. The assessment instrument collected from that student was returned to the classroom teacher and subsequently returned to the student. No data from that student's survey were recorded or analysed.

Student Survey Items and Results

The Student Survey was comprised of seven items. The items included a variety of questions that required students to read data from a graph, make inferences from the data, create graphs, or describe variation. Some of the questions included were designed to gather information about students' previous experience with using software packages and ability to ask questions about data. One of the objectives was to ensure that each of the dimensions in the theoretical framework, *Model of Learning Behaviour in EDA Graphing Environments*

(Figure 1.7), was addressed in multiple items of the Student Survey. Table 3.1 indicates the connections between the theoretical framework and the Student Survey items. As intended each of the dimensions of the *Model of Learning Behaviour* is addressed by more than one item.

Table 3.1.
Summary of Model of Learning Behaviour in EDA Graphing Environments and Student Survey Assessment Instrument Items

Dimension	Student Survey Items						
	1	2	3	4	5	6	7
Generic knowledge	✓	✓	✓	✓	✓	✓	✓
Being creative with data		✓		✓			
Understanding data	✓	✓	✓	✓	✓	✓	✓
Thinking about data					✓		✓

Another objective was to source items for the Student Survey from those used in previous research in statistics education. Items 4-7 met this criterion. These items were developed for research that explored students' statistical thinking and reasoning skills and are noted for their ability to elicit information about students' thinking at varying levels of understanding. The items cover a range of data displays, which include a column graph, scatterplot, case cards, and spread sheet formats, all of which resemble the data displays possible in *TinkerPlots*. In addition, the items sourced from research studies were administered to students from Years 3-9. This indicates that the items are appropriate for the students in this inquiry. Items 1-3 were developed to gather information about students' general knowledge of graphing and previous experiences with using graphing technologies. In this section, each item in the Student Survey is described and the method of analysis is explained. These are then followed by the students' results and example responses for the questions in each item.

Item 1 – What is a graph?

The questions in Item 1 (Figure 3.1) were designed to provide the opportunity for the students to describe what they knew about graphs. The aim was to draw out the students' core ideas about what they thought graphs were and what they were used for. Jones et al. (2000) used a similar strategy when they asked primary students the open-ended question, "What does this picture show you?" together with a variety of graphs and diagrams. The purpose of the question in the Jones et al. study was to assess the limits of the students' thinking in relation to describing data. In this inquiry, no diagrams were provided as visual clues to ensure the responses were not influenced by any particular graph type.

Question 1.1	What is a graph?
Question 1.2.	What are graphs used for?
Question 1.3.	Where have you seen graphs used?

Figure 3.1. Item 1 of the Student Survey.

Question 1.1. What is a graph? The responses for Question 1.1 were coded according to the P-U-M-R sequence presented in The Concept of Graph developmental sequence developed by Watson and Fitzallen (2010, p. 59). At the uni-structural level, the students' responses or diagrams represent one of the elements of a graph, such as scale or variation, but do not connect them with the data they represent. At the multi-structural level, responses or diagrams employ multiple elements of graphs to create a diagram that shows there is a connection between the elements selected but is incomplete. At the relational stage, the responses or diagrams employ multiple elements of graphs to create a meaningful picture or diagram that tells the story of the data. Inappropriate responses or no responses are at the pre-structural level. A summary of the results for Question 1.1 is given in Table 3.2.

Table 3.2.
Percentage of Responses for the SOLO Levels for Question 1.1

Pre-structural	Uni-structural	Multi-structural	Relational
31.0 (n=22)	40.8 (n=29)	15.5 (n=11)	12.7 (n=9)

The open-ended nature of the question elicited a variety of responses from the students. For the most part the students wrote short statements to explain their thinking and some students supported their statements with drawings of graphs. For Item 1 the pre-structural responses were non-statistical. The students described graphs as being geometric shapes or grids. They also referred to them as measuring tools, and as pictures for displaying information. These descriptors were not accompanied by further elaboration. Student 4 wrote, "A graph is a row of lines in any direction" and Student 21 explained, "A graph is a measuring ruler, a bit like that." Five of the students drew pictures of grids with no extra information like the one drawn by Student 57 in Figure 3.2.

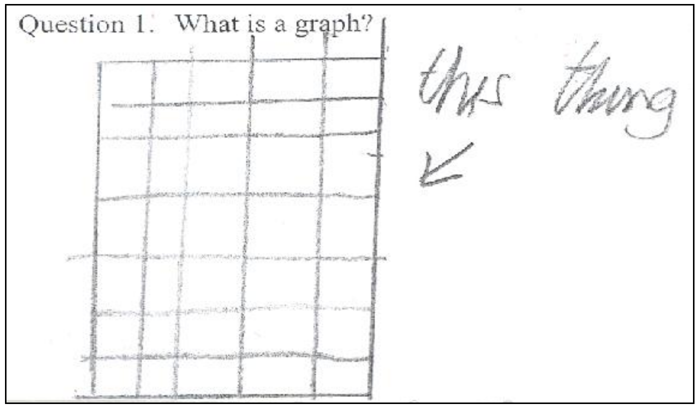


Figure 3.2. Pre-structural response to Question 1.1 drawn by Student 57.

At the uni-structural level the students' responses focused on the visual display generated by graphs and indicated that graphs had a purpose. Student 44 explained that a graph was "a diagram with a key to show information about something" and Student 24 explained "People use graphs to show information found out from surveys." Responses at the multi-structural level went further and acknowledged that the graphs served a specific purpose. Student 6 wrote, "A graph is a way of showing how many people like different things or different sports, colour etc." At the relational level the students' responses described graphs as being representations of data and gave examples of what they showed. Student 19's relational response is in Figure 3.3.

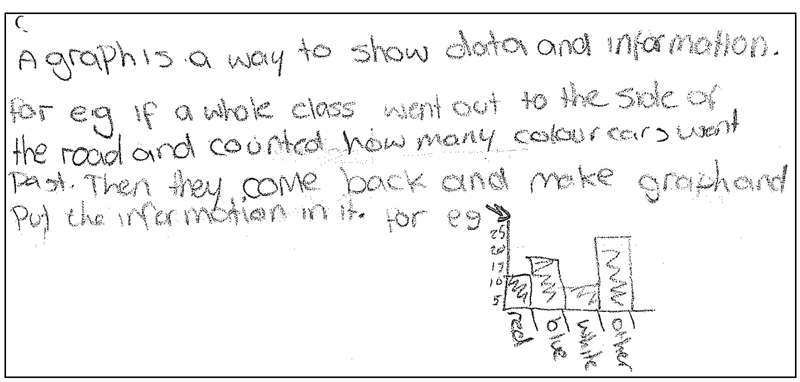


Figure 3.3. Relational response to Question 1.1 drawn by Student 19.

Question 1.2. What are graphs used for? The responses for Question 1.2 were coded, using the SOLO hierarchy. Corresponding to the SOLO model, uni-structural responses employ only one purpose for graphs and multi-structural responses employ multiple purposes of graphs. Responses at the relational level employ multiple purposes of graphs and make the connection between graph types and what they represent or make connections to the context of the graphs. A summary of the results for Question 1.2 is given in Table 3.3.

Table 3.3.
Percentage of Responses for the SOLO Levels for Question 1.2

Pre-structural	Uni-structural	Multi-structural	Relational
21.1 (n=15)	62.0 (n=44)	12.7 (n=9)	4.2 (n=3)

The students' responses that were coded at the pre-structural level focused on irrelevant information. Student 7, for example, suggested graphs were used "to make the world a better place; a better environment." The responses coded at the uni-structural level re-iterated some of the responses to Question 1.1. For the most part they involved describing graphs being used for surveys or as a way of making a tally or showing information. At the multi-structural level the students' responses provided more than one purpose for graphs without elaborating further. "To show information from a survey. Show how many people like something. To see your favourite thing," written by Student 14, was typical of the responses coded at the multi-structural level. The responses at the relational level went further and encompassed notions of data representation. The response written by Student 17 (Figure 3. 4) conveyed that he understood that graphs were used to represent data in different ways for different purposes. He also noted a meaningful context within which graphs are used.

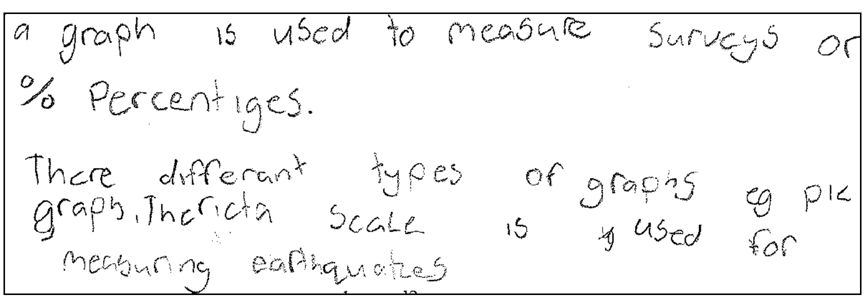

a graph is used to measure surveys or % Percentiges.
There different types of graphs eg pic graph, Thericta Scale is used for measuring eathquakes

Figure 3.4. Relational response to Question 1.2 written by Student 17.

Question 1.3. Where have you seen graphs used? This question was designed to gather information about the contexts within which students had encountered graphs. The students' responses were coded and grouped according to the content included in the responses. The groups of content were then clustered according to the common themes dominant in the students' responses to determine the categories of responses made. The dominant categories were School, TV, Books, Computer/Internet, and Applications. A summary of the results for Question 1.3 is given in Table 3.4.

Table 3.4.
Percentage of Responses for Dominant Categories Evident in the Reponses to Question 1.3

School	TV	Books	Computer/Internet	Applications
28.8 (n=34)	13.6 (n=16)	18. 6 (n=22)	23.7 (n=28)	15.3 (n=18)

"Who wants to be a millionaire?" and "Big Brother" were mentioned regularly in the responses coded for the category, TV. This highlights the influence popular culture plays in the development of ideas for young people. The responses coded for the category of School and Books provided little indication of what type of graphs were sighted or what they were used for. Student 40 elaborated a little when she noted, "I've also seen graphs in books. Non-fiction books of course." The category of Applications included responses that described the uses of graphs rather than where graphs had been encountered. Student 54 suggested, "heart beats" and "what colour hair" was suggested by Student 34. Some of the students made connections to meaningful contexts and indicated that the graphs were used for a purpose. One student wrote,

> Graphs are used on television programs such as who wants to be a millionaire, in a hospital to measure heart beats, measure earthquakes or what a person's favourites are. They can also be used as a way to present how many questions you get right in an exam or test. They are also used by nutrition doctors as a way to show the food pyramid. (Student 9)

Item 2 – Draw a graph

Following on from Item 1, the question in Item 2 (Figure 3.5) required students to draw a graph. No context or data were provided at this stage as, like Item 1, the aim was to elicit information that was generated solely from the students' thinking to ensure the responses were a demonstration of the students' fundamental ideas.

Item 2. Draw an example of a graph.

Any type will do. Put as much detail on the graph as possible.

Figure 3.5. Item 2 of the Student Survey.

Item 2 elicited information about the students' understanding of how data were organised in graphs. The graphs drawn by the students were predominantly column graphs and represented the frequency of attributes. Five students drew variations of pie charts and only one student drew a line series graph (Figure 3.6). Many of the graphs showed information related to familiar contexts such as the "number of pets" or "the colour of cars." A summary of the results for Item 2 is given in Table 3.5.

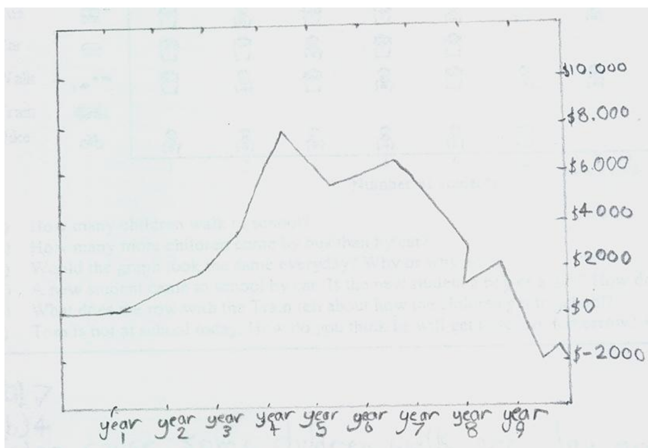


Figure 3.6. Time series graph drawn by Student 10.

Table 3.5.

Percentage of Responses for the SOLO Levels for Item 2

Pre-structural	Uni-structural	Multi-structural	Relational
7.0 (n=5)	9.9 (n=7)	25.4 (n=18)	57.7 (n=41)

Similar to the responses to Item 1, the student responses for this item at the pre-structural level were drawings of grids and lines that did not display data in any way. The uni-structural responses had indications of a scale and variation in the structure of the graph. Although some of the responses at this level had numbers on a scale or on the columns, often there were no connections made to the context of the data to give an indication of what the graph represented. The graph in Figure 3.7 drawn by Student 12 was coded at the uni-structural level. Although the graph is devoid of context and is an unconventional representation, it does show that the student had some understanding that graphs display variation.

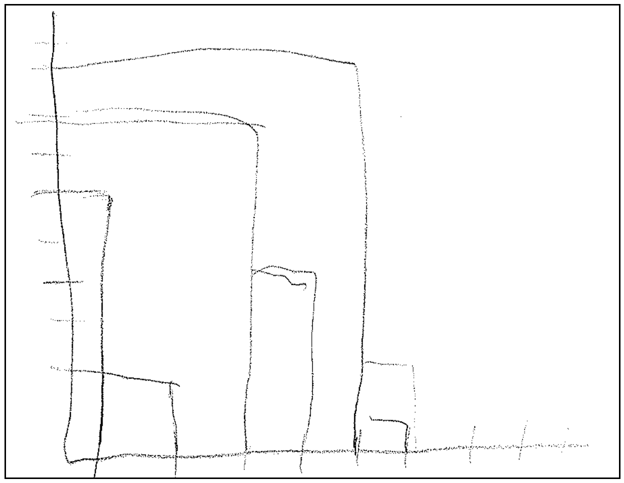


Figure 3.7. Uni-structural response to Item 2 drawn by Student 12.

At the multi-structural level the students' responses made connections to meaningful contexts and displayed variation in the data by varying the height of the columns. The graph in Figure 3.8 created by Student 58 is typical of the graphs coded at the multi-structural level.

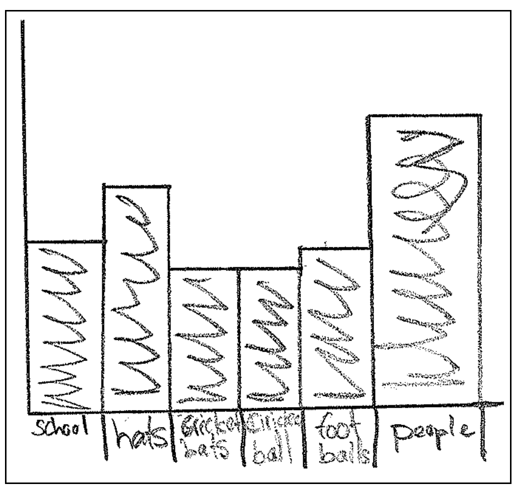


Figure 3.8. Multi-structural response to Item 2 drawn by Student 58.

Item 2 elicited more relational level responses than Item 1. At the relational level the data were related to a context, variation was evident in the structure of the graph, and additional information about the attributes, the purpose of the graph or interpretation of the graph were provided. Figure 3.9 shows a relational graph (Student 55), as does Figure 3.6.

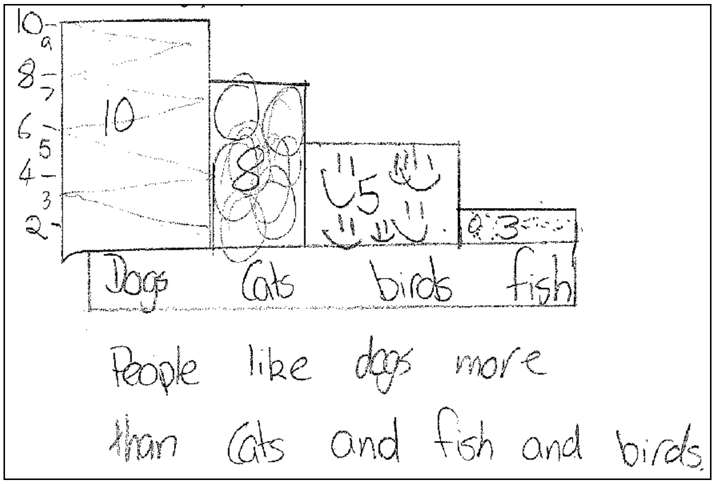


Figure 3.9. Relational response to Item 2 drawn by Student 55.

Item 3 – Use of Graphing Software

As the inquiry included exploring students' use of *TinkerPlots* it was important to establish if the students had prior experiences with educational graphing technologies. The series of questions in Item 3 was designed not only to find out if the students had used software packages to create graphs but also to explore their understanding of the purpose and context of creating the graphs. The questions in Item 3 are given in Figure 3.10. The responses to questions in Item 3 were sorted into categories relevant to each question and a frequency of occurrence for each category recorded.

Question 3.1.	Have you used the computer to draw graphs before?
Question 3.2.	What programs did you use?
Question 3.3.	Describe what sort of graphs you drew.
Question 3.4.	What were the graphs used to show?

Figure 3.10. Item 3 of the Student Survey.

Question 3.1. Have you used the computer to draw graphs before? The responses for Question 3.1 were sorted into "yes," "no," or "no response" categories. Forty-two percent of the students indicated that they had used the computer to draw graphs. This was exceeded by the percentage of students who indicated they had not used the computer for graphing purposes (53.5%). Three students did not respond to Question 3.1.

Question 3.2. What programs did you use? In the responses to Question 3.2, 15 students indicated that they had used *TinkerPlots* previously. This was not surprising as the classes that completed the Student Survey were drawn from schools that were given *TinkerPlots* as part of the MARBLE project. The other software packages that were mentioned in the students' responses were Publisher (n=4), Excel (n=3), Word (n=4), Microsoft (n=2) and KidPix (n=1). Among the responses, the students named software applications that are not usually used for graphing. The mix of software applications noted may be an indication that the students had not always paid attention to the software used. As the students named an application, the assumption was made that they had used the computer for graphing. Nine students wrote that they did not know what programs they had used and 46% of the students did not respond to Question 3.2.

Questions 3.3. Describe what sort of graphs you drew. Twenty-five percent of the students described the types of graphs they had drawn with the computer. The types of graphs named were bar graphs (n=4), pie charts (n=5), and graphs with squares or boxes (n=8). Six students responded, "don't know." Interestingly, all of these students had indicated that they had drawn graphs on the computer and had drawn a graph in response to Item 2.

The responses to Questions 3.3 highlighted the difficulty students have when they do not know the conventional terminology to articulate their thinking and knowledge. To describe the types of graphs drawn, many of the students used rudimentary language to describe the graphs. Student 40 wrote that she had drawn a graph with squares and went on to elaborate, "I drew a 1x1 graph before. But other times I've drew 1 by 4 or 1 by 5." This description suits a description of a table rather than a graph and it could be construed that the student is confused about what a graph may be. Another student wrote, "A square with dots and colours (Student 35). Similar to Student 40's response, this student's response did not articulate fully the conventional notion of a graph. She did, however, note that she had used *TinkerPlots* in her response to Question 3.2. It was, therefore, possible that her description of a square graph was describing the nature of a graph constructed in *TinkerPlots*. Among the responses, one student (Student 50) noted using a “school graph.” No additional information was given to determine the features that might characterise a school graph.

Questions 3.4. What were the graphs used to show? Only 30% of the students (n=21) responded to Question 3.4. Of those students, approximately 80% indicated that the graphs were used to show “How many…” of a particular category, such as “How many Smarties in a box?” as explained by Student 60 and “How many cars went past in an hour?” (Student 39). The only other responses explaining what graphs were used for were “A graph to show you what jobs I do around the house on weekdays” (Student 28), “I used it to show what I had to do for homework” (Student 45), “bicycle riding” (Student 61), and “Yes, doing it on cars” (Student 25).

Item 4 – Draw a Graph from Data

Item 4 was included in the Student Survey as it provided the opportunity for students to demonstrate the way in which they used data to construct graphs. The task, reproduced in Figure 3.11, was used by Moritz (2003b) to examine students’ approaches to creating graphs from numerical data. He conducted research that included 133 students from Years 3, 4/5, 7, and 9. The results from the research conducted by Moritz demonstrated conclusively that the assessment task was accessible across the full age range of the students and provided evidence of diversity of levels of understanding within each year of schooling.

A science class was studying temperature. They used a thermometer to measure the room temperature every 5 minutes for 30 minutes.

First they turned a heater on for 15 minutes.
Next they turned the heater off for 10 minutes.
Lastly they opened the window for 5 minutes.

They wrote down these numbers.

Time (Minutes)	5	10	15	20	25	30
Temperature (°C)	15	20	25	25	25	15

Draw a graph to show how the temperature changed over time.

Figure 3.11. Item 4 of the Student Survey (Reproduced from Moritz (2003b, p. 234)).

Moritz (2003b) coded the students’ responses to the task according to the degree to which the spatial features in the graphs constructed represented the four levels of the SOLO hierarchy for covariation developed for the task. At the pre-structural level responses do not show the data, but include information about the context of the data or provide non-statistical elements of a graph format, such as a blank set of axes. Uni-structural level responses represent at least one data series, either as a table of corresponding values or a graph of a single data series. Multi-structural level responses display both sets of values but do not use position in two dimensions to denote values of the two variables, either showing correspondence but lack an ordered scale to indicate variation, or showing variation but lack a direct correspondence of variables. At the relational level, responses use position in two dimensions to denote values of the two variables in bar graphs or line graphs (Moritz, p. 235). The P-U-M-R coding schema for this question developed by Moritz was used to code the responses to this item. A summary of the results for Item 4 is given in Table 3.6.

Table 3.6.
Percentage of Responses for the SOLO Levels for Item 4

Pre-structural	Uni-structural	Multi-structural	Relational
0 (n=0)	21.1 (n=15)	52.1 (n=37)	9.8 (n=7)

The graphs drawn by students in response to Item 4 were similar to those drawn by the students in the Moritz study (2003b) from which this item was sourced. In this inquiry the students responses were coded at the uni-structural level or greater. No responses were coded at the pre-structural level, however; 17% (n=12) of the students did not respond to Item 4. At the uni-structural level the graphs showed one variable with corresponding values from the table, as is evident in Student 43's response in Figure 3.12. Student 43 included the context of the information as a second attribute on the graph but did not indicate that the attribute was related to the variable time. The graph in Figure 3.13 drawn by Student 63 is typical of a response at the multi-structural level. At the multi-structural level the graphs showed both variables but often the data values lacked correspondence to the values in the table. At the relational level the graphs showed both variables and displayed the variation evident in the data or used coordinates to show discrete values of data. The graph in Figure 3.14 is an example of a graph coded at the relational level. The graph drawn shows the covariation between the two attributes and represents a conventional time series graph with appropriate scales on the axes.

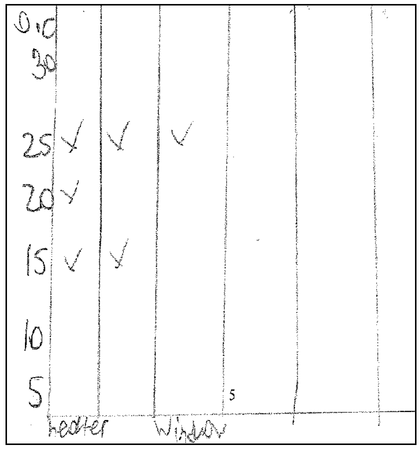


Figure 3.12. Uni-structural response to Item 4 drawn by Student 43.

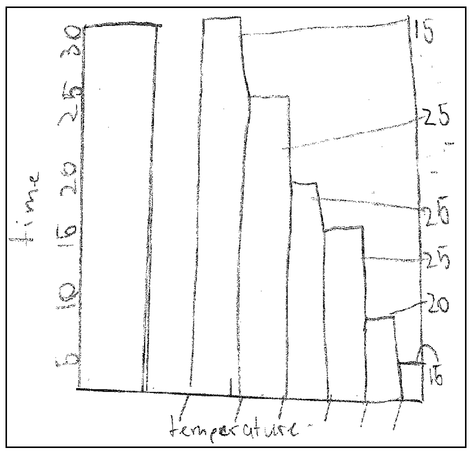


Figure 3.13. Multi-structural response to Item 4 drawn by Student 63.

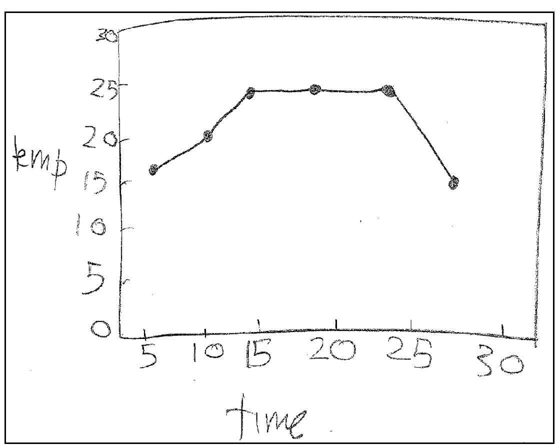
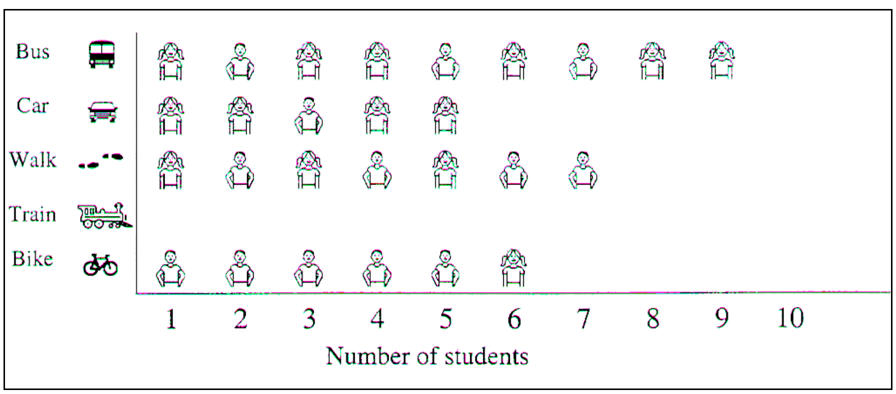


Figure 3.14. Relational response to Item 4 drawn by Student 38.

Item 5 – Interpreting a Pictograph

The task in Figure 3.15 was selected for the Student Survey as it provided the opportunity for students to demonstrate their graph interpretation skills, demonstrate their understanding of using graphs to make inferences, and show how they used the context of the data to draw conclusions. The task in this item was developed for a large chance and data survey administered to 730 students from Years 3, 5, 7, and 9 (Watson, Kelly, Callingham, & Shaughnessy, 2003). The task explored students' knowledge of graph reading, variation, and inference. The pictograph format was chosen to provide the opportunity for younger students to engage with the task. Questions 5.1 and 5.2 focused on eliciting information about the students' graph reading skills. The other four questions were designed as open-ended questions to provide students with the opportunity to explain their answers (Watson & Kelly, 2003). Watson and Kelly found that students often considered this task to be difficult as there were no "correct" answers to some of the questions. As a result, the task was successful in drawing out different levels of understanding in the students' responses.

How children get to school one day



Question 5.1. How many children walk to school?
Question 5.2. How many more children come by bus than by car?
Question 5.3. Would the graph look the same everyday? Why or why not?
Question 5.4. A new student came to school by car. Is the new student a boy or a girl? Explain your answer.
Question 5.5. What does the row with the Train tell about how the children get to school?
Question 5.6. Tom is not at school today. How you think he will get to school tomorrow? Why?

Figure 3.15. Item 5 of the Student Survey (Reproduced from Watson & Kelly, 2003, p. 722).

Questions 5.1. How many children walk to school? The students' responses to Question 5.1 were coded on a correct-incorrect basis only. The correct answer was "7," which most of the students identified correctly. The students' responses indicated that the majority of the students could read values from a graph. Only three of the students gave incorrect responses. Student 1 may have counted incorrectly as his response was "6." Student 66 said "none," and Student 31 did not answer the question directly. He did, however, draw the diagram in Figure 3.16. It was not clear which part of Item 5 the diagram was related to but it shows the student had an appreciation of variation within a data set and could extract information from a graph.

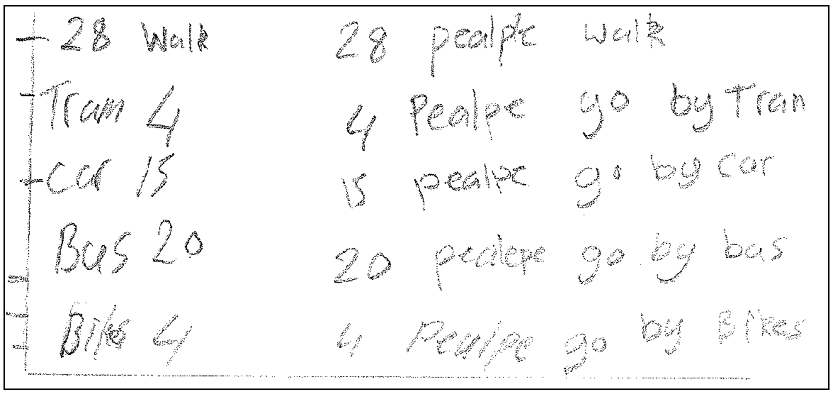


Figure 3.16. Response to Item 5 drawn by Student 31.

Question 5.2. How many more children come by bus than by car? The correct response to Question 5.2 was "4." Seventy-three percent (n=52) of the students answered Question 5.2 correctly. Three students, however, said the answer was "14." To get this answer the students most likely added the number of children travelling by bus to the number of children travelling by car. Another six students said the answer was "9." That answer is incomplete and is related to the number of children arriving by bus only. The students who gave this response read the value from the graph but did complete the calculation necessary to determine the correct answer. Five students gave "2" as their response to the question. The answer of "2" is correct if the question was asking the difference between the number of children who walked to school and the number who arrived by bus. The markings on the graph from Student 60 confirmed he had worked out that "2" was the answer in that way.

Student 63 gave an interesting response to this question. He wrote, "bus 2 more, car 2 less." It appears that the student compared the number of children walking to school with the number of children arriving by car and by bus. Although this was an incorrect response, it did indicate that the student extracted information from the graph and compared different data values. Five students did not respond to the question or gave inappropriate responses.

Questions 5.3. Would the graph look the same everyday? Why or why not? Question 5.3 was designed to elicit students' understanding of variation and the possibility of change (Watson & Kelly, 2003). Watson and Kelly coded the responses to the task according to a hierarchical coding schema that had three levels – Levels 0, 1, and 2. Responses at Level 0 are either inappropriate responses or no response. At Level 1, responses incorporate an implicit recognition of variation without acknowledging uncertainty and at Level 2 responses display a realistic recognition of variation and uncertainty. The coding schema developed by Watson and Kelly was used to code the responses to Question 5.3 in this inquiry. A summary of the results for Question 5.3 is given in Table 3.7.

Table 3.7.
Percentage of Responses for the Coding Levels for Question 5.3

Level 0	Level 1	Level 2
21.1 (n=15)	26.8 (n=19)	52.1 (n=37)

At the highest level, Level 2, the students used the language of uncertainty in their responses, such as "might" and "could." At this level, the responses included notions of variation and uncertainty that were embedded in meaningful contexts. For the most part, the responses incorporated contexts that were familiar to the students. In response to the question "Would the graph look the same everyday?" Student 9 wrote, "probably not because weather might take over for the walkers." Student 15 wrote, "No, because someone's car might break down, or it rains." Another popular context was related to changes due to parents having to work as expressed by Student 18, "No, because some children's parents might work and want to take their kids to school." Student 23 suggested, "If the parents have to work the child might have to walk." Only two students offered responses that incorporated data from the graph to make predictions about possible changes.

> No, because not every day people come to school by bus, for example. Because there is 9 today going by bus and by car there are 5 people going by car, but tomorrow there could be 9 people that come by car and 5 come by bus. (Student 6)

The responses coded at Level 1 included declarative statements that recognised variation but did not include notions of uncertainty. Student 38, for example noted, “No. Absent or different transport.” Student 45 also offered a Level 1 response when he stated, “No, because people have days off.” The responses coded at Level 0 included either no responses or statements such as “No,” with no supporting explanation. Only one student suggested “It would look the same” (Student 68).

Question 5.4. A new student comes to school by car. Is the new student a boy or a girl? Explain your answer. Question 5.4 required students to make a prediction about new data to be added to the graph. The responses for this question were coded according to the developmental model for Graph Interpretation developed by Watson and Fitzallen (2010). The Graph Interpretation model was developed according to the SOLO hierarchy (Biggs & Collis, 1982) and includes elements of Concept of Graph, Concept of Variation, Concept of Average, Context, and Critical Questioning Attitude. To accommodate the purpose of Question 5.4 the coding schema focused on the element of Concept of Graph. Inappropriate responses, no response, or responses that draw on context only without engagement with the graph are considered to be at the pre-structural level. At the uni-structural level responses involve extracting individual values from a graph to make decisions. At the multi-structural level, the responses consolidate the message from a graph by drawing together individual values or an individual value with other characteristics of a graph. Responses at the relational level combine multiple pieces of information from a graph to question or draw implications. Also at the relational level responses may incorporate an understanding of the context of the data. A summary of the results for Question 5.4 is given in Table 3.8.

Table 3.8.
Percentage of Responses for the SOLO Levels for Question 5.4

Pre-structural	Uni-structural	Multi-structural	Relational
50.7 (n=36)	26.8 (n=19)	16.9 (n=12)	5.6 (n=4)

Pre-structural responses for Question 5.4 included no response or non-statistical responses, such as “I don’t know” or “There is not enough information.” In addition, responses that did not provide evidence that the students had engaged with the graph to draw their conclusions were coded at the pre-structural level. These included responses that nominated a gender for the new student but did not provide an explanation for the choice made. At the uni-structural level the majority of the responses nominated that the new student would be “a girl I think, because I look in cars and a girl was at the end” (Student 35). Responses such as the one offered by Student 35 were considered uni-structural as there was no evidence that the students had engaged with the graph beyond stating the value of the last data point for the transport option of Car.

Responses coded at the multi-structural level showed an appreciation of the possible variation in the data and substantiated their choice by describing the pattern evident in the graph. Student 39, for example, suggested the new student was a “boy. Because it goes girl, girl, boy, girl, girl, and it must be a boy.” Responses at the relational level extracted information from the graph and recognised the possibility of variation and uncertainty of the context. This was evident in the response given by Student 5 who said, “I would estimate that he was a boy because there is only one boy [by car] but most of the boys come by bike but he hasn’t.”

Question 5.5. What does the row with the Train tell about how the students get to school? At the uni-structural level the responses showed an appreciation of the context as it appeared in the graph. At the multi-structural level, the responses consolidated the message from the graph by drawing together the context of the data and information from the graph. Responses at the relational level combined the context of the data and information from the graph to question or draw implications from the graph.

Question 5.5 was considered similar in nature to Question 5.3 (Watson & Kelly, 2003) and thus coded using the coding schema developed for Question 5.3. A summary of the results for Question 5.3 is given in Table 3.9. Overall, 31% of responses were coded at Level 0, half of which included references to the proximity of the train station and the school. The other half included statements like “Nothing” or “Not enough information.” Student 60, however, offered an alternative answer. She suggested that “no one likes to catch the train.” The majority of the students read the graph literally and said “No students caught the train” or similar. These responses were coded at the Level 1. Only two students offered responses that included notions of variation or used statistical language. These responses were coded at Level 2. Student 36 wrote, “None of the students caught the train on that day,” and Student 4 wrote, “Most children don’t take the train.”

Table 3.9.
Percentage of Responses for the Coding Levels for Question 5.5

Level 0	Level 1	Level 2
31.0 (n=22)	66.2 (n=47)	2.8 (n=2)

Question 5.6. Tom is not at school today. How do you think he will get to school tomorrow? Why? Watson and Kelly (2003) designed Question 5.6 to elicit students understanding of variation and uncertainty. Although those aspects of statistics were evident in the students’ responses to this question in the Student Survey, Question 5.6 was used to ascertain if students used data or their knowledge of the context to make decisions. With this in mind, the data were coded and then clustered according to four dominant categories: Category 1 – No response or an inappropriate response; Category 2 – An option selected but no explanation offered; Category 3 – An option selected with a non-statistical reason given; and Category 4 – An option selected and supported by data extracted from the graph. A summary of the results for Question 5.6 is given in Table 3.10.

Table 3.10.
Percentage of Responses for the Categories for Question 5.6

Category 1	Category 2	Category 2	Category 3
23.9 (n=17)	26.8 (n=19)	40.8 (n=29)	8.5 (n=6)

The responses in Category 1 included “Don’t know,” “can’t tell” or offered no response. In Category 2 “car” was the most popular option chosen followed by “bus.” In Category 3 the responses involved the use of context without any statistical justification or engagement with the data in the graph. The responses in this category drew on a variety of contexts and were, at times, creative. Student 29, for example, offered a variety of reasons for selecting “bus” as the mode of transport. She wrote, “Bus because Mum and Dad doesn’t have a car, has no bike, on crutches and is scared of trains.” In Category 4 the responses included data from the graph, such as “Probably by bike because you can see most of the boys ride their bikes to school” (Student 19).

Item 6 – Design Questions from Data Cards

The purpose of Item 6 was to determine if students could extract information from data cards and use that information to generate questions about the data. The use of data cards in the Student Survey was considered important as data are automatically represented in data cards in *TinkerPlots*. The format of the data cards in the task (Figure 3.17) was modelled on a set of data cards used by Chick and Watson (2001). Watson and Chick asked 27 Year 5/6 students to look for and show any interesting features about the data displayed on a set of data cards. The results of their study demonstrated that the use of data cards provided information about the students' thinking at varying levels across the three categories – observation of characteristics of data, interpretation of data, and creation of representations. The responses to Item 6 in the Student Survey were coded according to the P-U-M-R sequence described in Question 1.2. A summary of the results is given in Table 3.11

Table 3.11.
Percentage of Responses for the SOLO Levels for Item 6

Pre-structural	Uni-structural	Multi-structural	Relational
15.5 (n=11)	38.0 (n=27)	19.7 (n=14)	26.8 (n=19)

The information about individual students is on separate cards, like the ones below.

Item 6. What questions about the group of students could be answered by using the information on the cards?

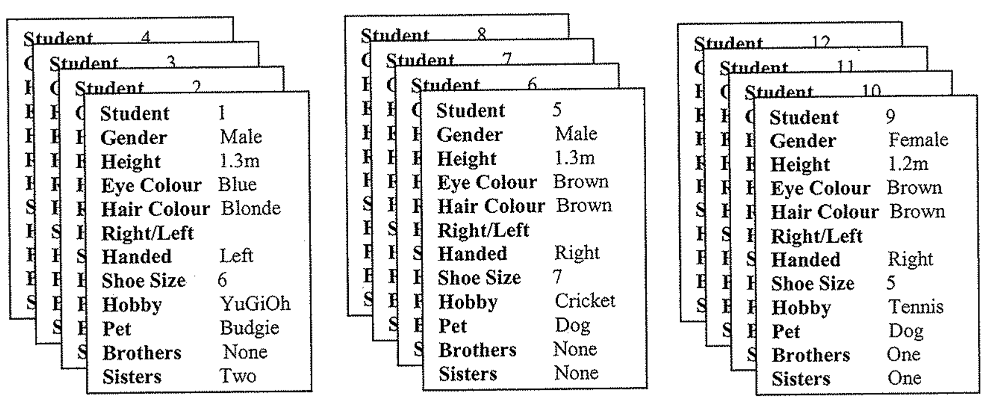


Figure 3.17. Item 6 of the Student Survey (Adapted from Chick & Watson (2001)).

Item 6 elicited a variety of responses that included knowledge questions that were based directly on the data on the cards, creative questions that were related to the context but unrelated to the data or questions that required making inferences or comparing groups. The pre-structural responses were non-statistical, inappropriate, or no response. At this level, Student 61's questions were related to the answers on the cards. He asked, "What YuGiOh do you have? What shoes do you where? What brand racquet do you have?" Also coded at the pre-structural level were the questions posed by Student 29. She asked, "Have Mum or Dad got glasses?"

At the uni-structural level the students posed questions that focused on the individual cases. Student 18 posed a number of questions that were typical of the questions coded at the uni-structural level. The questions took into consideration all the attributes listed on the data cards. The questions included:

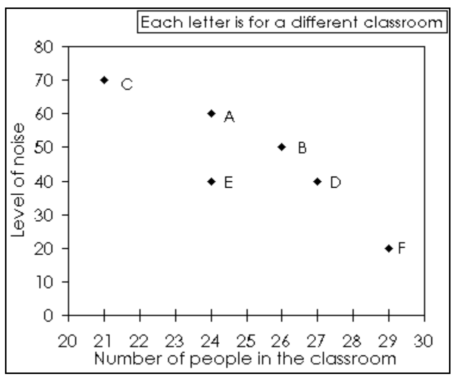
> Is it a boy or a girl? What colour hair? How many siblings? How high is the person? Does the student have any pets? Does the student have a hobby? What colour eyes? What hand does the student write with? What shoe size?

At the multi-structural level the students' questions focused on the group of cases and asked questions that were predominantly phrased as "How many … in the class?" The questions posed that were coded at the relational level were more creative and required inferences or comparison to be made, which draw together relevant elements from the data cards. These included Student 23's question, "Which two students are the same and how much are they taller than the other students?" Student 6 asked, "How many girls like dogs?" "Who are taller the males or the females?" Student 40 asked, "What is the most common hair colour? Do most people prefer sport than outside games? Do people have more brothers than sisters?"

Item 7 – Interpreting a Scatterplot

To determine the range of student experiences with graphs and to establish if the students were able to reason about covariation, the task in Figure 3.18 was included in the Student Survey. The task was used by Moritz (2003a) in an interview-based study of student reasoning about a scatterplot. The graph was used to assess students' verbal and numerical graph interpretation skills and knowledge. In the Moritz study, the students' interpretations of the data were observed at four SOLO levels of understanding for the numerical and verbal interpretations. According with the coding schema developed by Moritz (2003a), pre-structural level responses are either inappropriate responses or no responses. At the uni-structural level the responses are related to reading individual values from either axis of the coordinate system. The next level of response, multi-structural, requires the students to interpolate the data and use multiple data points to justify an answer. At the relational level responses need to consider the whole data set. Question 7.2 only required a numerical response that was either correct or incorrect and was coded accordingly.

Some students were doing a project on noise. They visited 6 different classrooms. They measured the level of noise in the class with a sound meter. They counted the number of people in the class. They used the numbers to draw this graph.



Question 7.1. Pretend you are talking to someone who cannot see the graph. Write a sentence to tell them what the graph shows. "The graph shows...."

Question 7.2. How many people are in Class D?

Question 7.3. If the students went to another class with 23 people, how much noise do you think they would measure? Please explain your answer.

Question 7.4. Jill said, "The graph shows that classrooms with more people make less noise". Do you think the graph is a good reason to say this? □ YES or □ NO Explain your answer.

Figure 3.18. Item 7 of the Student Survey (Reproduced from Moritz (2003a, p. 525)).

Question 7.1. Pretend you are talking to someone who cannot see the graph. Write a sentence to tell them what the graph shows. "The graph shows …." The responses to Question 7.1 coded at the pre-structural level were similar to the one offered by Student 30 who said, "I don't know how to do this kind of graph." Many other students did not provide a response at all. The responses at the uni-structural level focused on one data point or one attribute. Student 20 suggested the graph showed "What class was the noisiest" and Student 17 wrote, "The graph is measuring sound in decibels." Multi-structural responses included, "The graph shows that the noise level in Class C is 70 compared to the noise in Class F. It is 20" (Student 3). The relational level responses included "The graph shows the smaller classes make more noise" (Student 1) and "The graph tells you how much noise that different classes make and the number of students in that class, the classes with less people make more noise, according to the graph" (Student 10). A summary of the results for Question 7.1 is given in Table 3.12.

Table 3.12.
Percentage of Responses for the SOLO Levels for Question 7.1

Pre-structural	Uni-structural	Multi-structural	Relational
35.2 (n=25)	26.8 (n=19)	22.5 (n=16)	15.5 (n=11)

Question 7.2. How many people are in Class D? Question 7.2 only required a numerical response that was coded correct, incorrect or no response. Thirty-two percent of the students did not provide a response to this question. Ten students (14%) gave incorrect answers, five of which gave "40" as the answer. The students who gave this response read the value for Class D from the vertical axis, which was related to the noise level not the number of people in the class. The remaining 54% of the students gave the correct response of "27."

Question 7.3. If the students went to another class with 23 people, how much noise do you think they would measure? Please explain your answer. The focus of Question 7.3 was on estimating the noise level of a class with 23 students. Only two students offered responses at the relational level. Student 48 wrote, "I think it would make 65. The less students there are the more noise they make, so it would be the second highest, because it has the second least amount of students." Student 17 used the trend of the graph to make his decision. He wrote, "maybe 65 because it's going up in a line from 'A' to 'F'." At the multi-structural level the responses gave values that were based on the location of nearby individual data points, such as the response offered by Student 3, "I think it would be around 60, like Class A." Many students gave responses that were coded at the uni-structural level. Those students gave values for the noise level but did not justify their answer. Thirty-nine percent of the students did not respond to this question. A summary of the results for Question 7.3 is given in Table 3.13.

Table 3.13.
Percentage of Responses for the SOLO Levels for Question 7.3

Pre-structural	Uni-structural	Multi-structural	Relational
39.5 (n=28)	52.1 (n=37)	5.6 (n=4)	2.8 (n=2)

Question 7.4. Jill said, "The graph shows that classrooms with more people make less noise". Do you think the graph is a good reason to say this? Explain your answer. Question 7.4 elicited the most no responses or inappropriate responses for the questions in Item 7. The question required students to respond to a statement about the graph with a Yes or No answer and then explain their answer. Many of the students' responses at the pre-structural level were derived from their knowledge of the context and not related to the data in the graph. Student 19, for example, offered, "No because the students that were noisy that day might be quiet other days and the students that were quiet might be noisy other days." In addition, the responses that ticked Yes or No as a response and then offered no explanation were coded at the pre-structural level. Sixty-two percent of the students' responses were coded at the pre-structural level. At the uni-structural students used one characteristic of the graph to justify their conclusions. "Yes because you can see how much noise there is" (Student 24) is a typical example of the statements coded at the uni-structural level.

At the multi-structural level the students' responses included multiple values and information from the graph but did not extend to making general statements necessary to be coded at the relational level. Student 20 noted, "Yes, because F has 29 people, which is quieter than C, which only has 21 people." At the relational level the students gave responses that indicated they had identified the data points were related to the two attributes and there was a relationship between the attributes. This was evidenced in Student 48's response. She wrote, "Yes. You can see that the graph is going down. The class with more people make less noise and the classes with less people make more noise." A summary of the results for Question 7.4 is given in Table 3.14.

Table 3.14.
Percentage of Responses for the SOLO Levels for Question 7.4

Pre-structural	Uni-structural	Multi-structural	Relational
62.0 (n=44)	8.5 (n=6)	9.8 (n=7)	19.7 (n=14)

Summary of Student Survey

The Student Survey was designed as a formative assessment task for this inquiry and as such provided extensive information about the students' prior learning. The items in the Student Survey were selected to provide information about the students' prior learning in relation to the dimensions of the *Model of Learning Behaviour in EDA Graphing Environments* (Figure 1.7) as detailed in Table 3.1.

For the dimension of *Generic knowledge*, the results from the Student Survey suggest that the students can recognise the characteristics of data and graphs but often lack the language needed to express their ideas and explain their thinking. For the most part they were competent at reading values from a graph. The results also imply that the students had little experience in using graphing software to construct graphs. For the dimension of *Being creative with data*, the graphs drawn by the students provide evidence that most of the students were able to construct graphs from data but selected a limited range of graph types to present the data. They were, however, able to construct bar charts or frequency graphs readily.

When asked to describe data from graphs most of the students cited individual values but few of them expressed a comprehensive view of the data. For the dimension of *Understanding data*, the students' results suggest that they were inexperienced at making sense of data and graphs. Often questions about the data were answered with numerical or one word answers without justification or explanation. It appears likely that the students had little experience in identifying messages from the data. For the dimension of *Thinking about data*, the results from the Student Survey also suggest that the students were inexperienced at making inferences and recognising the limitations of the data. They were able to identify variation in a graph but did not provide enough evidence to suggest they were competent at determining the trend in a graph.

The Student Survey employed four tasks used in other studies – Items 4-7. Direct comparison of the results from those studies and the present inquiry was not possible as the year groups and ages of students were different, the application of the task was not the same, or a different coding schema was used to analyse the responses, the details of which for each item are provided in the descriptions of the items. It is important to note, however, that the typical responses to the relevant items in the Student Survey replicated the examples given in the research reports for each of the items (for examples see Moritz, 2003a, 2003b, 2004; Watson & Kelly, 2003; Watson et al. 2003). This is an indication of the robust nature of the items selected and supports the choice to use them in this inquiry to elicit knowledge about the students' prior learning.

Worthy of particular attention are the responses to Item 7. Item 7 had the highest number of no responses for the Student Survey. The item included a scatterplot displaying covariation. Student responses to Item 4 (Figure 3.11) indicate that many students were unfamiliar with the covariation scatterplot representation as many of them chose to construct bar charts for the data instead of scatterplots that would have displayed the data as a time series graph. This lack of experience with different graph types may have contributed to the students' inability to respond to Item 7. In addition, Item 7 was the last item on the Student Survey and some students may have become fatigued and not able to answer the questions in the item.

Selection of Students to Participate in Stage 4

One of the purposes of the Student Survey was to use the results to select the students to participate in the Sequence of Learning activities implemented in Stage 4. To make the selection, the two teachers of the Year 5/6 classes at the primary school were provided with the results from the Student Survey for their students. The teachers were asked to select six students each to be invited to participate in Stage 4 – Sequence of Learning and Outcomes. The teachers were asked to base their decisions on the results from the Student Survey and their knowledge of the students in order to select a sample that represented a spread of academic ability in each class. In addition, the teachers were asked to recommend students who would be able and willing to discuss their work and their thoughts with the researcher. Gender was not a criterion for selection.

The results from the Student Survey provided a descriptive account of the prior learning of the students. As it was intended to be a formative assessment task, the results from the Student Survey were not put together to provide a “total” for each student. The students’ results were arranged randomly when given to the teachers. This was done so the students’ responses were not ordered according to their performance on the Student Survey, which may have influenced the teachers’ decision about which students they recommended to participate in the inquiry. The aim was to inform the teachers of the outcomes for their students so that they could use the results to inform their own teaching practice. The results from the Student Survey did not contribute to the students’ end-of-year summative assessment.

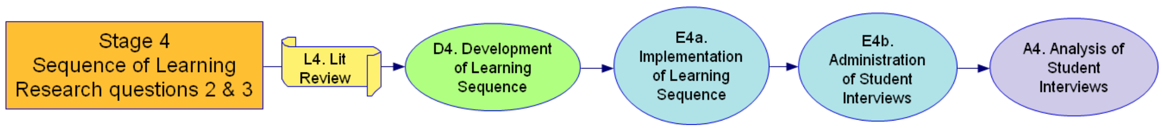
All the students recommended by their teachers were then asked by their teachers if they would be happy to participate in the learning activities using *TinkerPlots* with the researcher. Parental consent was gained from the parents of those students. All of the twelve students selected agreed to participate in the inquiry and completed all the tasks required. Their names are: Blaire, Jake, James, Jessica Johnty, Kimberley, Mitchell, Natasha, Natalie, Rory, Shaun, and William.

Concluding Remarks

The student responses on the Student Survey provided rich, descriptive information about their understanding of graphs and graph creation. The Student Survey was successful in providing a comprehensive view of the students’ understanding of graph creation and graph interpretation skills across all the dimensions of the *Model of Learning Behaviour in EDA Graphing Environments* (Figure 1.7). The valuable information elicited about the prior learning of the students was used to plan the implementation of a Sequence of Learning Experiences that was designed to support students’ development of statistical thinking and reasoning using *TinkerPlots* in Stage 4 of the inquiry. In addition, the results were used to select the students who participated in the Sequence of Learning Experiences implemented in Stage 4.

Stage 4

Sequence of Learning and Outcomes



The purpose of Stage 4 of the inquiry is to develop Student Profiles that characterise the students' development of data analysis skills in the *TinkerPlots* learning environment. The dimensions of the *Model of Learning Behaviour in EDA Graphing Environments* (Figure 1.7) are used as organisers to build the Student Profiles from the data collected from the Student Interviews. The Student Profiles are used to present the data from the Student Interviews and provide a descriptive account of each of the students' participation in the Student Interviews. The Student Interviews were conducted at the end of the implementation of a sequence of learning experiences that was developed to provide the opportunity for the students to learn how to use *TinkerPlots* for data analysis and develop an understanding of covariation.

This stage of the thesis describes the development of the sequence of learning experiences by using theoretical frameworks from the literature on the teaching and learning of graphing and covariation. The development of the learning sequence is also informed by knowledge of student prior learning gleaned from the Student Survey administered in Stage 3. The methodological issues related to conducting and analysing the Student Interviews to construct the Student Profiles are addressed in this stage of the inquiry. The participants for this stage of the inquiry were selected in Stage 3.

Developing a Sequence of Learning

To facilitate learning experiences that develop higher-order thinking skills, teachers develop sequences of activities that draw on students' knowledge of and experience with authentic contexts and then shift that thinking towards representing and considering the key concepts in abstract ways (Van de Walle, 2007). Programs of learning are an interpretation of the curriculum and should follow the main principles of the curriculum. The emphases and activities in learning programs represent the teachers' judgements about what they consider important to guide their particular class to optimise student learning (Marsh, 2004). The judgements made for the sequence of learning experiences in this inquiry were informed by the literature on teaching and learning of statistical concepts and the curriculum discussed in this section, in conjunction with knowledge of students' prior learning determined in Stage 3.

In the case of the development of covariation skills within *TinkerPlots'* EDA software environment for the novice learners in this inquiry, the learning process is multi-dimensional and many aspects of graph understanding need to be considered. These include student understanding of: characteristics of data (Biehler, 1997), characteristics of graphs and graph conventions (Kosslyn 1989), graphical and statistical representations afforded by *TinkerPlots* (Watson & Donne, 2009), using *TinkerPlots* to create a variety of graphs (Fitzallen & Watson, 2010), interpreting graphs to identify covariation (Konold & Khalil, 2003; Moritz, 2004), and using an understanding of covariation to make informal inferences (Watson & Donne, 2008). To maximise the opportunities for students to develop an understanding of these aspects of graph creation and graph interpretation, a learning sequence that embraces a broad view of learning whilst still attending to the finer details of statistical thinking and reasoning is necessary.

Structure of Learning

As an organiser for structuring investigations made by students, the instructional design theory called Realistic Mathematics Education (RME) is useful to guide student learning that focuses on the development of abstract thinking (Stephan, 2009). RME allows students to be active participants in the learning process and promotes the development of learning experiences that start with rich contexts that provide real-life sources of mathematics. "The contexts thus serve not only as a source but as an area of application as well" (Heuvel-Panhuizen, 2009, p. 12). Students do not actually have to have experienced the situation but must be able to make meaningful connections to the context. RME also suggests instruction should build students' reasoning gradually from the concrete to the abstract. This is achieved through the use of manipulatives, diagrams, and other imagery to reinforce students' reasoning. The final component in the RME organiser is that students should create models and use symbols to represent their concrete activity. These models are then used as reasoning devices for more abstract thinking. An essential part of RME instruction is the active participation of students in this process by making artifacts that are meaningful and useful to them (Bakker, 2004; Stephan, 2009).

RME provides a general structure that is a useful starting point for designing a sequence of learning experiences. Within this structure instruction should be intentionally designed so that students reorganise their thinking progressively toward more abstract ideas (Stephan, 2009). It also recognises the importance of context – a vital aspect of statistical investigations and it accommodates the use of *TinkerPlots* as it places an emphasis on student construction of models and symbols to facilitate abstract thinking.

Model of Statistical Learning

Comprehensive learning about graph creation and graph interpretation skills together with developing an understanding of covariation encompasses more than knowing how to recognise and draw the constituent parts of graphs described by Kosslyn (1989) in Figure 1.1. It involves statistical investigations that include all the steps in the *Model of Statistical Investigation* developed by Watson (2009) detailed in Figure 4.1. This model aligns with the developmental progression proposed by the *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report: A Pre-K-12 Curriculum Framework I* (Franklin et al., 2005). Similar to Watson's *Model of Statistical Investigation,* the *GAISE Report* curriculum framework recognises that any statistical investigation is underpinned by variation (Franklin et al., 2005).

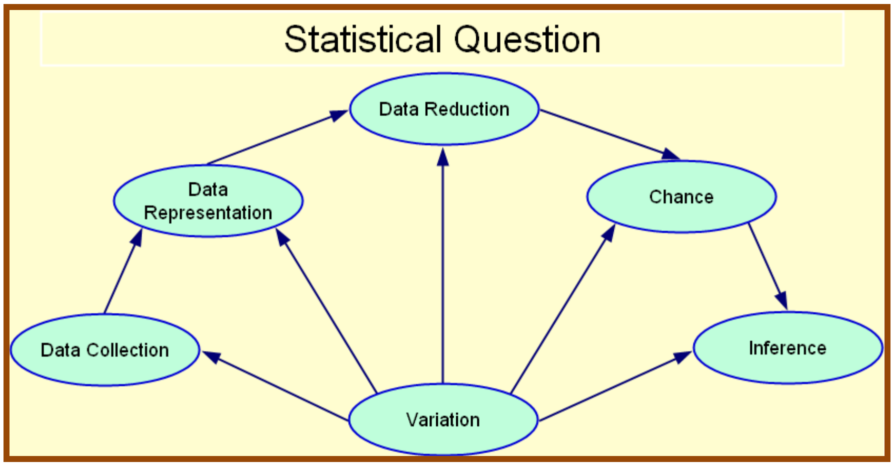


Figure 4.1. Model of Statistical Investigation (Reproduced from Watson, 2009, p. 91).

Watson's (2009) model recognises the importance of context by promoting that the starting point of a statistical investigation should be directed towards posing a statistical question. The question sets the scene for an investigation and draws in the context of the data. The Data Collection step provides the data that can be represented in a number of forms – numerical, pictorial or graphical. The data are often reduced using statistical calculations of measures of centre and spread or graphical representations such as a box-and-whisker plot. These representations or others are then used to make inferences about the data that answer the question posed at the onset of the investigation. Part of the inference step is recognising the uncertainty of the conclusions drawn. This emerges from the Chance step of Watson's model.

It is ideal to include all of Watson's (2009) steps from the *Model of Statistical Investigation* in a sequence of learning experiences as the steps encompass the aspects of working statistically detailed in the *Tasmanian Curriculum Mathematics-Numeracy* (DoET, 2007) and the *GAISE Report* curriculum framework (Franklin et al., 2005). As well as accommodating "actively collecting, organising, summarising and interpreting data" (p. 63), Watson's model also recognises the importance of context through the use of an over-arching statistical question to guide an exploration of data. The notion that not all conclusions can be made with the same level of confidence and the underpinning idea that variation is fundamental to all statistical investigations is embedded in Watson's model.

Reforms in Statistics Education

The increasing availability of new technologies that generate and use statistical information and the growing awareness of the value of statistical information to inform decisions and influence policy has led to reforms in statistics education (Moore, 1997). As a result, three main general reforms have emerged – content, pedagogy, and technology. Moore proposes that "the most effective learning takes place when *content* (what we want students to learn), *pedagogy* (what we do to help them learn), and *technology* reinforce each other in a balanced way" (p. 124). It is the synergy between these three characteristics of learning environments that facilitate learning. These three characteristics are discussed in turn as they relate to this inquiry.

Content

The content selected to be the focus of learning activities should be guided by where the core mathematical concept is situated in the progression of learning outlined in curriculum documents and based on an understanding of the teaching and learning of the core mathematical concept (Marsh, 2004). This inquiry explored where covariation was situated in curriculum frameworks and recommendations from the literature about sequencing learning about graph creation and graph interpretation.

Curriculum content. In 1989, the *Curriculum and Evaluation Standards for School Mathematics* presented the U. S. National Council of Teachers of Mathematics' (NCTM) view of what should be valued in school mathematics. This document was updated in 2000 as *Principles and Standards for School Mathematics*. The *Principles and Standards* are instructional programs for students from prekindergarten through to Year 12 and are indicators of what students should be enabled to do. The Data Analysis and Probability Standard puts an emphasis on developing statistical reasoning skills and calls for students to formulate questions and collect, organise, and display relevant data to answer questions; to select and use appropriate statistical methods to analyse data; and to develop and evaluate inferences and predictions that are based on data (NCTM, 2000).

To complement the *Principles and Standards* (NCTM, 2000), the American Statistical Association (ASA) commissioned development of the guidelines and curriculum framework detailed in the *GAISE Report.* The guidelines were developed to provide a conceptual structure for statistics education in order to give a coherent picture of the developmental nature of the Data Analysis and Probability Standard in the *Principles and Standards*. The aim was to support teachers, who may not have had an opportunity to develop a sound knowledge of the statistical concepts underlying the practices of data analysis.

The GAISE Report curriculum framework (2005) has four problem solving processes that acknowledge the role of variation in statistics: Formulate questions, Collect data, Analyse data, and Interpret results. It provides a developmental sequence of learning experiences that incorporates all four problem-solving steps across three levels of development – levels A, B, and C. Covariation, or "covariability" as it is termed in the framework, is included as a point of focus in Level B of the developmental framework. It is, however, applied implicitly in all levels of the framework when recommendations are made for the application of scatterplots. At Level A scatterplots are used to "observe association between two variables (p. 31)" and "look for patterns and trends" (p. 33) in data. At Level B the use of scatterplots is recommended for "measuring the strength of association between two quantitative variables" (p. 48) and "modelling linear association" (p. 51).

Following on from the lead from the U.S., statistics content was formally introduced into the curriculum in the Chance and Data Strand when *A National Statement on Mathematics for Australian Schools* was released in 1991 (Australian Education Council [AEC]). At that time, the purpose of the *National Statement* was to promote consistency in curriculum outcomes across Australia. This document provides a curriculum framework that describes the progression of learning typically achieved by students during the compulsory years of schooling (Years K-10) detailed in the *Mathematics – A Curriculum Profile for Australian Schools* (AEC, 1994).

The Chance and Data strand of the *Mathematics Profile* has four organisers that apply when conducting statistical investigations: Collecting data, Organising data, Displaying and summarising data, and Interpreting data (AEC, 1994, p. 13). In the primary years (Years K-6), graphing activities focus on students representing univariate data in pictographs and bar graphs. The curriculum outcomes for early secondary students (Years 7 and 8) allow for the continuing use of univariate data with a variety of graph types but also extend to the interrogation of bivariate data. The descriptors include "Displays one-variable and two-variable data in plots" and "Reads and describes information in histograms, plots," with elaborations that comprise "Represent two-variable data in scatter plots and make informal statements about relationships," "Informally interpret relationships and reach conclusions from scatter plots" and "Write or present an accurate summary of the information displayed in a range of tables and graphs" (p. 93).

Although not adopted in full by all Tasmanian schools, the *National Statement* (AEC, 1991) heavily influenced curriculum documents produced by the Department of Education, Tasmania (DoET) from that time on. The most recent iteration is the *Tasmanian Curriculum Mathematics – Numeracy* (2007). Included in this curriculum are the five content strands from the *National Statement*, one of which is Chance and Data. The *Tasmanian Curriculum* is the curriculum framework most relevant to the implementation of this inquiry as the inquiry investigates students from Tasmanian schools.

The scope and sequence of learning opportunities detailed in the *Tasmanian Curriculum* (DoET, 2007) consists of 15 stages of learning, which provide a continuum of learning from Kindergarten to Year 10, as shown in Table 4.1. The *Tasmanian Curriculum* encourages teachers to plan learning opportunities across a range of stages for each year group, in order to allow time for students to demonstrate understanding following the learning of new concepts and challenge students in their learning, as well as consolidate ideas. To achieve these goals the complexity of the descriptors in the stages increases as the stages progress.

Table 4.1.
Stages for opportunities to learn: The Tasmanian Curriculum – Mathematics-Numeracy (DoET, 2007, p. 11)

Year level	Kinder-Prep	Years 1 and 2	Years 3 and 4	Years 5 and 6	Years 7 and 8	Years 9 and 10
Opportunities to learn	Stages 1-4	Stages 3-6	Stages 5-9	Stages 7-11	Stages 9-13	Stages 11-15

The early stages of the *Tasmanian Curriculum* (DoET, 2007) have two main foci: reading graphs, and creating personally meaningful data representations. Different data representations or graph types are introduced strategically throughout the curriculum with the introduction of pictographs introduced in Stage 2, bar charts in Stage 4, and tally tables in Stage 5. No new graphs types are introduced after Stage 5 until Stage 10. The emphasis in the curriculum, however, shifts from reading graphs to the interpretation of graphs, which allows for a consolidation of using basic graph types to answer questions about the data. The notion of variation emerges throughout these stages as well, although not emphasised explicitly.

Stage 10 is a pivotal moment in the *Tasmanian Curriculum* (DoET, 2007) as many new graph types are introduced. Line graphs, stacked dot plots, scatterplots, and stem-and-leaf plots are introduced in this stage together with the mean, median, and mode. The use of discrete and continuous data is also a focus point. The increase in intensity of using a larger range of more complex graph types aligns closely with the progression proposed by Rangecroft (1991b). To complement the other representations, box-and-whisker plots and two-way tables are introduced in Stage 11 and frequency tables and histograms in Stage 13. The extensive choice of graph types from Stage 10 on potentially provides "the opportunity to use the graph type most appropriate for the data collected and the questions explored" (Watson & Fitzallen, 2010, p. 25).

As well as the increase in the availability of a range of graph types, Stage 10 sets the scene for more meaningful data analysis activities (DoET, 2007). The inclusion of the descriptor "actively collecting, organising, summarising and interpreting data" (DoET, 2007, p. 63) extends the use of graphs from merely determining values and identifying differences in data to developing "working statistically" outcomes. Working statistically is about carrying out investigations that have more than one step and that encompass the stages shown in Figure 4.1 (Watson, 2009, p. 61). This emphasis continues in each progressive stage, through to Stage 15.

The *Tasmanian Curriculum* incorporates technology into the curriculum as a key tool to support students' learning and notes:

> Students learn about and use ICT in mathematics–numeracy to develop their skills of problem solving, communicating and reasoning. They identify data needs and use ICT to locate and access data. They use the power of ICT to organise, manipulate and transform data and develop their own interpretations and understandings. (DoET, 2007, p. 7)

In relation to the Chance and Data strand, direct reference to the application of technologies occurs in three stages. In Stage 9 it is recommended that technology is used "to create graphs and other data representations and select the most appropriate format for the given context" (p. 51). Information and communication technologies (ICT) are mentioned again in Stage 14 in relation to interpreting the correlation coefficient to verify the fit of a linear rule, and in very general terms on Stage 15 – "Using statistical packages, spreadsheets or databases to make judgements" (p. 88).

Although covariation is not mentioned explicitly in the *Tasmanian Curriculum* (DoET, 2007), it could be construed that the development of covariation skills occurs when students use scatterplots to make inferences during a statistical investigation as indicated in Stage 10. As Stage 10 aligns with Years 5/6 of schooling, including covariation in the sequence of learning experiences in this inquiry is appropriate. The choice to use *TinkerPlots* as a learning tool is also supported by the specific introduction of using technology to construct different graph types in Stage 9.

Sequencing the delivery of content. Although the models of graphing (Shaughnessy et al., 1996; Friel et al., 2001; Mooney, 2002; Moritz, 2004) utilised in *Stage 1 – Development of a Model of Learning Behaviour* provide information about the key characteristics of graph creation and graph interpretation skills, they do not provide explicit information about the sequence in which the concepts should be introduced to students. In contrast, Watson and Fitzallen (2010) provide hierarchical sequences of graph creation and graph interpretation that have the potential to assist teachers to the develop sequences of learning experiences.

The graph creation and interpretation hierarchy involves four sequences of development within three cycles of learning. Collectively, the four sequences of development provide an order in which to introduce the fundamental notions of graph creation and interpretation. As it is not associated with any particular grade level, the application of the graph creation hierarchy can be based on the knowledge and experience of students at various stages throughout the extent of the students' education. The Concept of Graph is the first sequence (Figure 4.2). This includes the basic elements of a graph: Attribute, Data, Variation, and Scale. The Concept of Graph suggests an order that describes the process of developing an understanding of the characteristics of a graph and the different levels of understanding that may be realised when drawing parts of a graph together to tell a story about the data. At the relational level, "students can create a meaningful picture/diagram with appropriate scale and tell the story of variation in the data and what it means for the attribute displayed" (p. 59). As a foundation for graph creation, The Concept of Graph underpins all the proceeding cycles.

	The Concept of Graph			
Relational stage	**Combines all elements to create a representative graph**			
Multistructural stage	**Links 2 or 3 elements to create basic graphs**			
Unistructural stage	**Uses single elements unlinked**			
Elements	Attribute	Data	Variation	Scale

Figure 4.2. Developmental sequence for the basic Concept of Graph (Reproduced from Watson & Fitzallen, 2010, p. 59.).

The second cycle of graph creation includes two sub-cycles: The Concept of Graph for Multiple Attributes (Figure 4.3) and The Concept of Graph for Large Data Sets (Figure 4.4). These sub-cycles are underpinned by the consolidated Concept of Graph from the first cycle. The Concept of Graph for Multiple Attributes emphasises the different types of attributes, introduces a variety of graphs, and establishes the need for two dimensional representations to accommodate the measurement of more than one attribute for each case. At the relational level, "students can recognise among the graph types they understand, the one appropriate for particular associations between attributes and apply it to tell the story of the data" (Watson & Fitzallen, 2010, p. 59).

	The Concept of Graph for Multiple Attributes			
Relational stage	**Chooses and or creates appropriate graph for attributes and explains their application**			
Multistructural stage	**Creates graphs from elements: split dot plots, time series, line graph, scatterplots**			
Unistructural stage	**Builds elements, cannot combine into complete graphs**			
Elements	**The Concept of Graph**	Types of attributes	2-D scaling	Relationship of two attributes to single case

Figure 4.3. Developmental sequence for the Concept of Graph for Multiple Attributes (Reproduced from Watson & Fitzallen, 2010, p. 60.).

The Concept of Graph for Large Data Sets (Figure 4.4) introduces large data sets, more complex graphical representations, and data reduction methods. It is also underpinned by consolidated graph concept from the first cycle. At the relational level, "students have the ability to select the appropriate graph for a particular data set and attribute" and establish an "integrated understanding of the various representations" used to interpret large data sets (Watson & Fitzallen, 2010, p. 60).

	The Concept of Graph for Large Data Sets				
Relational stage	**Chooses and or creates appropriate graph for attributes and explains their application**				
Multistructural stage	**Creates graphs from elements: histograms, cumulative frequency graphs, ogives, frequency polygons, box plots, pie charts**				
Unistructural stage	**Builds elements, cannot combine into complete graphs**				
Elements	**The Concept of Graph**	Percentage	5-number summary	Area representing frequency	Equal interval grouping

Figure 4.4. Developmental sequence for the Concept of Graph for Large Data Sets (Reproduced from Watson & Fitzallen, 2010, p. 60.).

The third cycle – Informal Decision Making for Graphs (Figure 4.5) potentially takes the graph creation process one step further and leads to students making informal inferences. It involves making decisions about creating and choosing appropriate graphs for a particular purpose and recognising the potential different representations have in assisting to draw conclusions and make informal inferences beyond the data. At the relational level, "all elements need to be integrated for completely justified conclusions to be reached" (Watson & Fitzallen, 2010, p. 61.). This third cycle is underpinned by the all the previous cycles.

<table>
<tr><td></td><td colspan="4">Informal* Decision Making for Graphs</td></tr>
<tr><td>Relational stage</td><td colspan="4">Combines all elements as required by the question to reach an informal conclusion</td></tr>
<tr><td>Multistructural stage</td><td colspan="4">Combines two or more elements to reach partial informal conclusions for questions about a data set</td></tr>
<tr><td>Unistructural stage</td><td colspan="4">Appreciates elements in isolation, has difficulty combining</td></tr>
<tr><td rowspan="2">Elements</td><td rowspan="2">Concept of Variation</td><td>Concept of Graph for Multiple Attributes</td><td>Concept of Graph for Large Data Sets</td><td rowspan="2">Concept of Average</td></tr>
<tr><td colspan="2">Concept of Graph</td></tr>
</table>

*The term informal is used here to distinguish the hierarchy from one that would involve formal statistical tests.

Figure 4.5. Development sequence for Informal Decision Making for Graphs (Reproduced from Watson & Fitzallen, 2010, p. 61.).

Watson and Fitzallen (2010) stress the importance of providing time for students to consolidate their understanding of graph creation and the role different graph types play in the data analysis process. Not only do students need to understand the elements of a graph but also they need to make connections to the context of the data to make meaningful conclusions and justifications. Watson and Fitzallen go on to say, "The concept of graph is built over time from seeing many different examples in different contexts. It is building this repertoire that makes it possible to choose and use graphs when given data handling questions" (p. 61).

Watson and Fitzallen (2010) provide a further hierarchical sequence – Graph Interpretation (Figure 4.6) – that follows on from the previous four sequences. This sequence is flexible enough to be applied as a second cycle of learning after consolidation of any of the Concept of Graph cycles and attends to the factors, such as outliers, that potentially impact on the shape of a graph. It also includes a questioning attitude element, which is required for critical thinking and development of statistical literacy skills necessary to interpret unusual or potentially misleading graphs so often encountered in the media (Watson, 2006). The complexity of graph interpretation tasks is determined by which cycle of Concept of Graph applies to the data and questions asked.

	Graph Interpretation				
Relational stage	**Ability to question or draw implications from the graph by combining understanding of the elements present**				
Multistructural stage	**Consolidating the message in the graph based on the elements present**				
Unistructural stage	**Appreciation of the single elements as they appear in the graph presented**				
Elements	**Concept of Graph (Basic or Advanced)**	**Concept of Variation**	**Concept of Average**	**Context**	**Critical Questioning Attitude**

Figure 4.6. Developmental sequence for Graph Interpretation (Reproduced from Watson & Fitzallen, 2010, p. 61.).

Pedagogy

Effective teaching practices adopt pedagogies that identify and build on prior knowledge, make real-life connections, develop deep understanding, align with curriculum guidelines, and promote the active participation of students in the learning process (Bakker, 2004; Moore, 1997; Stephan, 2009; Van de Walle, 2007). Such practices encapsulate a holistic view of learning and actively engage students in the learning process at a deeper level (Ramsden, 2003). They draw on learning practices that value developmentally-appropriate facilitator-supported learning that is directed by learners' needs (Van de Walle, 2007). In order to assist students to construct personally meaningful conceptions of mathematical topics, providing the opportunity for them to explain and justify personal solution methods through conversational interaction is paramount (Fraivillig, Murphy, & Fuson, 1999). Greenfield (1984) supports this view in relation to learning in general, and Lovett and Greenhouse (2000) reiterate the stance in relation to teaching and learning about statistical concepts.

In addition, well developed pedagogical content knowledge (Shulman, 1986) is necessary to understand how to teach particular concepts. Shulman suggests that the research that provides the evidence about the instructional conditions necessary to transform student thinking should be considered when developing pedagogical practices.

Statistics education. New techniques in data analysis (Emerson & Hoaglin, 1983; Emerson & Strenio, 1983; Tukey, 1977) provide alternatives to the more traditional procedures that are dominated by probability-based inference. This shift puts an emphasis on using EDA techniques as diagnostic tools for thinking and reasoning about data (Moore, 1997). These new tools, when coupled with knowledge of the context of the data promote critical thinking, a fundamental element of statistical thinking (Pfannkuch & Wild, 2004). Embedded within this is the need to focus on the big ideas of statistics, such as association, distribution, covariation, and inference (Garfield & Ben-Zvi, 2004). The new pedagogies adopted in statistics education focus on the active participation of students in the learning process (Moore, 1997).

Teaching and learning - Data analysis and covariation. Complementing the developmental sequences offered by Watson and Fitzallen (2010) are recommendations from Konold and Khalil (2003). Konold and Khalil recommend that data analysis activities for middle school students should include comparing two groups; judging the relationship between two attributes; and the understanding that as a sample grows, measures of group characteristics from that sample become more stable and thus more informative (p. 5).

Also considered important is using real data from a context that students can relate to (Gould, Kreuter, & Palmer, 2007; Langrall, Nisbet, Mooney, & Jansem, 2011). Making strong connections to the context of the data supports students to recognise that data are needed in order to make judgements and draw conclusions (Pfannkuch & Wild, 2004). Groth (2006) also supports the notion of using real data as he contends that students should be exposed to the “messy” side of data analysis (p. 46). Gould et al., however, warn that some large data sets can be unmanageable for the novice learner as cleaning the data is costly in terms of the time it takes. Although large data sets may be problematic, using real data and considering the context of the data is vital to understanding the quality of the data and the relevance of the data to the problem, in order to make sense of data distributions (Pfannkuch & Wild). The benefits of using real data are increased motivation and stimulation of statistical thinking. These benefits were reported among the outcomes of using real data sourced from the International CensusAtSchool project in a study of two classes of middle school students (sample size not reported) (Connor, Davies, & Holmes, 2006).

Pivotal to development of an understanding of covariation is the ability to look at data sets as an aggregate. It is, therefore important students are supported to make the transition from thinking about the characteristics of individual cases to looking at data sets from a global perspective (Ben-Zvi, Grafiled, & Zieffler, 2006). Cobb, McClain et al. (2003) suggest students tend to use small vertical slices of distributions to determine if there is trend evident in the data. This is achieved by treating the slices as univariate distributions, which are then compared to identify the trend and make the transition to thinking about covariation by comparing the data across a number of slices (Cobb, McClain, et al., 2003). This is similar to the method of analysis suggested by Konold (2002), whereby the mean of successive groups of data highlight the trend evident.

Technology

The use of educational technology in the classroom has the power to make teaching and learning more effective and efficient (Ministerial Council of Education, Employment, Training, and Youth Affairs, 2005). When used effectively, technologies play an integral role in the learning process and have the potential to change what concepts are explored and how they are taught (Ben-Zvi, 2000; Chance, Ben-Zvi, Garfield, & Medina, 2007). Technologies are not necessarily the focus of the learning but provide an environment for problem-solving and foster knowledge building. Software tools can enhance the teaching and learning of statistics (Biehler, 1997) but careful planning is needed to capitalise fully on their potential for establishing student-centred learning activities (Behrens, 1997; Chance et al., 2007; Finger et al., 2007; Rubin, 2007).

> Technology has also expanded the range of graphical and visualization techniques to provide powerful new ways to assist students in exploring and analyzing data and thinking about statistical ideas, allowing them to focus on interpretation of results and understanding concepts rather than on computational mechanics. (Chance et al., 2007, p. 3)

Educational technological tools such as graphing software packages, provide the opportunity for students to create and change representations and "… this may lead to exploration of important mathematical concepts" (Goos & Cretchley, 2003, p. 153). The features of graphing software enable students to move efficiently from tabular and graphical representations to the visualisation of data. They also allow students to design multiple graphical representations of the same data set quickly and more accurately than with pen and paper (Ben-Zvi, 2000; Chance et al., 2007; McGehee & Griffith, 2004). An added advantage is that interactive graphing software, such as *TinkerPlots*, maintains connections between the data and the visual representations (Konold & Higgins, 2003).

Some educational packages make large data sets more accessible and "encourage students to explore data sets in depth, to allow the data to tell the story to the student" (Chance et al., 2007, p. 9) and promote "students' active knowledge construction, by 'doing' and 'seeing' statistics" (Ben-Zvi, 2000, p. 128). To take advantage of the potential of statistical graphing technologies, all educational factors such as teaching, end goals, instructional activities, tools, and assessment need to be aligned (Bakker, 2004) in order to ensure students' development of an understanding of statistical concepts is optimised.

Sequence of Learning Experiences

In this inquiry, the Sequence of Learning Experiences focused on developing students' understanding of the big statistical idea of covariation. It integrated data analysis activities using *TinkerPlots* with appropriate pedagogies that were informed by the curriculum, previous research on development of covariation skills (Cobb, McClain et al., 2003; Garfield & Ben-Zvi, 2004; Moritz, 2004; Ross & Cousins, 1993; Zieffler & Garfield, 2009), best practice in developing statistical concepts through statistical investigations (Konold & Khalil, 2003; Ben-Zvi, 2004; Watson, 2009), understanding of student development of graph understanding (Watson & Fitzallen, 2010), and knowledge of student prior learning determined in Stage 3. It was also influenced heavily by the teaching experiment developed by Cobb and Gravemeijer (2004) and utilised by Ben-Zvi (2006). One of the aims of the teaching experiment was to progressively shift students' focus from the properties of individual cases to properties of aggregates as the sample size increased. The statistical concept explored in the former study was distribution, whereas inference was the focus of the latter study.

The fundamental notions of the teaching experiment based on using "growing samples" employed by Cobb and Gravemeijer (2004) and Ben-Zvi (2006) were applied to the exploration of covariation in this inquiry. The students were gradually introduced to sample sizes from the same population. For each sample, they were asked to make sense of it as well as learn about the features of *TinkerPlots* that could support or influence their descriptions of the data and explanations of the conclusions drawn about the sample. Students were prompted to recognise the features of distributions and variation, and compare their hypotheses regarding larger samples with their observations in the data in the smaller samples (Cobb & Gravemeijer). Data sets of large and small sample size were compared and contrasted to explore the stability of variation and spread in the varying samples (Ben-Zvi). They were also encouraged to think about how certain they were about their conjectures and conclusions drawn. In addition they were encouraged to use their knowledge of the context to support their thinking and were required to explain how the context of the data influenced their thinking (Watson, 2006).

Lesson Protocols

The Sequence of Learning Experiences developed for this inquiry was made up of six lesson protocols. Each lesson protocol was designed to encompass two 45 minute sessions and was based on data sets uploaded into *TinkerPlots.* The students' workings were saved as *TinkerPlots* files. Table 4.2 lists the sample size of each of the data sets for each lesson protocol as well as the teaching focus for each lesson protocol.

In Lesson Protocol 1 the data were generated by the students participating in the inquiry. The students measured their height, belly button height, and foot length and recorded their gender. All the data sets used by the students included these attributes. Starting with data collected by each of the students provided a foundation that allowed the students to understand the connection between the various data sets and the context of the data. Throughout the learning sequence students were asked to consider the way in which the context of the data influenced or supported decisions and statements made about the data. The students were gradually introduced to different samples of data of varying size as the Sequence of Learning Experiences progressed. The attributes of the data, however, remained the same for all data sets.

The data in Lesson Protocol 2 and 3 were generated by one of the classes from which the students participating in the inquiry were drawn (Class A). Lesson Protocol 4 used data generated from the second class (Class B) from which the students participating in the inquiry were drawn as well as the data from the Class A. In Lesson Protocol 5 the students were introduced to a data set (Class D) generated by the Australian CensusAtSchool data base using the random sampler accessed from the education section of the Australian Bureau of Statistics (www.abs.gov.au/censusatschool). The Class D data were used in conjunction with the Class B data to compare the two classes for a number of different attributes in Lesson Protocol 5. Looking at the three data sets collectively in order to compare classes was the focus of Lesson Protocol 6. The class' data were recorded as separate collections in the lesson protocol files to ensure the students understood the context of the data being explored and how they were generated. As the students worked through the lesson protocols, skills learnt and ideas explored from previous lessons were revisited. Although each lesson protocol had a particular focus, the aim was to build students' knowledge of and experience with data analysis and understanding of covariation by continually constructing various graph types and using a number of different data sets in *TinkerPlots*. Connections to the context of the data were made explicit by asking students to justify the conclusions drawn and the hypotheses made based on the context of the data and how it were generated. At the end of the six lesson protocols the students individually participated in a Student Interview that required them to complete data analysis activities related to a larger data set generated from the CensusAtSchool data base. Figure 4.7 shows the use of the different data sets across the six lesson protocols of the Sequence of Learning Experiences and the Student Interview.

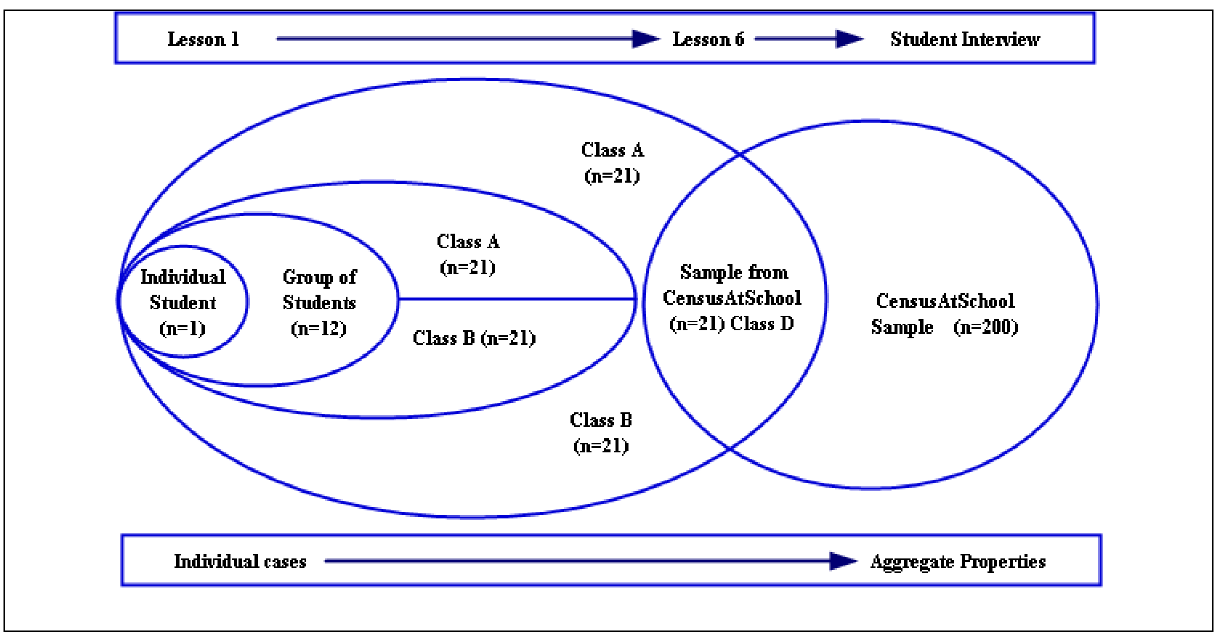


Figure 4.7. Change in sample size across the Sequence of Learning Experiences.

Table 4.2.

Sample Size and Lesson Focus for each Lesson Protocol

Lesson Protocol	Sample Size	Lesson Focus
1	Data generated by the students participating in the inquiry (n=12)	How to use the features of *TinkerPlots* and creating stacked dot plots with bins on the horizontal access initially, changing the size of the bins. Creating horizontal and vertical value bar graphs. Describing the case-based representation of data. Making connections between the data in the plot and the data on the cards. Accessing reference lines. Exploring variation in hand span. Discussing the variation in a small sample. Conjecturing about the likelihood of other samples of similar size having the same or different values based on the context of the data. Establishing how the context of the data influenced the data by considering how and from whom the data were collected.
2	Data generated by one of the students' classes – Class A (n=21)	Constructing graphs. Creating stacked dot plots and making the transition from the axis split into bins to using a continuous scale. Determining the range of the data. Describing variation and spread. Comparing values of data points. Using the mode as a reference point to describe the variation evident. Sorting data. Exploring the variation in foot length. Comparing the spread and variation evident in the graphs and suggesting reasons for difference based on the context of the data. Hypothesising about where data points would be placed in the graph if new data were collected.
3	Data generated by the students' classes – Class A (n=21) and Class B (n=21)	Constructing various graph types. Creating split stacked dot plots with bins and continuous scales on the horizontal axis. Comparing groups. Identifying outliers. Accessing hat plots. Discussing the differences between the two data sets. Using the context of the data to justify decisions about outliers and support conclusions drawn from the graphs about the data.
4	Data generated by the students' classes – Class A (n=21) and Class B (n=21)	Constructing various graph types for various attributes. Creating scatterplots. Describing the spread of the data. Identifying outliers. Removing outliers and describing the changes to the spread and distribution of the data. Using hat plots to compare groups when comparing the data from the two classes. Using the context of the data to justify decisions about outliers and support conclusions drawn from the graphs about the data.
5	Data generated from the CensusAtSchool at the same year level – Class D (n=21) and data generated by Class B (n=21)	Constructing various graph types for various attributes. Creating scatterplots. Identifying the trend. Hypothesising about the placement of additional data points based on the context of the data. Making connections between the data points and value of the data points using reference lines. Describing the relationship between height and belly button height. Comparing Class B and Class D for various attributes and using the context of the data to justify decisions.
6	Data generated by Class A (n=21), Class B (n=21), and Class D (n=21)	Looking collectively at plots from the three data sets and comparing classes using a variety of graph types. Identifying similarities and differences in the spread, distribution, mean, mode, median, and hat plots. Making suggestions for causes of variation from an understanding of the context. Thinking about the messages in the data and making inferences from the data based on the context of the data.

Figures 4.8 to 4.12 show the data sets used in each lesson protocol and a typical plot generated from the lesson protocols. For the graph in Figure 4.8 from Lesson Protocol 1, the lack of bars between 23.5cm and 26cm indicates that no student in the data set had a foot length for those values. The grey data point to the right of the graph indicates there is missing data. It shows that one of the students in the sample did not record a value for his/her foot length. In Figure 4.9 from Lesson Protocols 2 and 3, two cases in the sample did not include a measure of their hand span.

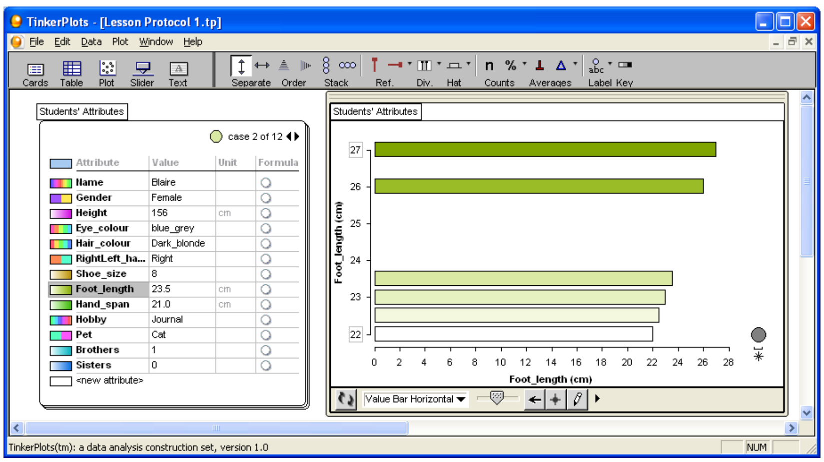


Figure 4.8. Data set and typical plot for Lesson Protocol 1.

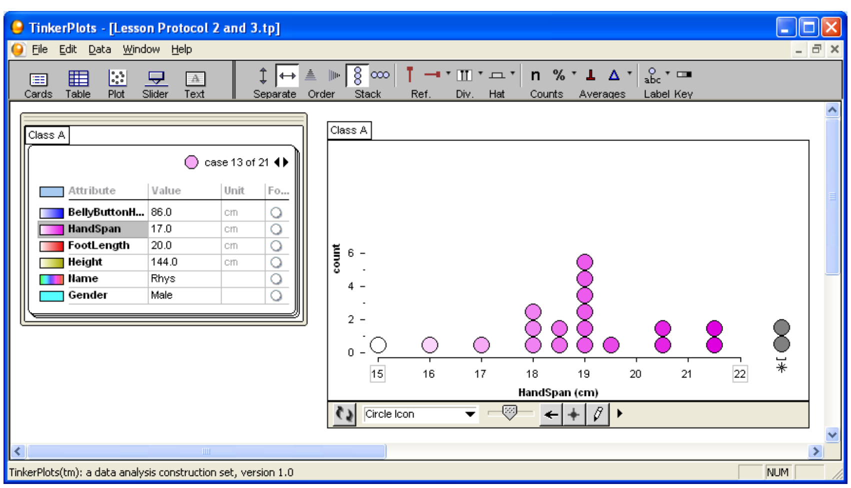


Figure 4.9. Data set and typical plot for Lesson Protocol 2 and 3.

In Figure 4.10 from Lesson Protocol 4, comparing distributions is done by comparing two different data cards rather than by adding an attribute of “Class number” and gathering all the data cards in one data stack. Although such an approach is more practical and allows for learning key aspects of data organisation and flexibility in representing data, the decision to use two separate stacks of data was to ensure the students understood the data were generated from two different samples. In making this decision the students’ lack of experience of working with data organised in data cards was considered. The aim was to ensure the students could identify with the data sets by knowing what and who it they were related to and by making direct connections between the data sets and how it they were represented in the cards and in the graphs.

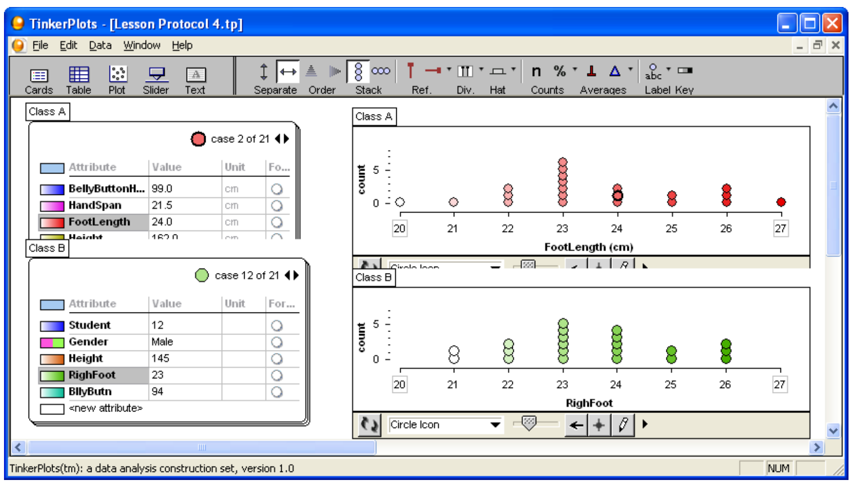


Figure 4.10. Data set and typical plot for Lesson Protocol 4.

Figure 4.11 shows vertical and horizontal reference lines added to a graph generated from the data used for Lesson Protocol 5. These reference lines were generated by *TinkerPlots* when first added to the graph, as displayed in Figure 4.11. From those positions, the students moved the reference lines to values that were of importance to them. Often the students used the reference lines to determine the value of the mean for an attribute. In Lesson Protocol 5 the reference lines were used to establish that each data point on a scatterplot was related to the values of two attributes.

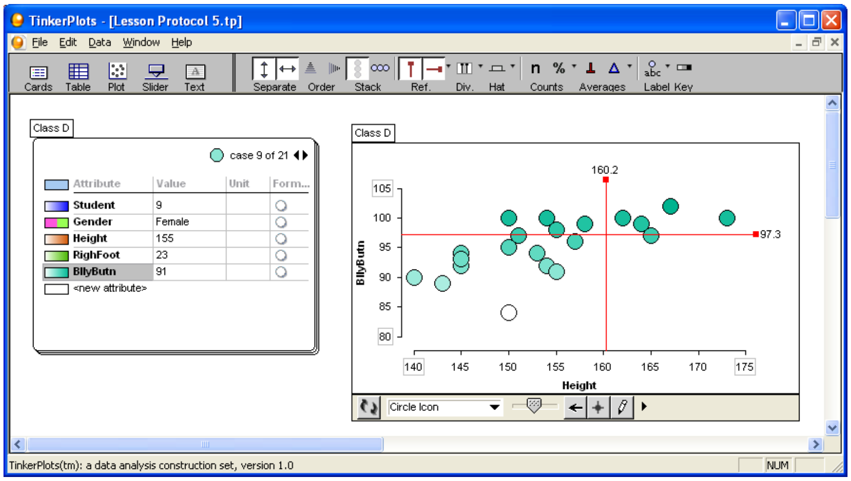


Figure 4.11. Data set and typical plot for Lesson Protocol 5.

Comparing groups was also part of Lesson Protocol 6 (Figure 4.12). This provided the opportunity for the students to explore the variation in height for females and males across three samples of the same size generated from three different sources.

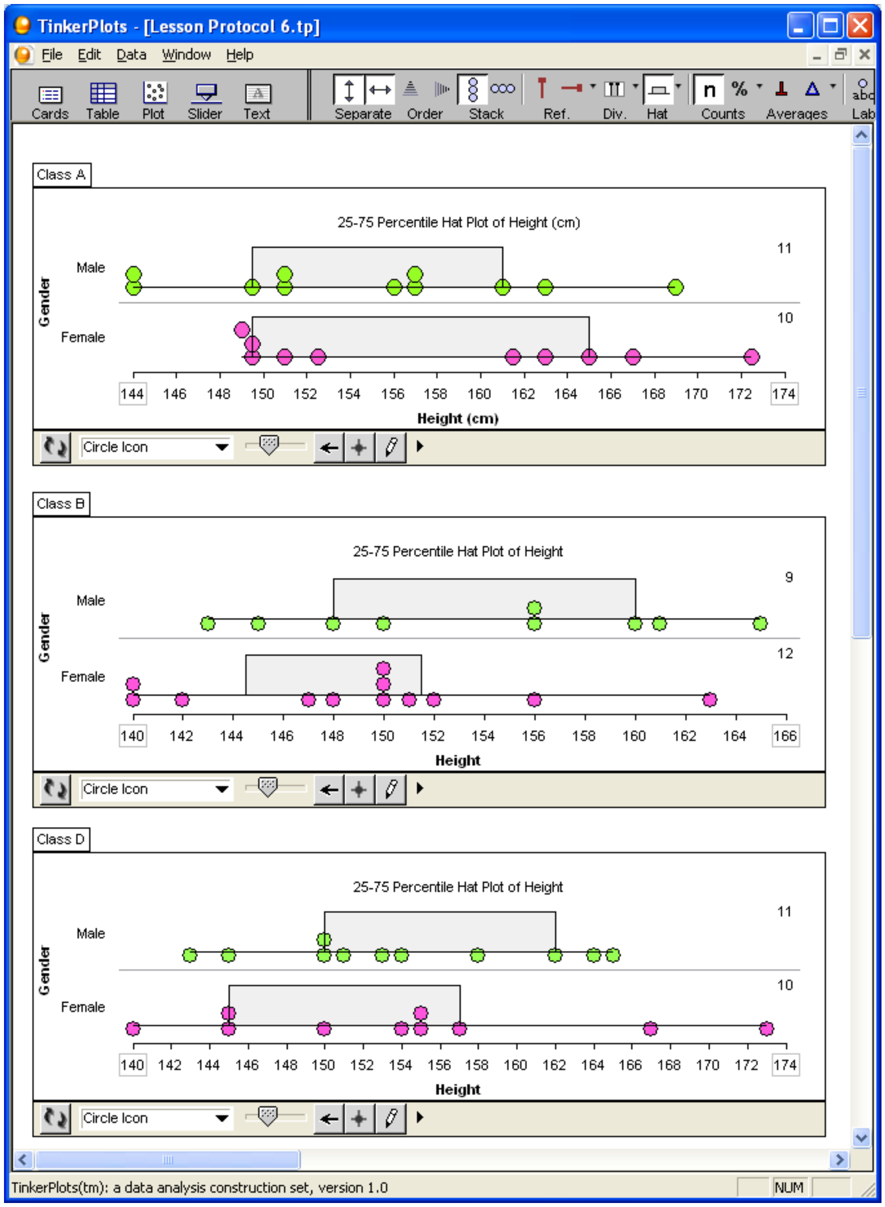


Figure 4.12. Typical plots for Lesson Protocol 6.

Implementation of the Sequence of Learning

Both pragmatists and constructivists recognise that learning is conducted in diverse settings, making it appropriate for investigating one-on-one teacher/researcher and student scenarios (Hyslop-Margison & Strobel, 2008). With respect to learning both pragmatist and constructivist researchers aim is to create learning environments that can be studied in depth (Cobb & Steffe, 1983; Steffe & Thompson, 2000). Working in more naturalistic contexts raises many issues related to methodology when the researchers are also the designers of the educational environments in which the research is conducted (Cobb, Confrey et al., 2003). There are difficulties developing effective instructional interventions that reflect objectivity and take into account the researcher's biases (O'Donnell, 2004). The advantages, however, are many. Cobb and Steffe (1983) maintain that "there is no substitute for experiencing the intimate interaction involved in teaching children" (p. 19) and assert that "the activity of exploring children's construction of mathematical knowledge must involve teaching" (p. 20), which is essentially a pragmatists stance (Hyslop-Margison & Strobel, 2008). Although this compromises the independence of the researcher and challenges the assertions that can be generated from educational design research studies (Barab & Squire, 2004), more importantly, observing and sharing students' experiences is crucial in formulating an explanation of those experiences (Cobb & Steffe, 1983).

Despite the difficulties, conducting research in close proximity with students is vitally important. The researcher not only contributes to the learning context but also experiences the students doing the mathematics over prolonged periods of time (Cobb & Steffe, 1983). Guba and Lincoln (1989) describe the importance of prolonged engagement in terms of building a rapport and getting behind "fronts" that participants may present. This is essential if researchers want to delve into the intricacies of student thinking and learning. In this inquiry the researcher acted as teacher during the implementation of the sequence of learning experiences and the administration of the Student Interviews. Any reference to the teacher should be recognised as the researcher as the teacher and researcher are one in the same person. Being able to deliver the learning experiences made it possible to interpret the students' actions as they occurred as well as ask questions to encourage and guide student learning and discover student thinking (Steffe, 1991). Of particular importance to this inquiry was the opportunity to experience first hand what students did and how they did it, when working with *TinkerPlots*.

The implementation of the sequence of learning experiences was conducted twice a week over a period of six weeks. The students worked in pairs at the one computer with the teacher guiding the students through the learning activities. The pairs of students were withdrawn from class to participate in the inquiry. The teacher worked with two students at a time. The lesson protocols, which included the data sets, instructions, and questions the students were required to complete, were set up as files in *TinkerPlots*. The sessions took 45 minutes each and were conducted in a quiet room away from the regular classroom environment. During that time the students took turns at working at the computer. While one student handled the mouse, the other student contributed to the learning experience by giving advice, asking questions of both the other student and the teacher, and making suggestions. Conversations between the students and the teacher were encouraged as were conversations between the two students. Worth mentioning is the way in which each of the students encouraged and supported the other student they were working with during the learning sessions. There were no exceptions to this. It was also commendable that all of the students approached the activities in a positive manner and engaged in the activities with enthusiasm. At the end of the implementation of the sequence of learning experiences the students worked individually with the teacher to complete a Student Interview protocol set up in *TinkerPlots*. The Student Interviews constitute the data used in this thesis. Delivery of the sequence of learning experiences was documented but analysis of the data is not included in this thesis.[1]

Student Interviews

Interviews provide insights that are valuable and go beyond the data collected in everyday conversation (Kvale, 1996) and to go even further, semi-structured interviews provide a depth of data that is difficult to gather by other means (Fontana & Frey, 2003). Semi-structured interviews provide the opportunity for ideas expressed to be clarified and expanded. In this way rich, descriptive data are collected, which is a hallmark of educational design studies (Akker et al., 2006; Cobb, Confrey et al., 2003). Silverman (2003) warns, however, that although qualitative data collection methods, such as interview, provide valuable information about "how people see things," the methods often ignore the importance of "how people do things" (p. 359).

---

[1] The method of data collection for the lesson protocols was the same as is elaborated in the Student Interview section.

In this inquiry a novel data collection approach was taken to address the concern expressed by Silverman (2003) and capture the complex nature of working at the computer to conduct data analysis activities. A strategy was developed whereby the students worked with a data set in *TinkerPlots* to create graphs and answered questions about the graphs, directed from the teacher, as they worked through the creation process. The aim was to simultaneously conduct semi-structured interviews, to draw out the students' thinking while at the same time capturing the way in which the students interacted with *TinkerPlots*, collecting the artefacts produced by the students. Gathering data collectively with these three purposes in mind recognised the value of conducting interviews as a strategy of inquiry and the importance of using student work as evidence of learning (Hyslop-Margison & Strobel, 2008).

In order to capture the students' thinking while working with *TinkerPlots*, a Student Interview protocol was developed. The Student Interview protocol was set up as an activity sheet in *TinkerPlots* and was designed to be a culminating performance of understanding (Perkins & Blythe, 1994), which provided the opportunity for the students to demonstrate what they had learned while working through the activities in the sequence of learning. The Student Interview protocol was open-ended and required the students to draw on their skills and understandings developed throughout the sequence of learning experiences. The Student Interview protocol required the students to explore a data set by constructing graphs, drawing conclusions about the data based on the graphs, and describing the relationships among attributes identified in the graphs. As the students worked at the computer, they were asked questions by the teacher that required them to describe their graphs and explain how they came to their conclusions about the data from the graphs.

As the students worked through the Student Interview protocol set up in *TinkerPlots*, another software package, *Adobe Captivate 3* [*Captivate*] (Adobe Systems Incorporated [Adobe], 2007) was used to create a video of the students' activity as they worked on the computer. The *Captivate* files generated showed the screen of the computer as seen from the user's perspective. The digital video files recorded all of the movements of the cursor on the screen as the students worked at the computer. This made available a chronological record of how and when the students accessed the features of *TinkerPlots* to construct and interpret graphs. *Captivate* also supplied audio data of the conversations the students had with the teacher during the interview. The audio data from the *Captivate* digital files provided evidence of what the students said and the on-screen video data provided evidence of what the students did.

The data set used in the Student Interview protocol was generated from the Australian CensusAtSchools data base (www.abs.gov.au/censusatschool). This provided a random sample of 200 Year 5/6 students from across Australia. The attributes in the data set were: gender, foot length, height, and belly button height. The data set was imported into a *TinkerPlots* file, which was then used by the students to complete data analysis activities and answer questions about the data. The teacher's role was to support and guide the students though the activity and to ask the students questions in order to give them the opportunity to clarify statements and elaborate on responses to questions.

The twelve students worked through the Student Interview protocol individually with the teacher. The interviews were approximately 30 – 45 minutes in length and were recorded on digital files generated on the computer by the software package *Captivate*. The Student Interview Protocol used is reproduced in Figure 4.13. The file used by the students only included the data cards and the plot window. The questions and additional information in Figure 4.13 were conveyed verbally by the teacher to the students during the sessions as the students created and interpreted graphs. The focus of the questions on the statistical objects and the software mechanics as well as the interpretation of the data was intentional in order to gather evidence about the students' interaction with *TinkerPlots* and the way in which it supported and influenced the students' development of understanding of covariation.

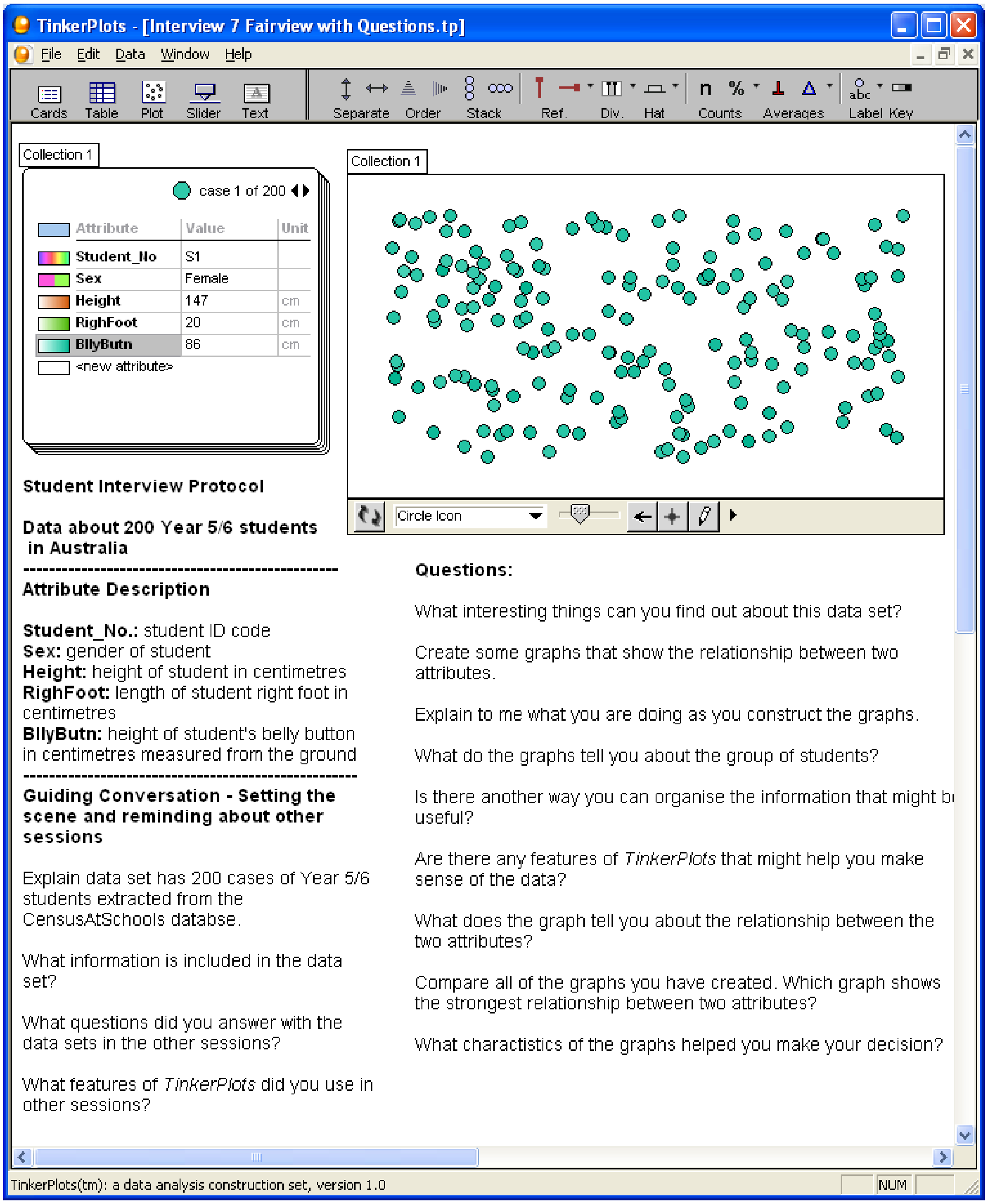


Figure 4.13. Student Interview Protocol.

Analysis of Student Interviews

At the heart of qualitative data analysis is the task of discovering themes and applying them to raw data (Ryan & Bernard, 2003). The data being analysed may take a number of forms, including interview transcripts and video footage. In thematic analysis the task of researchers is to identify a limited number of themes that adequately reflect the essence of their data. Miles and Huberman (1994) suggest that researchers start with general themes derived from reading the literature and add subthemes as they become more familiar with the data. The clustering of the themes and the subthemes forms the basis for drawing and verifying conclusions (Miles & Huberman). In keeping with the pragmatic approach taken for this inquiry, both the themes and the subthemes were drawn from the literature.

For the purposes of this inquiry, the themes used to code the data from the Student Interviews were derived from the *Model of Learning Behaviour in EDA Graphing Environments* (Figure 1.7). The dimensions of the *Model of Learning Behaviour – Generic knowledge, Being creative with data, Understanding data,* and *Thinking about data* – served as themes for the coding process and the behaviours in the dimensions of the *Model of Learning Behaviour* served as subthemes for exploring the data further. The dimensions were appropriate codes for the interview data as they were derived from the literature, as suggested by Miles and Huberman (1994).

The data for this inquiry were generated from the Student Interviews conducted at the end of the implementation of the sequence of learning experiences. The audio data from the Student Interview videos generated by *Captivate* were transcribed verbatim. The audio transcripts were then synchronised with the on-screen capture video data generated by *Captivate* and descriptors of the students' actions at the computer were added to the audio transcripts. The audio transcripts together with the descriptors of the students' actions that were added to the transcripts were coded according to the four dimensions of the *Model of Learning Behaviour in EDA Graphing Environments* (Figure 1.7). The data coding process involved line-by-line scrutiny of the interview transcripts. Following the coding process the data were grouped according to the four dimensions of the *Model of Learning Behaviour* and then the grouped responses were summarised using the behaviours of the dimensions as subthemes. The results were synthesised to construct the individual Student Profiles, which provide a comprehensive description of each student's reasoning about covariation and ways of working with data in the learning environment afforded by *TinkerPlots*, according to the four dimensions of the *Model of Learning Behaviour*.

Student Profiles

Qualitative descriptive studies demand that authentic accounts of events be given. Descriptions need to ensure a comprehensive summary of events in everyday terms is an accurate account of the meanings participants attribute to the events (Sandelowski, 2000). To avoid misrepresenting the data “better descriptions of the nature and role of context and how it interfaces with an individual … will contribute to the advancement of the study of learning and the use of design research” (O’Donnell, 2004, p. 260). In line with this recommendation, the use of descriptive data in the form of quotations by participants illustrates and substantiates research findings as descriptive data make qualitative studies credible (Patton, 2002). This inquiry takes advantage of the rich, thick, descriptive data collected from the Student Interviews and uses quotations extensively to evidence and support findings.

The use of open-ended questions in the Student Interview yielded responses that unveiled students’ thinking as they constructed graphs with *TinkerPlots*. The digital files generated by *Captivate* (Adobe, 2007) as the students worked at the computer provided video evidence of events that maintained the connection between the interview transcripts and the context of the learning environment afforded by *TinkerPlots*. Interpretation of the transcripts was facilitated by viewing the video data at the same time (Patton, 2002). In presenting the results in the Student Profiles verbatim quotations are used to convey the findings and to provide descriptions that convey the sense of the setting, thereby reflecting the experiences of the students.

In this section, the twelve Student Profiles are presented. Each of the Student Profiles includes a *Student introduction* that is followed by a summary of the student’s work according to the four dimensions of the *Model of Learning Behaviour in EDA Graphing Environments* (Figure 1.7).

Blaire

Student introduction. Blaire was very nervous when she started the interview session. This resulted in her being hesitant about getting on with the tasks required. In the beginning she asked many questions to clarify what she was asked to do. As the session progressed she found it necessary to ask rudimentary questions about *TinkerPlots* such as "So now I can make it smaller, right?" Often no response was needed from the teacher when Blaire posed a question as she made decisions for herself as part of the asking process.

Generic knowledge. Blaire's confidence grew as she worked through the tasks. It was evident she was able to access the basic features of *TinkerPlots* freely. She was able to insert a hat plot, average icon, and reference line without assistance. She stacked the data and used the drawing tool to draw lines on the graph. Blaire required assistance with higher level technical tasks such as deleting cases, swapping axes, inserting numerical values for the average, and changing the scale of graphs. She readily recognised the range of the data for particular attributes from the scale of the graphs and was able to read data values from the cards for the different attributes as well as read individual data values from the graphs.

The language Blaire used to describe features of graphs and the characteristics of the data was informal. This did not pose a problem as she was able to make herself understood by pointing to the characteristics of graphs on the screen. Her actions were accompanied by supporting statements like "But … this one goes further back that way than what this one does." Another example of the informal language she used was "Could I mix two of them together or just use one?" This question followed instructions from the teacher to create a graph that showed the relationship between two attributes of the data set. Blaire also had difficulty understanding some words used by the teacher such as "trend" and "relationship." On one occasion she asked "What do you mean by 'trend,' exactly?"

Being creative with data. Blaire created many different graphical representations to compare attributes and to look at the relationship between attributes. Over the course of the session she used multiple representations, creating stacked dot plots for one attribute and sorting by another, both vertically and horizontally, and covariation plots with either continuous or segmented scales. Figure 4.14 shows a split stacked dot plot with bins before Blaire changed the horizontal scale to be continuous.

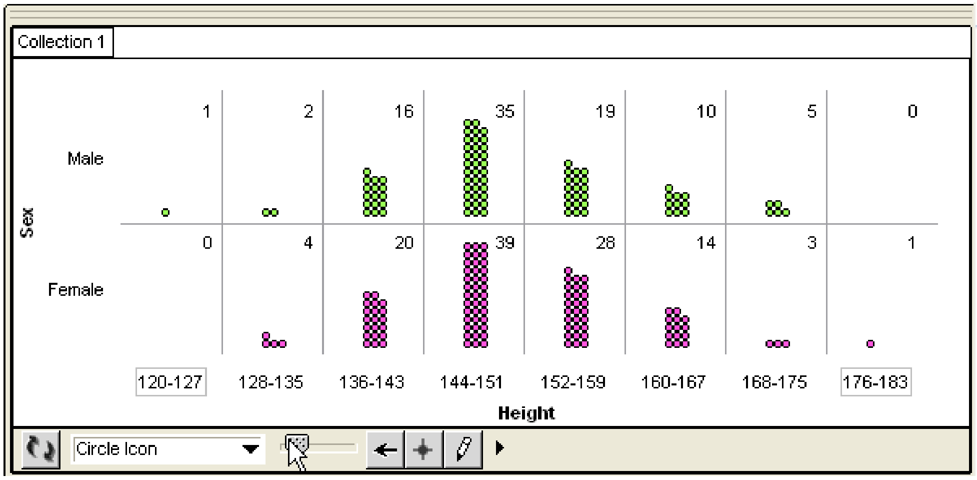


Figure 4.14. Blaire's split stacked dot plot with bins.

Blaire was not adventurous in her use of *TinkerPlots*. She often needed suggestions from the teacher to prompt her to access additional tools in *TinkerPlots* to stimulate her thinking. Questions like, "What is the range of the middle 50% for females?" and "Can you see?" were also posed to get Blaire to start thinking about the story the data may have to tell.

Identification of multiple features of the graphical representations was prevalent in Blaire's comments. She was able to describe the data in the graphs from a general perspective as well as identify individual data values from the graphs, hat plots, and the data cards. Blaire identified clusters of data and recognised outliers. She described the range of a data set by citing specific values and used the reference lines to determine the range of a hat plot. Although some prompting was provided by the teacher in the first instance, Blaire used data summaries such as the hat plot and the mean to consider the relationship between two attributes. Of the four graphs created, two were covariation graphs, one was a horizontally split stacked dot plot, and the other was a vertically split stacked dot plot. The mean was inserted on all the graphs and upon creating a new graph she deleted what she considered to be outliers.

Understanding data. Blaire's comments that showed her understanding of the data were dominated by references to the structure of the graphs, the position of the individual data values, and the application of the average. The structural features of the graphs such as clusters and outliers were of particular interest to her. She removed outliers whilst justifying her decisions by saying,

> So, we'd have to exclude that one, wouldn't we? [pause] I'll just get rid of that one, oh! [pause] Then it makes it easier to understand because then the axis went up to 80 centimetres but that one wasn't actually a real realistic one.

Although the scale of the graph changed Blaire was still able to connect with the data and make sense of the changed representation. As she worked, Blaire used the position of clusters of data to identify outliers, as well as to justify her thoughts about the average of the different attributes. This was evident when she commented, "Cos it can tell you the rough average of … yeah, of why the clusters are there in the first place" and "Well, it's roughly the average because there's most people in the area." On one occasion Blaire used a cluster of data on a graph to make a decision that was at odds with her interpretation of the hat plot placed on the graph. She stated, "Even if the hat plot doesn't actually say so, this looks like it's slightly taller [pointing to a cluster on the graph]. The average looks slightly taller on the guys than the girls."

Thinking about data. Connections to the context of the data were characterised by comments like "Of course you would already know that, as you get taller your belly button height would get higher, and with your height as well" and "That umm, it doesn't really make a difference what, what your gender is, you'll roughly have the exact same height until you get older." Although these comments show that Blaire had an intuitive understanding of the context, she also used evidence from the graphs to support her arguments. When remarking on the height data in a split stacked dot plot sorted by gender, Blaire compared the mean height of the males with the mean height of the females and determined that the males were taller. When asked what evidence supported her ideas, she indicated that her decision was based on the position of the mean on the axes, showing that she relied on the numerical value of the mean to make her decision.

When asked to describe the relationship between foot length and belly button height in a covariation graph she created, Blaire again relied on the mean and the clusters of data to support her thinking. She said,

> That their umm, foot, foot length is similar to their belly button length. Like, if they've got, say you've got 15 [foot length] and 120 centimetres [belly button height] right, in between is roughly 85, but most, yeah like most people, their foot length is in the middle, same with their belly button height.

Kimberley

Student introduction. Kimberley worked throughout the session with the same enthusiasm and eagerness that she had shown during other sessions. When she was introduced to the data set, she was surprised by the size of it. Her reaction, "Whoa!" was an indication that she had not encountered or had limited previous experience with large data sets. This proved challenging but did not deter her from completing the tasks required. Unlike Blaire, Kimberley asked very few questions. She did, however, show a sense of curiosity about the data set, asking if it was from the students at her school, from what grade, and how old the students were? By asking these questions, Kimberley made connections to the context of the data. It seemed that understanding what the data were related to was important to her. During the session, she did not ask any questions about how to use *TinkerPlots*. Nor did she ask any questions about how to create the graphs or add features other than questions like "Can I put the count in?" and "Could I please put some reference lines in?"

Generic knowledge. The lack of questions about *TinkerPlots* posed by Kimberley was due to her familiarity with the software and her understanding of the graphical representations created. It was clear that she knew what she was being asked to do and required no clarification of terms or instructions on the application of the standard features of *TinkerPlots* that she used to create graphs. She did, however, need reminding that a continuous scale was required on one of the axes when using the hat plot. Although not requested, the teacher gave Kimberley instructions on how to change the scale on an axis as a way of omitting data from a graph. Later in the session, she was able to apply this knowledge by changing the scale of the axes of two covariation graphs to be the same. This facilitated direct comparison of the graphs.

As Kimberley created the graphs she accessed many of the features of *TinkerPlots*. She added hat plots and reference lines to all of the graphs but only included the mean and the median on one occasion. Of particular interest, was the way Kimberley used the software environment to work between the graphs and the data cards. At the beginning of the session, Kimberley clicked through the data cards, checking the information on the cards. She continued to use the cards by coming back to them when she wanted to know the value of a data point on a graph. It seemed she preferred to access the information from the cards rather than read values from the scale of the axes.

Although Kimberley used hat plots to summarise the data, her descriptions about them focused on the range of the crown of the hat plot more than the position of the hat plot in relation to the spread and density of the data. Her attention was drawn to individually spread out data points and she was not inclined to focus on clusters of data like Blaire had done. Although the mean and median were added to one of the covariation graphs, Kimberley did not use them to describe the data.

Being creative with data. The size of the data set influenced the process Kimberley followed to create her graphs. Initially, she would bring down a plot window and add the attributes of interest to create a graph. She then added the count and then proceeded to drag and drop the icons to change the bin width on the two axes (Figure 4.15). Kimberley stopped regularly to evaluate how the graphs had changed and was more purposeful in using the bins than Blaire, who preferred to move quickly to using a continuous scale. Sorting the data into bins helped Kimberley identify outliers as gaps in the data were recognised easily. Her aim was to construct a representation that she could, in her words, "read." She did this by changing the size of the plot window and the size of the icons as well as stacking the data so that the data points were not overlapping. She continued to work with the graphs until they were clear and uncluttered (Figure 4.16). It was not until she got the graphs to that stage was she comfortable with describing the graphs in general terms. Kimberley was bothered by the density of the data when the data points overlapped.

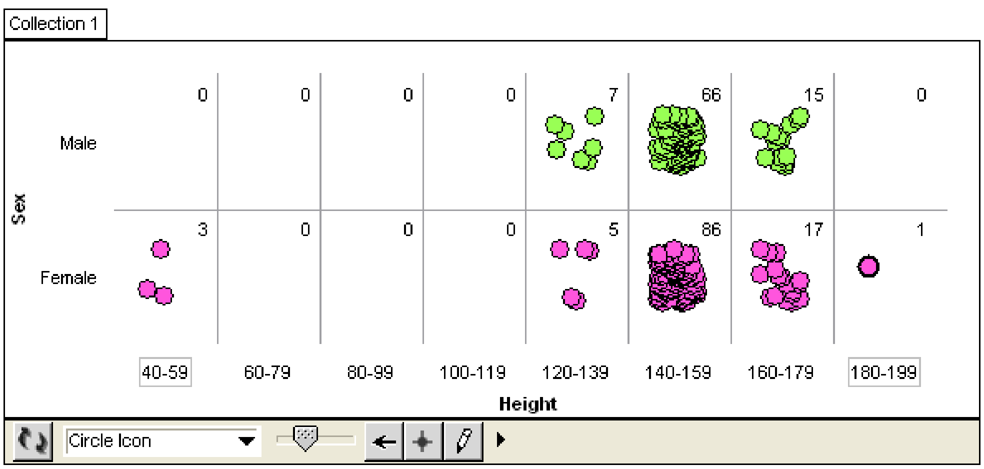


Figure 4.15. Kimberley's split stacked dot plot with bins.

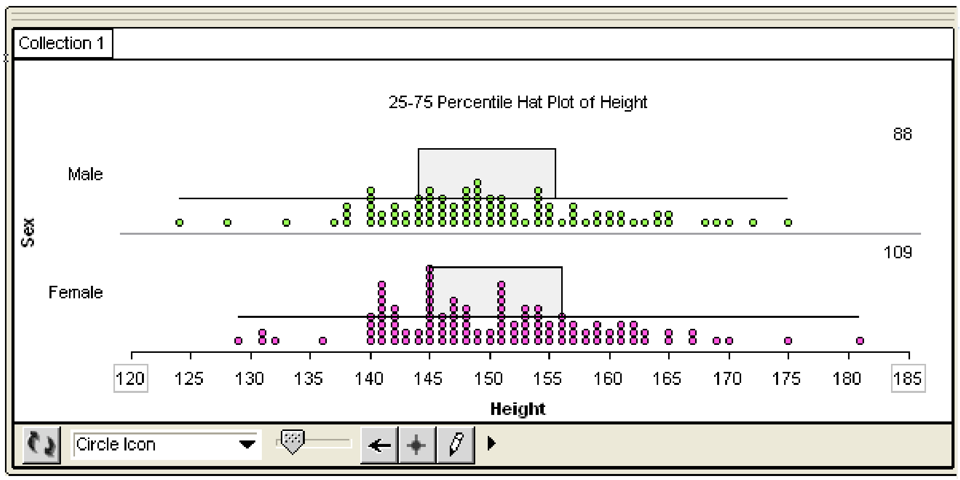


Figure 4.16. Kimberley's plot with data points separated and uncluttered.

Kimberley created two horizontally split stacked dot plots that showed the association between two attributes: one for gender and height and another for gender and belly button height. She also created two covariation graphs that showed the relationship between foot length and height and foot length and belly button height. The count, hat plots, and reference lines were added to all the graphs and outliers were deleted.

Understanding data. The first graph Kimberley created showed the association between gender and height. She changed the graph many times before settling on the representation shown in Figure 4.16. As the graph changed she was able to describe the different representations and was able to get different information from the different representations. She described the first graph generated from the default settings in

TinkerPlots (Figure 4.17) saying "Umm, out of the girls how many are from 100 to 199 measurements? Ah… oh! The males' feet are bigger! There's none [males] from zero to 99 centimetres, there's three female." After a few changes to and the removal of outliers from Figure 4.15, she used the graph to determine that the females were taller. When asked "Why are you saying the females are taller?" She replied "Because there are more up in the high area than there are in the lower area," she then went on to change the horizontal axis to a continuous scale and inserted hat plots (Figure 4.16). At this point she put in reference lines, which she used to determine the values for the range of the crown of the hat plot for both the males and the females even though it was obvious from the graph that the crown of the hat plot was the same for each gender. She determined from the measures that the range for both genders was 11cm and declared that there was no difference between the males and females height but then commented, "No, it was just that these guys were down the bottom and these girls were up the top" inferring that the males were shorter and the females were taller. Although this showed that she understood how to use the crown of the hat plot to compare the middle 50 percent of the data set, she had not demonstrated that she knew how to use the brim of the hat plot to describe the lower and upper 25 percent of the data set. Kimberley inserted hat plots on the covariation graphs but in these instances, did not use them to describe the data.

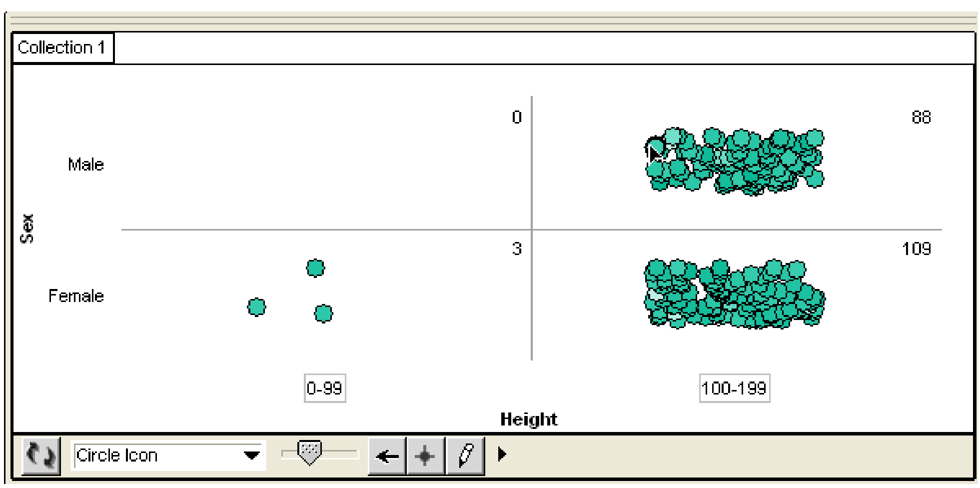


Figure 4.17. Plot generated from the default settings.

Kimberley's ability to understand a graph and use it to describe the data was influenced by its appearance. As noted previously, she was bothered by overlapping data. The graph in Figure 4.18 was greeted with the exclamation "Ergh!" Immediately, Kimberley stacked the data horizontally, added the count, and broadened the size of the plot window to create the graph in Figure 4.19. When asked if that was useful she stated "Yeh, so you can read it." But when asked what the graph showed in terms of the relationship between foot length and belly button height she conceded, "Nothing" and changed the graph to a covariation graph with continuous scales on both axes. As she did this, she went through the same process of changing the size if the icons and the plot window to separate the data points. It was evident she wanted to see each individual data point on the graphs.

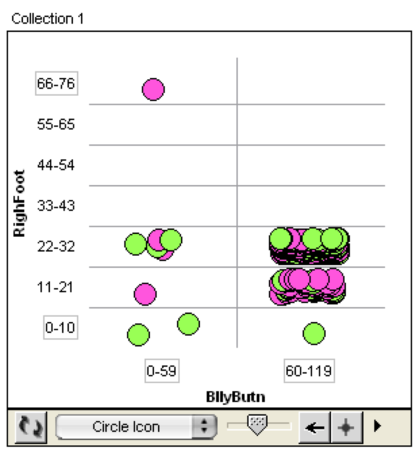


Figure 4.18. Kimberley's plot with overlapping data points.

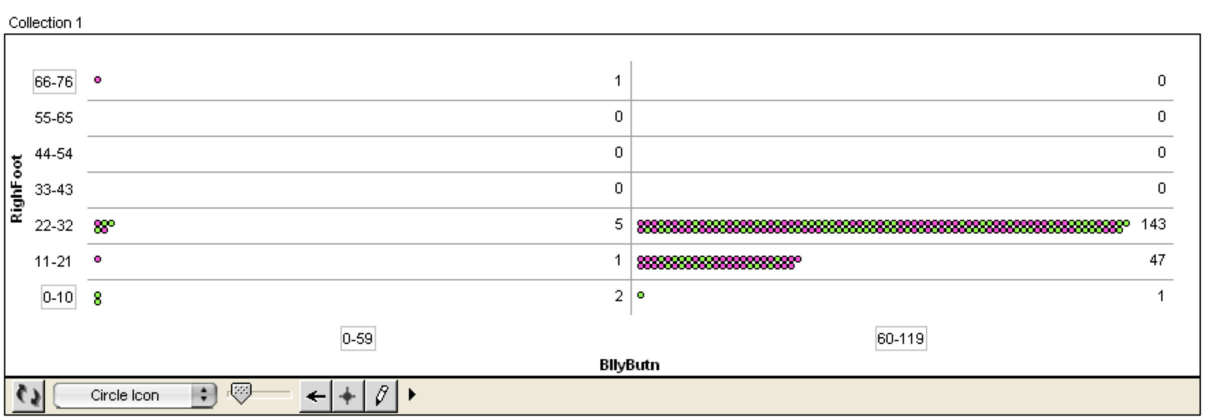


Figure 4.19. Kimberley's plot with data stacked horizontally.

Thinking about data. Although Kimberley had an intuitive sense about the context of the data, she did not always rely on that to describe the data. She allowed the data to tell their own story and was open to what they could tell her about the cases in the data set. On one occasion she stated, “Oh, the girls’ foot length is the longest! I wouldn’t have thought that!” She did, however, draw on her understanding of the context when making decisions about outliers. As she deleted a data point with a foot length value of 8cm she said, “No one I know in Grade 5/6 has that.”

From the two covariation graphs Kimberley created that looked at the relationship between foot length and belly button height and foot length and height, she determined that when the students were taller they had larger feet and when they were shorter they had smaller feet. When asked which graph showed the strongest relationship she declared that the relationship between foot length and belly button height was the same as the relationship between foot length and height. She did this by comparing the two graphs in Figure 4.20. These graphs were the only ones where she left the data points overlapping.

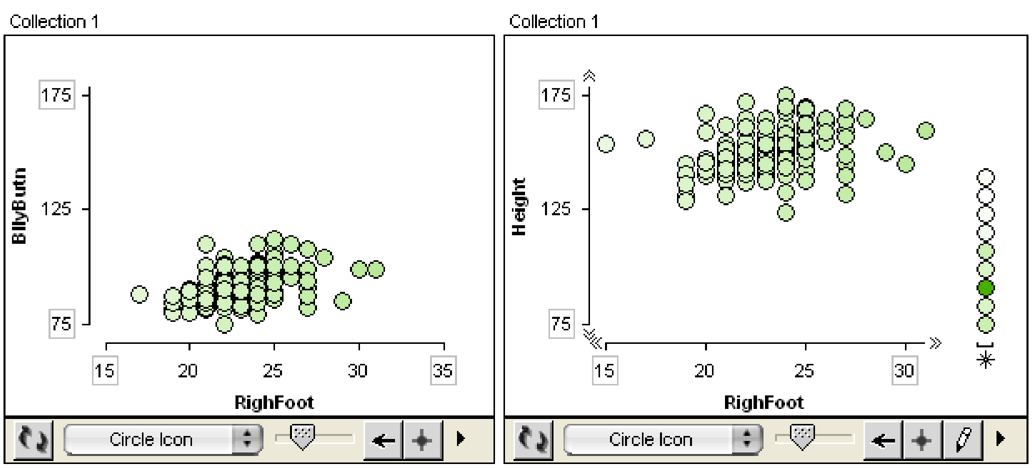


Figure 4.20. Kimberley’s covariation plots.

When considering the association between gender and belly button height, Kimberley compared the range of the crown of the hat plots as she had done previously with the gender and height graph. On this graph the crown of the hat plots for the males was broader than the females and covered the females crown completely as shown in Figure 4.21. Kimberley thought carefully about the graph saying, “Graph … [whispers “50%”] is that there appears to be … oh, that there's a bigger range of the males 50% than the females." At that point she went quiet, continued to think and then declared "[voice becomes animated] actually, yeah, because umm … [more quietly] what's it mean when they've got bigger range, yeah that … [louder again] oh, there are shorter people in the males so therefore there must be more taller females." Again, she was able to make comparisons from a graph and come up with reasonable conclusions but like before, she only drew on the crown of the hat plot to make the comparison.

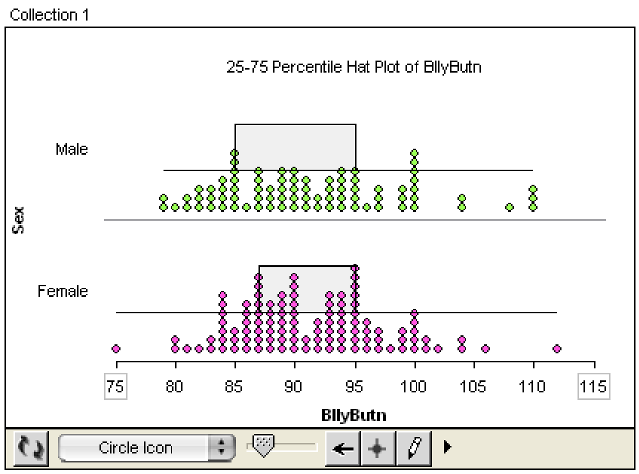


Figure 4.21. Kimberley's plot for comparing male and female belly button height.

James

Student introduction. On the day of the interview James had a heavy cold and sniffed intermittently throughout the interview. Although he was unwell, he worked consistently and completed all the tasks. During the session it became evident he was a confident user of *TinkerPlots* and had developed set routines for constructing stacked dot plots and covariation graphs. James did this efficiently and required no instructions on how to use *TinkerPlots* and did not ask for assistance with accessing the features he applied to graphs. After being introduced to the data set he did not inquire about the cases or the context of the data.

Generic knowledge. Efficient and minimalistic are the terms that characterise best the way in which James used *TinkerPlots*. When constructing graphs, his aim was to move quickly to the representation that he had in mind. As he was doing this he was able to identify when a graph was not how he wanted it and was able to make changes smoothly as he worked. The following quotes illustrate the thinking and working process adopted.

> Umm right foot and belly button [J continues working] … I need to add that into there. Now I've got right foot and right foot! [J does something to the graph] … Belly button!
>
> Yeah … [J talks to himself as he prepares the graph] … It's gone now! I don't know why it isn't continuous! Now it is! … It's gone from the bottom. That's better. Stack that one. Click that.

James identified correctly the range of the data, the mode, the mean, and described correctly what a hat plot represented. He did, however, in the initial stages, have difficulty expressing his knowledge about statistical terms and graphs correctly. When referring to covariation graphs he explained "like graphs, like two-way graphs, with the height and the foot length." At first he called the mode a "modem" and struggled to find the word to describe outliers, stating "Out … Outstand, outstanders or something." It was clear he understood the terms and the graphical representations but had not internalised the formal language used to describe them. As the session progressed and he was given assistance to establish the proper terms he used them fluently.

Being creative with data. James approached the task of creating graphs with an attitude that was direct and to the point. This was evidenced by the way he constructed graphs. Most of the time he did not take the opportunity to view the graphs at different stages as the axes made the transition from segmented bins to a continuous scale. His preference was to move quickly to a continuous scale. James did, however, construct one stacked dot plot that had bins on the vertical axis for belly button height and split by gender on the horizontal axis (Figure 4.22).

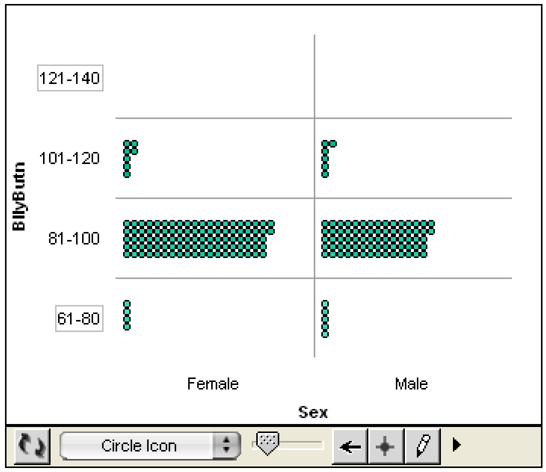


Figure 4.22. James' plot with vertical axis split into bins and the data stacked horizontally.

The first graph James constructed was a stacked dot plot for the attribute, height. To do this he inserted a plot window, added the attribute height to the horizontal axis, changed the scale of the axis to a continuous scale, changed the size of the icons, and deleted outliers. James performed these actions autonomously and applied the same procedure when constructing covariation graphs. For other features of *TinkerPlots* such as the mean and the hat plot he was more selective. He only applied these to the graphs when he thought they might be useful, unlike Kimberley, who used them just in case they proved to be useful. Like Kimberley, he added reference lines to a graph to determine the values of the range of the crown of the hat plot and changed the size of the icons to prevent the data points from overlapping. The drawing tool was the only feature that the teacher made a suggestion to use.

Although James changed the size of the data icons in the stacked dot plots until they were all separated, he did not use the same strategy for the covariation graphs. Also, he deleted all the outliers from the stacked dot plots but did not delete all of them from the covariation graphs. Instead of deleting them, James used the drawing tool to section off the data that he considered were outliers (Figure 4.23).

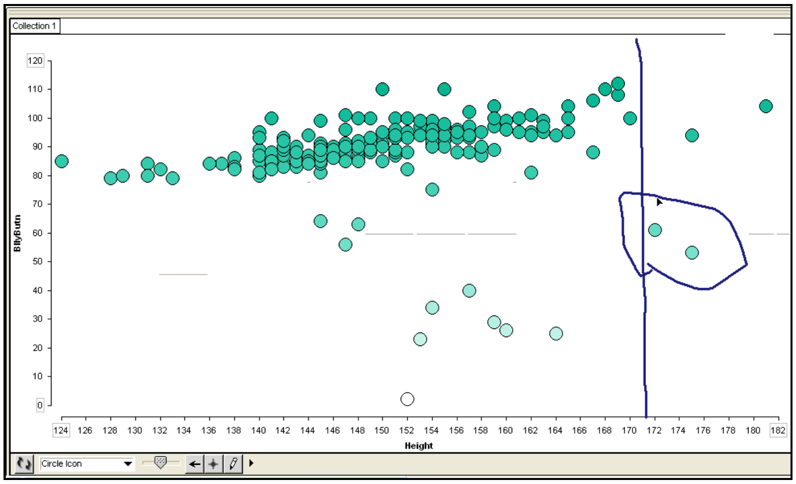


Figure 4.23. James' plot with some outliers circled.

Understanding data. The structure of the graphs James created influenced his interpretation of the data. When asked to describe the relationship between the belly button height for males and females represented in Figure 4.22 graph, James determined it was the same for both genders. He qualified this by saying that the number of males and females in each bin were much the same, therefore there was no difference. James also expressed the notion of sameness to describe the relationship between belly button height and foot length (Figure 4.24). He said,

> Umm ... they're all around about the same. Like … between 22 and 24, and 110 and, 110 and 80. Like there's … that's sort of like a pattern that goes like that [James uses drawing tool]. That, like, all, like even the area where it's not circled, it goes the same.

Further questioning on this point from the teacher elicited that, in this instance, the sameness James was referring to was related to a foot length having many different heights on the graph and that it was replicated in much the same way for each of the foot lengths within the range circled.

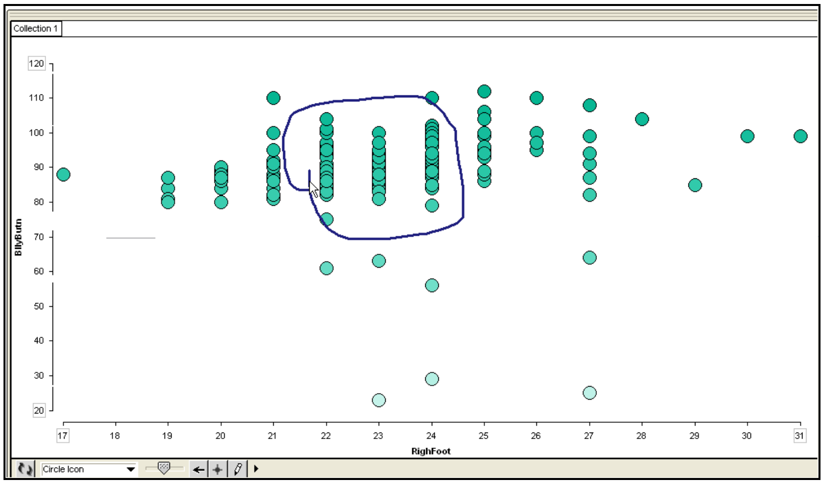


Figure 4.24. James' scatterplot with data he considered to be the "same" circled.

When asked to describe the relationship between belly button height and height, James described is as "The higher the, like … the taller the, longer. Yeah, and it's like the same ... 'cos it's kind of going up in a diagonal direction." For this graph (Figure 4.25), the idea of sameness was related to the steepness of the slope of the trend and the idea of the data going up together.

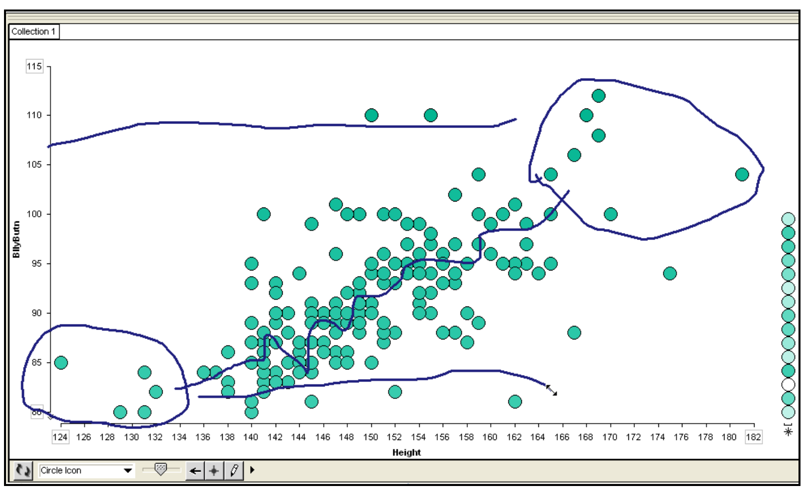


Figure 4.25. James' scatterplot with the trend identified and outliers circled.

The structure of graphs also influenced how James determined the mean of a data set. On the two occasions when James intuitively determined the mean, he came up with sensible values that were not the modes. In Blaire's case, she referred to a value being the average but actually identified the mode. For a stacked dot plot showing the distribution of the belly button height data (Figure 4.26), James determined that the mean was 92cm. When asked how he worked out the mean, he said "Because there's a lot around that area. Because at 95 you've got the … you've got like the most, the mode. And then at 90 you've got … around there you've got like the second [most]" and went on to indicate that he thought the mean would be somewhere in the middle. After James gave this explanation, he inserted the mean icon on the graph and found that the mean was 91cm.

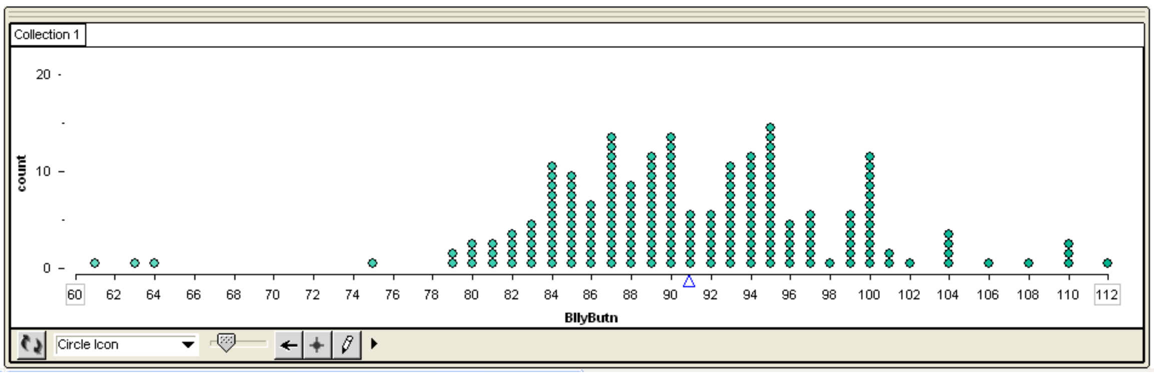


Figure 4.26. James' stacked dot plot displaying the mean.

Thinking about data. James recognised that as the belly button height of the cases increased so did the height. He determined this from the graphs and did not rely on his intuitive understanding of the context to make the decision. This was evidenced in the way he gave different interpretations for two different representations of a covariation graph of the two attributes. Figures 4.23 and 4.25 display the two graphs that showed the relationship between belly button height and height. The data set was the same for both graphs but the graph in Figure 4.23 had fewer outliers deleted from the graph than the one in Figure 4.25. From the graph in Figure 4.23, James determined that there was no relationship between the two attributes stating, "There's not much of a thing, a relationship." He qualified this by saying, "Them ones are like taller, but their belly buttons are only 50 and 60. And these ones, them two are taller … and like their belly buttons are up in the 100." James had determined that there was no relationship between the two attributes as cases with tall heights could have either short or high belly button heights.

By removing more outliers, James constructed the graph shown in Figure 4.25 before the pencil lines were added. To remove the outliers, James changed the end point of the vertical axis. This changed the distribution of the data as shown. When he was asked if the changed graph was useful to describe the relationship between the two attributes, James said that it was. He went on to say, “Because they’re shorter, they’ve got a smaller belly button [circled data in bottom left hand corner of graph]. It’s getting higher [belly button height] and it’s also getting higher [height].” As he said this, he circled the data in the top right hand corner of the graph. James then drew the line between the two circles, explaining “That, like, all, like even the area where it’s not circled is, it goes the same.” The line James drew was an indication not only that James recognised the upward trend of the relationship between the two attributes but also that he recognised the variation in the data. This was supported by comments he made about the line going up and down with the data. The two horizontal lines on the graph were drawn to show the regions of the graph that did not follow the trend indicated by the diagonal line.

The context of the data then became important to James when he compared particular data points on a graph. On one such occasion he said, “Like that seems more like what you’d think and … that would be about right. Like this one’s 115 and 170. Which would seem right.” He also drew on the context when he made judgements about outliers. An exclamation of “Umm … them ones must be babies!” followed the discovery that the foot length of a case was very small. He went on to reason that “Yeah, maybe they should have been 18 instead of 8 [cm].” This comment revealed that James had made the conclusion that some of the variation in the data was likely due to an error made when entering the data.

Jessica

Student introduction. Jessica approached the tasks with uncertainty and hesitation. Despite this, she created graphs readily and with ease but had difficulty finding the words to express what the graphs represented and the story the data had to tell. She understood what she was asked to do and used *TinkerPlots* successfully to create graphs but when the focus of the questioning moved to interpreting the graphs, she struggled to articulate her thinking. Although Jessica found the tasks challenging, she expressed an immense sense of achievement and was very proud of the work she produced. Jessica was the youngest member of the group of students in the study.

Generic knowledge. Jessica worked confidently when constructing graphs despite adopting a cautious approach when responding to questions. She constructed graphs quickly and required no assistance with accessing any of the features of *TinkerPlots*. Among her skills was the ability to create covariation graphs that showed the relationship between two attributes. She also knew how to remove outliers by changing the endpoints of the horizontal axis of the graph or by using the hide case function without assistance or prompting. Jessica was as adroit with *TinkerPlots* as James.

It was evident Jessica had not integrated the language of graphs and data into her vocabulary. As a result, she struggled constantly to convey her understanding verbally. Although it was clear that Jessica could use *TinkerPlots* well, the teacher found it necessary to encourage her to think about which features might be of assistance when she struggled to answer questions about the data. At one point, Jessica stated with great anguish, “Well, like … ’cause there’s like heaps of information there … like you probably expect … these like in this group like you probably expect them … umm … I know what I’m trying to say, I just can’t say it.”

As noted, the *TinkerPlots*’ interface posed no problems for Jessica. She recognised that the data represented in the case cards were directly related to the data in the graphs and could read data values from either representation. Interestingly, at the beginning of the session, Jessica used reference lines to determine the value of data points and did not refer to the cards at all. Figure 4.27 is an example of her application of a reference line. When asked if she knew other ways of accessing the information, Jessica indicated that she could get the same information from the case cards.

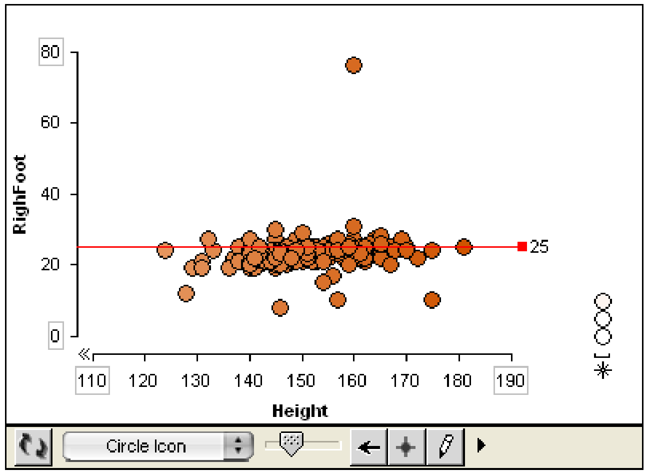


Figure 4.27. Jessica's scatterplot with a horizontal reference line.

Being creative with data. Having developed a good procedural knowledge of *TinkerPlots*, Jessica was able to create a variety of graphs. She constructed covariation graphs to explore the relationship between belly button height and foot length, and foot length and height. She also created horizontally stacked dot plots to show the association between gender and height, and gender and belly button height. Jessica added reference lines to all of the graphs to read values but only added a hat plot to two of the graphs and used the mean once.

Although Jessica could use *TinkerPlots* competently there were occasions when the teacher found it necessary to make suggestions about how she might proceed. On one such occasion, when struggling to articulate her thinking, Jessica was prompted to use the drawing tool to highlight the region of the graph she was trying to explain. She then used the tool to circle the region of the graph that she was focusing on (Figure 4.28). This enabled her to go on to say,

> Like in this group of people … most of their feet are between 19 cm and … between 19 cm and … 25 cm … for their foot length, and they're mostly … in between … like 112 cm from their belly button to the floor to … I don't know, about 82.

Using the drawing tool helped Jessica to translate her thinking into words and values that she used to explain her thinking.

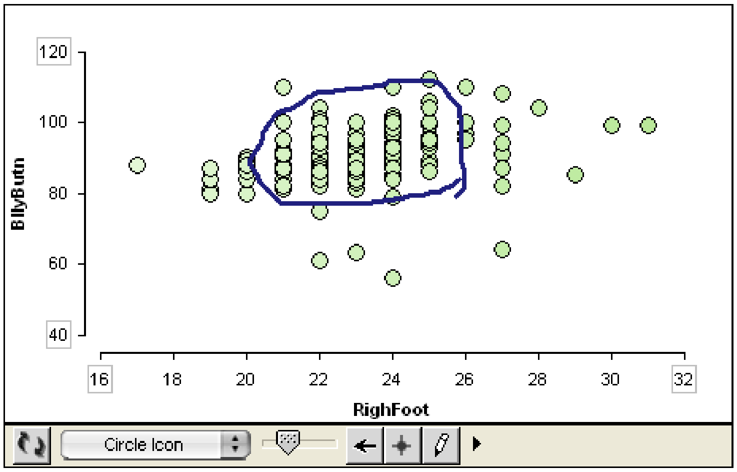


Figure 4.28. Jessica's scatterplot with the main cluster of data circled.

Understanding data. The way in which Jessica used the reference lines, hat plot and mean were similar. When she used them her focus was on the particular values that they generated and in the case of the hat plot and the mean, she did not take advantage of the potential they had to summarise the data. Jessica was able to explain adequately what the crown and the brim of the hat plot represented but focused on the endpoints of the brim when describing what they showed. Jessica commented,

> The male hat plot, the brim goes like way out there … and like, but that only goes to there, but with the female it only goes out to there 'cos that's where the smallest is and it goes out to here … there.

When asked to elaborate further on the graph in Figure 4.29, she added,

> Well, females look like they're taller, like belly button to the floor, 'cos there umm like largest is 112 cm whereas the boys is 110, so I suppose that's only 2 cm difference but with like this one here this belly button size is 63 and that's … 23 cm difference from the boys.

When she added the mean to the graph, she said, "Well, where the average is there's … 5 boys that are near the average and 2 girls that are the average as well." This demonstrates that she used the mean as a specific value and she had not established that mean was a representative value.

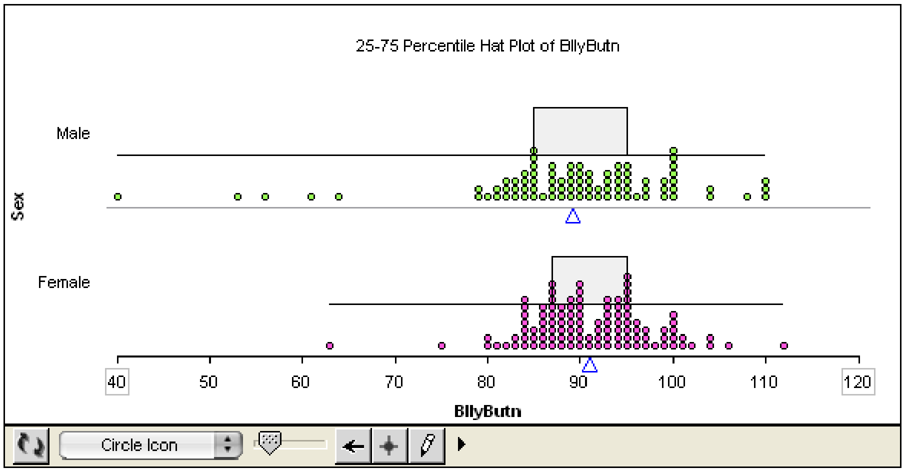


Figure 4.29. Jessica's split stacked dot plot that compares the difference between the ranges of the brims of the hat plots for male and female belly button height.

One of the graphs Jessica created displayed the association between gender and height (Figure 4.30). She went through the process of separating the gender data into the two categories on the vertical axis and then stopped to comment, "It shows that … the smallest people are all females." She then changed the horizontal axis to a continuous scale and proceeded to hide cases to remove the outliers (Figure 4.31). At that stage, Jessica was asked if she could determine if there was any difference between the heights of the males and females. She decided that they were the same and explained, "Like you kind of get a small group of girls and boys over there, with like, and they're the same height. And then you get them all bunched up over there … and then they just kind of go down." Her explanation shows that she was able to use the shape of the graphs to make her decision and indicates that she was starting to establish an appreciation of how the spread and distribution of the data are useful.

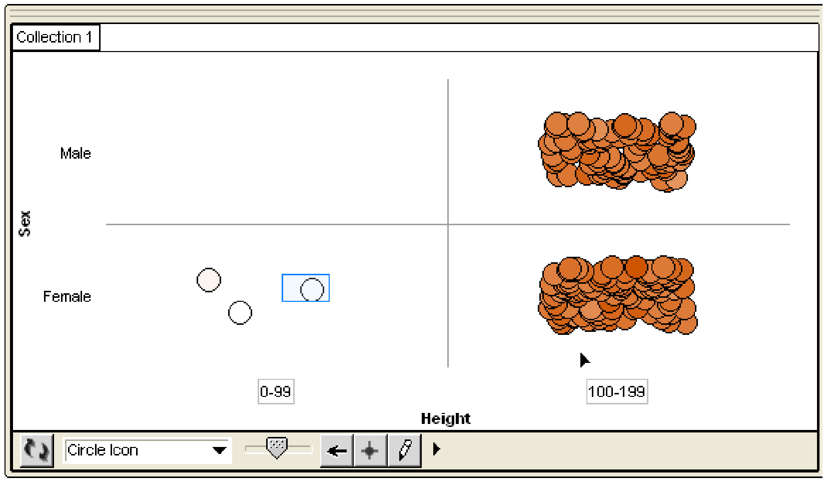


Fig. 4.30. Jessica's initial dot plot displaying gender and height.

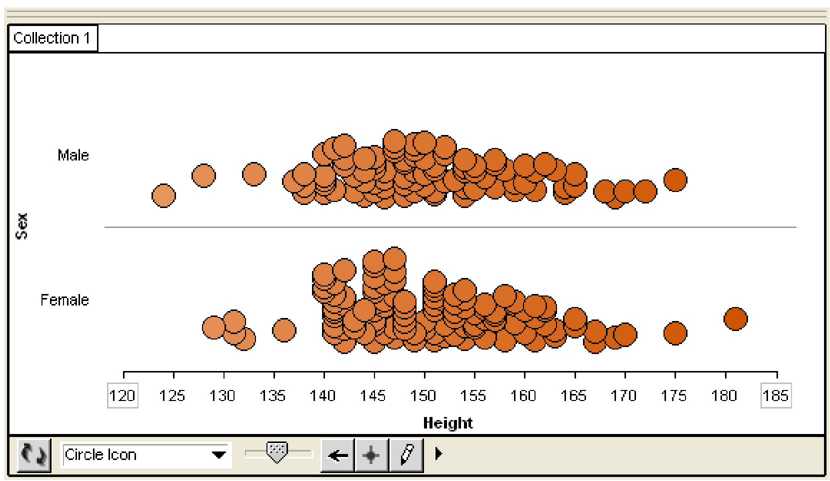


Figure 4.31. Jessica's dot plot from Figure 4.30 changed to a continuous scale on the horizontal axis and outliers hidden.

Most of the time, Jessica selected individual data points to describe graphs when she was asked to discuss the relationship between two attributes. The following comments are indicative of the way she used the information in graphs.

> There are less both female and male that actually have their belly button to the floor 59 cm or below.
>
> Like that only … one person which is a female has their umm … belly button size … like belly button to the floor between 19 cm and below.
>
> Umm … [long pause] … Well we know that there's more, both male and female, with their belly button to the floor within 80-99 cm.

When asked if a graph showed a trend, Jessica focused on the middle of the clusters similar to the way in which Blaire had done. Her descriptions included, "Like there's a big bunch of people like in like the middle, kind of in there." She was able to support this further with statements that indicated she understood what a trend was but chose to describe the story the graphs had to tell by citing specific values of data points.

Thinking about data. Interpreting the graphs was challenging for Jessica and there were not many instances where she integrated her knowledge of graph structures and types with the context to make inferences about the data. Towards the end of the session, the teacher asked Jessica to revisit a covariation graph that she created earlier (Figure 4.32). When she made the graph, Jessica talked about the middle of the group and did not talk about the attributes in relation to each other. As she viewed the graph for the second time she established there was a trend in the graph and said, "Because it's like got a straight line going up [tracing her finger diagonally across the screen]." Jessica's response to further questioning about what that meant in terms of the relationship between the attributes included,

> And … we can just see like as the people have grown, their foot length, like one of the people up here could have been that and their foot length probably would have been that and they've just gone up. That yeah, as you grow, your foot length usually grows too.

This response suggests Jessica had taken her initial sense of what the graph showed and connected it with her intuitive understanding of the context to create a story that explained an upward trend. Her explanation was feasible in terms of what is known about how people grow over time but Jessica's explanation was divorced from the data, which only reported the height and foot length measurements of students at one point in time.

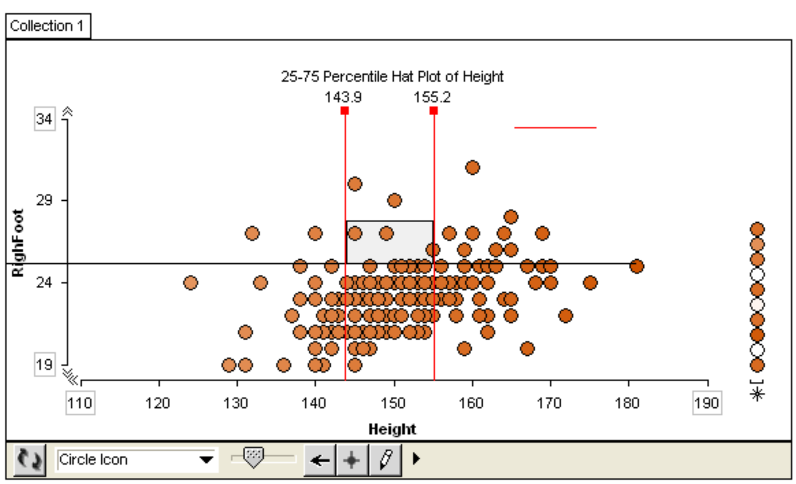


Figure 4.32. Jessica's scatterplot with a hat plot and reference lines inserted.

Jake

Student introduction. Jake worked throughout the session with a sense of uncertainty. He was tentative in his approach to using *TinkerPlots* and required many prompts from the teacher to use the software to change graphs, after their initial construction. When Jake was questioned about what the graphs showed, his uncertainty was reflected in the lengthy periods of silence, which characterised the session. In order to support Jake's interpretation of graphs and his understanding of how to use *TinkerPlots*, the teacher took the opportunity to use these instances as teaching moments and explained things such as why data icons appeared at the side of a graph when the end points of an axis were changed. The aim was to assist Jake to understand what he was asked to do to prevent him from becoming overwhelmed and unable to offer ideas and information when requested.

Generic knowledge. Dragging down a plot window and adding attributes to a graph were tasks that Jake completed easily. He also knew how to construct basic graph types, such as covariation graphs or stacked hat plots. In both cases, he chose to change the axes directly to continuous scales. After the graphs were in this initial state, Jake did not attempt to change the graphs, even when he struggled to explain what they represented. Often, the teacher made suggestions about changing the graphs. When suggestions were made, such as decreasing the size of the icons, no instruction about how to do it in *TinkerPlots* was required. Jake was familiar with how to access most of the basic the features of *TinkerPlots*, including the hat plot, but was not self-motivated to do so. He did, however, require instruction on how to remove outliers and change the end points of an axis.

Jake had mastered using the main features of *TinkerPlots* but there were times when he was challenged by what *TinkerPlots* did. His lower confidence was evident when the changes he made did not result in the representation he expected. When asked to construct a graph that showed the relationship between two attributes, he dragged the attributes gender and belly button height into a plot window. After changing the belly button height axis to a continuous scale, he tried to do the same with the gender axis (Figure 4.33). He then commented "I'm not sure whether it will work though." When the gender axis did not change beyond splitting into two bins, Jake asked "So how did we change that before?" At this point, the teacher intervened and explained to Jake why the scale on the gender axis did not look like the scale on the other axis.

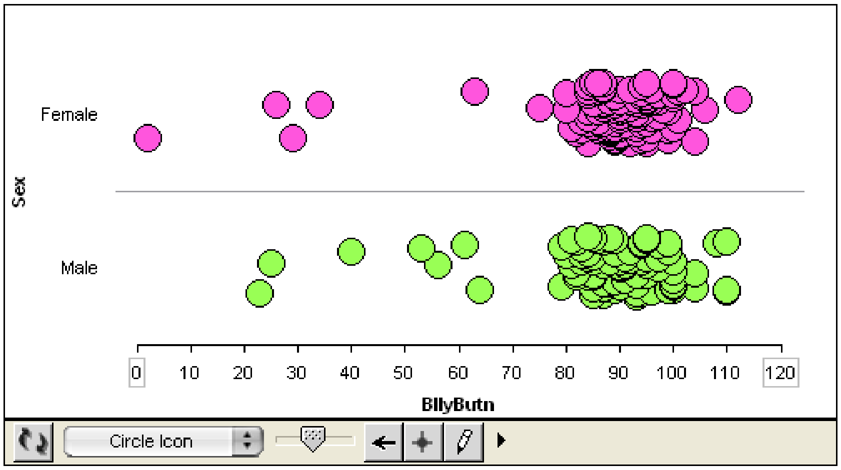


Figure 4.33. Jake's plot displaying gender and belly button height that he tried to transition into a scatterplot.

Being creative with data. Like other students, Jake was able to initiate the construction of a graph without assistance. He did, however, require much prompting from the teacher to think about how *TinkerPlots* could be used to change graphs once they were created. There were instances when Jake tried different things like stacking the data and adding a hat plot without instruction but these instances were isolated and not followed up with additional efforts to change the representation. The first graph he constructed was a stacked dot plot for the attribute belly button height. He dragged the icons to create a continuous scale and then commented, "It wouldn't really work if I stacked it, probably, because there's too many." As he said this he stacked the data. This action did not change the graph enough for Jake to be able to interpret what it showed. The left hand graph in Figure 4.34 shows the graph before the transition and the right hand graph shows after the transition had occurred. Although he was unable to discern information from the graph, he did not initiate further changes that may have been of assistance. After suggestions from the teacher, Jake changed the size of the plot window and the icons. He was then able to describe the mode, and the range of a cluster of data in the graph.

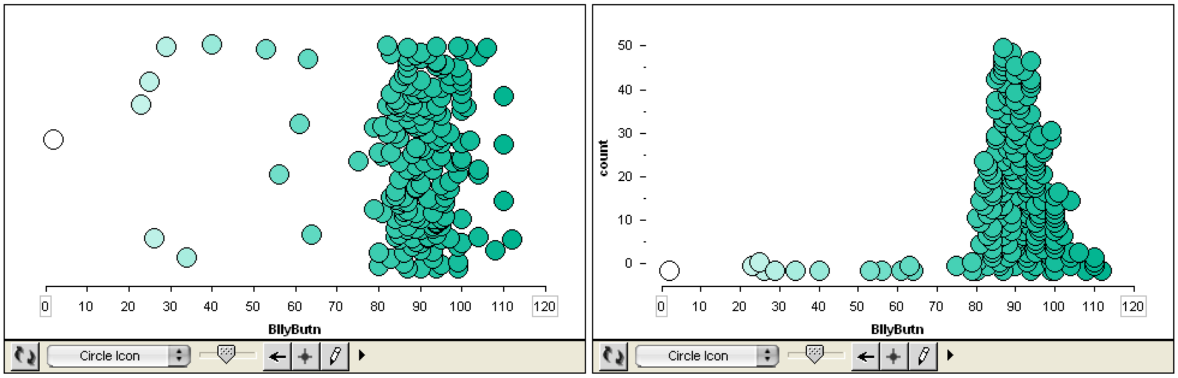


Figure 4.34. Jake's two dot plots that shows the transition from unstacked to stacked data.

Jake created another stacked dot plot for the attribute height, a covariation graph to show the relationship between foot length and height, and a horizontally split stacked dot plot for gender and belly button height. As noted, he initiated the construction of the different types without assistance but required suggestions from the teacher to access features such as the hat plot, and instructions on how to remove outliers. Jake did not access the measures of centre functions available. He did, however, determine the mode from the graphs.

Understanding data. When asked to explain what a graph showed, Jake's descriptions focused on specific values. For a stacked dot plot for the attribute height, he said, "The range is between 124 and 185, oh no, 181. Umm, the mode is 145." When asked, "What does the mode tell us?" he replied, "That's the most people, most people have that…" The value Jake identified as the mode was correct but his explanation showed that he had not established an understanding that reflected the accepted definition of the mode but had established a good sense of what it was related to.

Jake also focused on specific values when he used a hat plot to determine what a graph represented. After placing a hat plot on a stacked dot plot for the attribute height he commented, "Umm … 50% of the group are between 40, 144 and 156 centimetres. And 25% is between 124 and 144, and that one's between, the other 25% is between 156 and 181." It was clear Jake understood the structural nature of the hat plot and how it was related to the values in the graph but he was unable to express what the hat plot showed about the group of students from a global perspective.

Thinking about data. Discussions about the graphs revealed that Jake was aware that the data points in covariation and association graphs represented two attributes. His initial interpretation of the graphs demonstrated that he had intuitions about what they showed but he was unable to justify his thinking. This was evident when he commented on a covariation graph that showed the relationship between foot length and height (Figure 4.35). When asked if there was a relationship between the two attributes Jake said that there was no relationship. When questioned further he said, "Because … it doesn't really go up any more." He then went on to explain, "Hmm … the … the foot's not really getting any bigger. It's staying kind of the same." Jake's response showed that he had a preconceived idea that a relationship between two attributes was characterised by an upward trend. For this particular graph, Jake did not remove outliers.

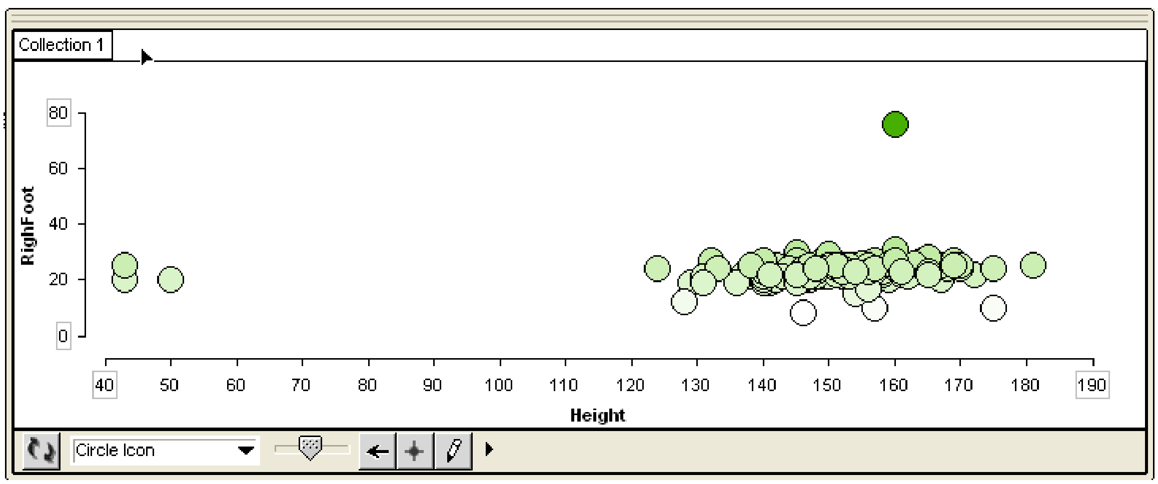


Figure 4.35. Jake's scatterplot that he determined did not show a relationship between the two attributes.

Later on in the session, Jake was asked to construct another graph that showed the relationship between two other attributes. The graph he created was for gender and belly button height, as shown previously in Figure 4.33. Following suggestions from the teacher, Jake stacked the data and then remarked, "Umm, they've got the same sort of trend. Like … they're both … I could put ... [Puts a hat plot on the graph] … they're both, their 50% like, is around about … 85 to 95, around about." It was evident Jake recognised that the middle 50% of the data for both attributes covered the same range but he struggled to articulate what this meant. His response did not connect with the context of the data nor did he draw on his previous knowledge of males' and females' heights or foot lengths to make sense of the data.

The connections Jake made with the context of the data were restricted to when he was thinking about outliers, like many other students. To a certain extent, his focus was primarily on the numerical values of data points and was devoid of context. When encountering a graph for the first time, he recognised that certain data points were isolated from the main cluster of data. From there he removed outliers from a graph when he determined that the value of individual data points was "unrealistic." These judgements were based on Jake's perception of size of the measure and were not supported by thinking about what measure would be feasible for students in Years 5 and 6.

Mitchell

Student introduction. Mitchell adopted a laid back and unassuming approach to the tasks he was asked to complete. He did, however, work confidently and quickly with *TinkerPlots* and explained with modest confidence what they represented. He instigated the application of *TinkerPlots* autonomously and required no assistance to use the features he selected.

Generic knowledge. As other students had done, Mitchell navigated the *TinkerPlots* interface easily and accessed many of the main features without instruction. He created a number of different graphs and inserted reference lines, hat plots, the mean and median, and used the drawing tool. He removed outliers from graphs by selecting and hiding individual cases as well as highlighting numerous data points and hiding the cases. Mitchell did not remove outliers by changing the endpoints of the axes.

Mitchell had no difficulty articulating his ideas and describing graphs. Most of the time he used statistical terms appropriately but occasionally used rudimentary language to describe a graph. To describe the relationship between two attributes in a covariation graph he used the words spread, outliers, and clusters. In contrast, when describing a stacked dot plot for the attribute foot length he commented, "Well, there's like a big [with emphasis] lump [shows this with his hand, arching like a hill] in around … 23." As was common for the students in the study, Mitchell referred to the mean as being the average.

Being creative with data. The way in which Mitchell worked was notable. When creating new graphs he talked to himself about what he was doing and verbalised his thinking process as he worked. He made decisions about what features of *TinkerPlots* to add and their usefulness as he worked. By working in this way, Mitchell applied a trial and error strategy that allowed him to try different things and make judgements about what they showed, as they were applied. Early on in the session while dragging the horizontal axis of a covariation graph for the attributes height and belly button height to a continuous scale, he said, "Alright. So move that down and that there ... Right? Crikey!" The "crikey" exclamation was in response to seeing the large data set in a graph for the first time. He then went on to add, "Oh … there you go, like that! Oh, you can't exactly do that can you? No. Wait a sec ..." as he added a vertical reference line and then realised he needed it to be horizontal. Later on, when constructing a graph to show the association between the attributes height and gender, he moved fluidly from making decisions about outliers to adding the mean to the graph and then immediately using the mean to draw a conclusion about the data. He commented,

> Well yeah, they're 40cm high again so … I think we'll delete them ones. And hide the cases … yeah, anyway [gives a flourish of his hand]. Yeah, umm, put the average in. They're both the same. So there's no real difference between male and female height.

Understanding data. There were times when Mitchell's explanations did not have the same timbre of confidence about them that his application of *TinkerPlots* exhibited. On occasions he offered a hesitant answer like, "They're a bit different," but then proceeded to try something else like adding a hat plot or the mean to a graph. Often, changing the graph by adding these features gave Mitchell evidence that made him more confident about his initial thoughts. In much the same way that other students had done, Mitchell used the features of *TinkerPlots* to affirm his thinking but he went further and also used them to build up his understanding of and justification for his first impressions. This enabled him to be more assertive about his conclusions.

The understanding conveyed by Mitchell, in terms of his interpretation of graphs, was based primarily on the physical location of data in the graphs. For the graph in Figure 4.36, Mitchell came to the conclusion that "Yeah, I'd say males have got … only just, but they have got a bit longer foot length … than the females." Further interrogation revealed that

Mitchell came to that conclusion because the cluster of data for the males was higher on the graph than for the females. Confirmation of this conclusion was provided when Mitchell added the mean to the graph. In response he said, "Oh yeah.… Yeah. So the male average is higher than the female average." Although the values were close together, Mitchell did not read the numerical values and based his judgement on the position of the means on the graph without considering that the difference in the means for males and females was 1.3 cm.

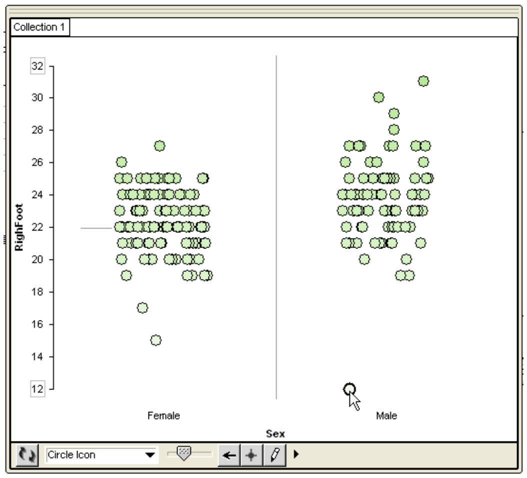


Figure 4.36. Mitchell's dot plot with data clustered.

Thinking about data. Although Mitchell's interpretation of graphs focused on the physical structure of graphs, he was able to draw together many characteristics of graphs when thinking about the data. He did not, however, use his knowledge of the context to influence his interpretations. When thinking about the graph on the left hand side of Figure 4.37, Mitchell added the mean and then commented that there was "no real difference between male and female height." He then added the median and decided that "they're both pretty much the same as well. The males are a little bit lower than the females." Following this he added a hat plot and remarked, "They're a bit different.... Yeah, so, really if you take all of them into account, the females are higher than the males." Mitchell continued to work on the graph by adding a reference line and using it to determine the range of the crown of the hat plot (right hand graph, Figure 4.37). After working out that the difference between males and females for the upper limit of the crown was 0.5cm and 1cm for the lower limit, he remarked, “So really there’s not that much of a difference.” His final conclusion concurred with his initial statement.

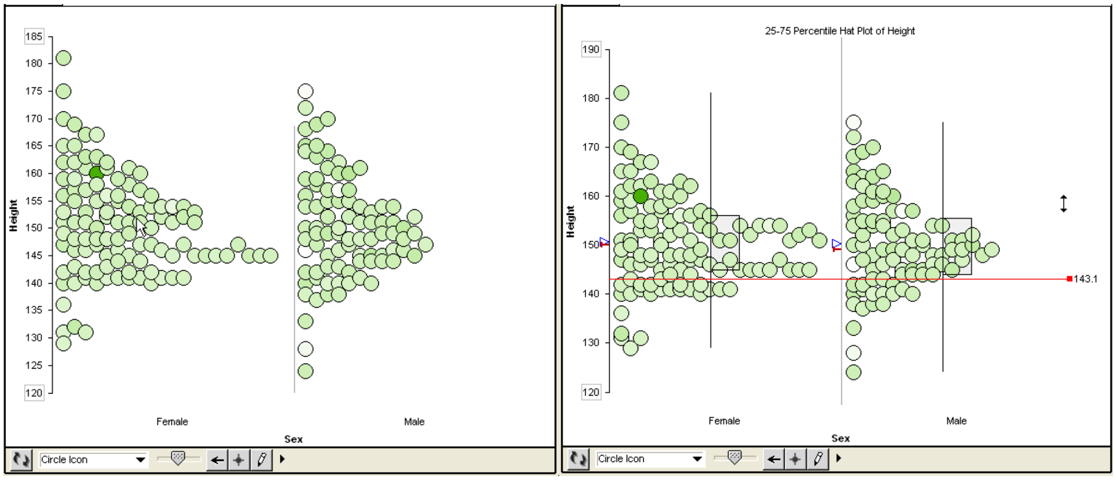


Figure 4.37. Mitchell’s initial stacked dot plot and the plot after adding the mean, median, hat plots and reference line

As well as using specific values to make decisions about what graphs represented, Mitchell was able to interpret graphs from a broader perspective. At the end of the session, four of the graphs he created were placed on the screen so that they could be viewed together. The graphs are labelled in Figure 4.38. Mitchell was asked to decide which of the graphs showed the strongest relationship between attributes. He selected the graph that showed the relationship between height and gender [Graph 1]. He contended, “Because, well, they’re pretty much both exactly the same, the two groups.” After that he was asked to look at the other three graphs and decide which of those showed the strongest relationship. The following quote was Mitchell’s response. As he went through his explanation, Mitchell pointed to the graphs he was referring to.

> Well, there’s a few outliers with these [Graph 2] so I’d probably go with that one [Graph 3]. There’s like pretty much a cluster there and a cluster there, and there’s not many, not as many outliers [Graph 3]. Not spread out as much [Graph 3]. I mean there’s a couple on this one but not as many [Graph 2]. Like, on this one they’re more spread out like. I mean there’s a big cluster in the middle and then there’s just little ones, like, outside like (Graph 3). Same as this one, there’s a line that goes through there, then there’s like little ones sticking out all over the place [Graph 4].

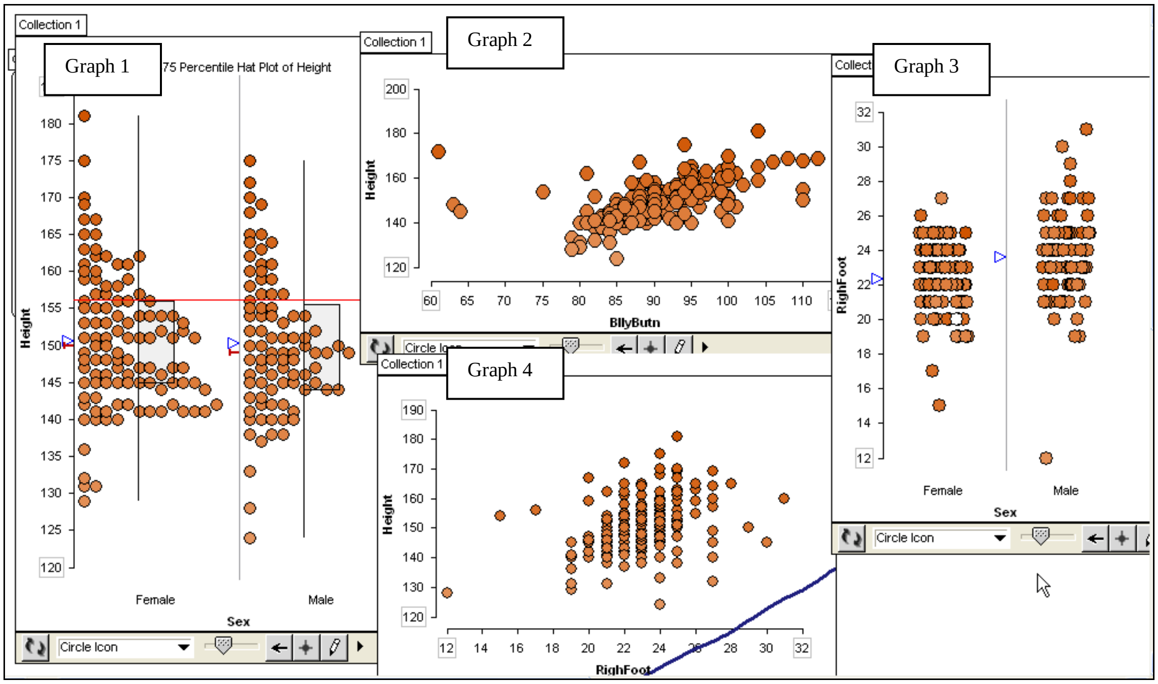


Figure 4.38. Mitchells' four plots that he compared to choose the one that showed the strongest relationship between two attributes.

Natalie

Student introduction. The session with Natalie started slowly. She showed uncertainty about using *TinkerPlots* and when describing and explaining what graphs showed. She did, however, warm to the task and develop confidence in both her actions and her explanations as the session progressed. In order to facilitate this, the teacher intervened on many occasions and made suggestions about how to use *TinkerPlots* and asked questions to stimulate her thinking. Although, Natalie found the tasks challenging, she was not adverse to the tasks and enjoyed what *TinkerPlots* had to offer. At the end of the session she made the following comment, which was unsolicited.

> You can do it with anything.... Because you've got, say if you've got homework to do and it's like based on 25% of the class wears hats, you'd have to sit down for *ages* if you didn't have *TinkerPlots*, and like 'bout for 25 or 35 minutes trying to work out everything, when you can just go on here and halve the class up into the graphs.

Generic knowledge. Natalie's understanding of the characteristics of graphs, characteristics of attributes, and statistical language was under developed. She did not use the proper names for different parts of graphs and often used statistical terms and expressions inappropriately. When asked to describe a graph, Natalie offered, “Umm, err, it’s a belly button one, it’s a line from 5 to 115.” The line Natalie was referring to was the horizontal axis. On another occasion she referred to the median icon in *TinkerPlots* as the “upside down T.” When she applied it to a graph, she did not use it to describe the data.

When constructing graphs, Natalie accessed the basic features of *TinkerPlots*. She knew that using the capabilities of *TinkerPlots* could lead to the development of a graphical representation that was useful and was able to access many of the features for that purpose. In an effort to get to a representation that was meaningful for her, Natalie tried numerous options. It was her preference to drag the axes of the graphs to continuous scales and increase the size of the plot windows in order to spread out the data. As with a number of the other students, Natalie liked the data icons to be spread out and not overlap.

Being creative with data. Most of the time, Natalie adopted a “let’s see what happens” strategy when using *TinkerPlots*. This was evident when she dragged the attribute gender into a plot window. As she did this, *TinkerPlots* split the attribute into the categories male and female, automatically. From there, Natalie tried dragging the icons to get a continuous scale on both the horizontal and vertical axes, adding the count, stacking horizontally, and stacking vertically. All these actions were to no avail as Natalie realised that the representations that resulted were not useful. After giving Natalie the opportunity to explore how she could change the graph, the teacher intervened and discussed with Natalie why categorical data similar to gender could not be displayed on a continuous scale.

At the beginning of the session, Natalie found it challenging to describe the graphs she created. She was, however, able to respond to questions posed by the teacher that directed her to think about particular parts of graphs. Her difficulty was that she could not explain what a graph showed, in a general sense, but endeavoured to focus on particular characteristics and talk about them. On one graph, she highlighted an outlier, read the data card and stated, “It stands out that the one that’s got the biggest foot is the female.” On another graph she identified the mode and added a hat plot. These characteristics gave Natalie information but she only cited values of data points and ranges along with how many data points were within the ranges. She did not use the information to discuss the graphs or the data further.

Understanding data. It was obvious that Natalie had had experience with statistical terms and expressions but had not established a good understanding of what these terms and expressions meant in order to apply them in new situations. In her explanation of a graph that showed the association between gender and foot length (Figure 4.39), Natalie indicated,

> Yeah, umm, we *can* like, talk about a difference but we … know they can't be the same because there's a different amount of people. But there can be like, a bimodal partner with them, if like, the boys have got … they've got a bimodal partner, and the girls have got bimodal partners.

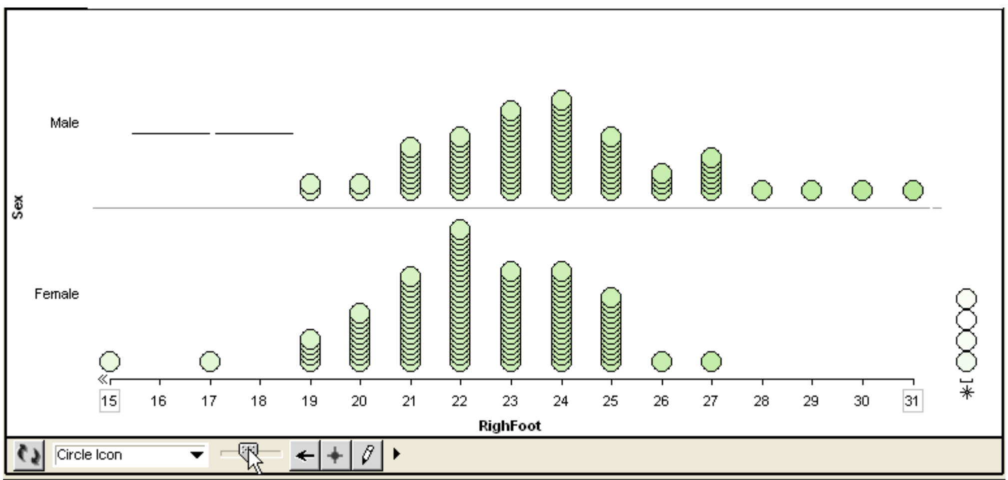


Figure 4.39. Natalie's split stacked dot plot displaying the association between gender and foot length

Natalie had confused the ideas of the mode and a graph being bimodal. In the questioning that followed the comment, Natalie identified correctly that a bimodal graph would have "two [values] with the same mode" and that the mode was "The most of like, the same amount." She explained that her bimodal partners referred to two values for foot length having the same number of data points. In her explanation she described the different levels of the stacked dot plot as separate graphs – one for the males and another for the females. She identified that the males' bimodal partners were 22cm and 25cm, and the females were 23cm and 24cm. Natalie checked her thinking by counting how many data points were in the relevant columns. For her, the recognition that some of the values on the foot length axis had the same number of data points stacked above was not related to her limited understanding of the mode. Even after being asked to identify the mode for the males and females individually and doing that successfully, Natalie did not give any indication that she realised that the bimodal partners she identified were not the mode. Her interpretation of a graph being bimodal was related to a graph having two columns of data the same height and she did not make the connection that it was related to the mode having two values.

Thinking about data. Natalie made good connections to the context of the data when she made judgements about outliers. At one stage, she commented about belly button height, "It's an outlier because of, like, you wouldn't think it would be 85 because she's 124 and I'm taller than her and mine is only 65." For two other outliers she suggested that the people taking the measurements had made an error. She identified that a 2cm belly button height meant that the person had measured the width of the belly button and justified a case with foot length of 76 cm as being an outlier by saying, "That … more, or most people's foot size is based around or most of them are like between … I'll just see … but he's probably measured his foot wrong." She then went on to explain, "Well, a ruler only goes up to 31 centimetres, so the foot would be like a ruler length and … two ruler lengths and … No, that has to be wrong."

The connections Natalie made to the context of the data focused on specific values and did not extend to helping her think about the data in broader terms. When asked if there was a relationship between two attributes in a covariation graph, Natalie responded by saying,

> These ones here [points] their height is in between 125 to 130, it cuts off here at 149 or 139, and you would think that with the height and everything, you would roughly be, it's 149 so … or 138, and their belly button height is 85, they've probably measured right up the top like about 2 centimetres over, and they're saying "that's it so, umm, we'll just write that down," but it probably isn't.

Natalie's interpretation of graphs was influenced by the aesthetic appeal of graphs. When she was asked to determine which of four graphs showed the strongest relationship between two attributes (Figure 4.40), Natalie misinterpreted the question and proceeded to describe which graph was the "best." Her justification was devoid of the context of the data and did not make reference to what the graphs showed.

> Well, like, say that you're going to like a meeting or something and you're discussing stuff, people wouldn't go for stuff like this [points to Graph 2] because it's not as spread out and as neat, but with this [points to Graph 1] you can learn more and you can see the stacks.

Natalie did not comment on the value of Graphs 3 and 4.

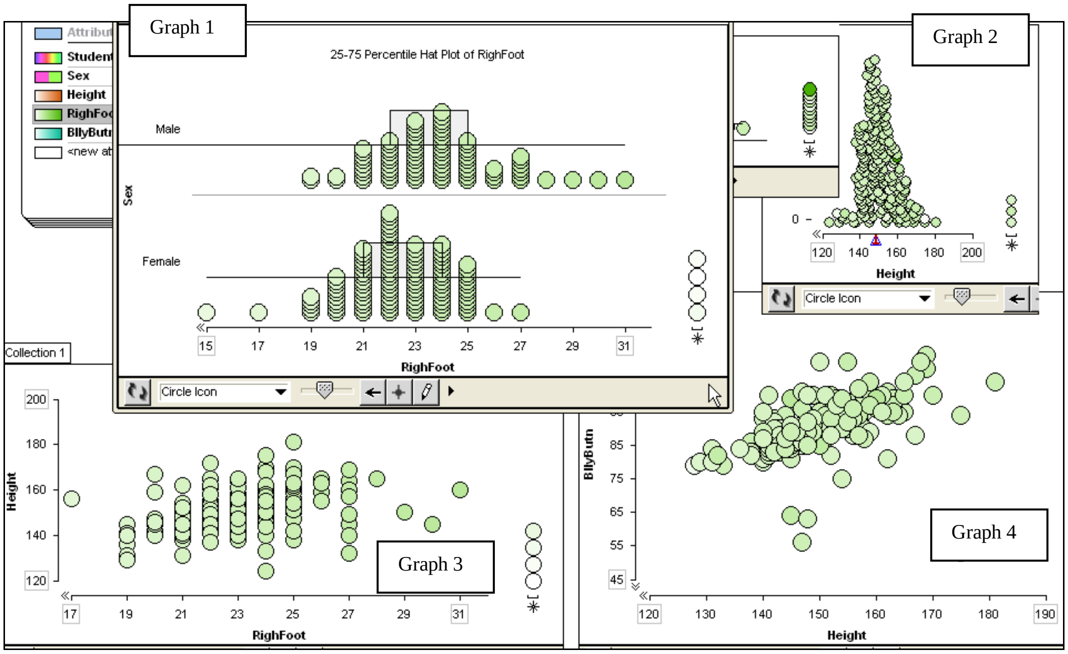


Figure 4.40. Natalie's four plots that she compared to choose the one that showed the strongest relationship between two attributes.

Rory

Student introduction. Rory worked steadily throughout the session completing all the tasks required. His approach, however, was passive and compliant. He followed instructions willingly but was not self-motivated to explore the application of *TinkerPlots* independently. Like Jake, Rory did not initiate many changes to graphs after the initial construction phase. He completed what was required and then waited for instructions or further questions from the teacher. This meant that he did what he was asked to do but he did not take the opportunity to be adventurous.

Generic knowledge. Rory had knowledge about *TinkerPlots* but lacked the confidence needed to initiate the application of its features independently. It became obvious during the session that Rory knew how to use the basic features of *TinkerPlots* to create a variety of graphs. He worked competently to create graphs with one or two attributes, change the axes to continuous scales; insert hat plots, the count, and the mean; and delete outliers by hiding cases. Often, the application of these features only occurred after prompts were given by the teacher.

Although Rory used *TinkerPlots* competently, from a skills perspective, there were some aspects of the *TinkerPlots* environment that he found confusing. When constructing a graph to look at the relationship between gender and height, Rory dragged the attribute height on to the horizontal axis of the plot window. He then wanted to change the attribute to gender so he clicked on the attribute gender on the case card. This changed the colour of the data points in the graph but did not change the attribute represented in the graph. He commented, “This looks like height ... but with different colours!” At this point the teacher explained how *TinkerPlots* could display the data from two attributes simultaneously. He then dragged gender into the plot window. This change was of no assistance to Rory but he persisted and after some thought, dragged the attribute gender onto the vertical axis and the attribute height onto the horizontal axis to create the association graph that he was aiming for. The three graphs in Figure 4.41 show the sequence of changes made by Rory.

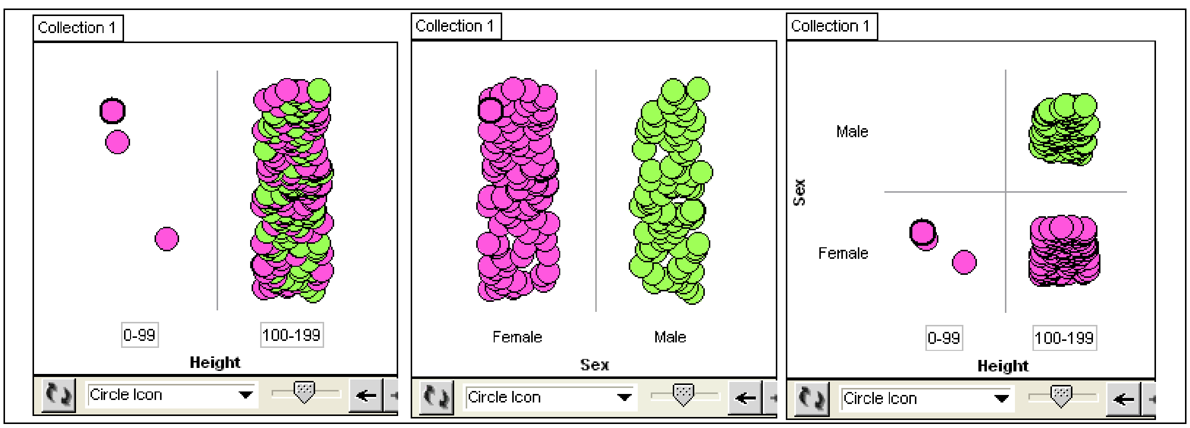


Figure 4.41. Rory’s plots that show the transition from an initial plot to a plot that displayed the attributes gender and height.

On another occasion Rory became confused when he was deleting some outliers. To do this, he highlighted a data point and then he clicked on ‘Hide unselected cases’ when he only intended hiding the case highlighted. This resulted in a graph with only one data point. Rory was surprised by what happened and did not know how to proceed. The teacher found it necessary to intervene and explain what happened as well as give instructions on how to reverse the command.

Being creative with data. Rory was demonstrated that he was a competent user of *TinkerPlots*. He constructed stacked dot plots, split stacked dot plots, and covariation graphs and accessed the features of *TinkerPlots* as described in the *Generic knowledge* section. His competence, however, did not lead him to use *TinkerPlots* creatively to interpret data. Like Jake, once Rory dragged a plot window on to the screen and created a graph, he did not attempt to change the representation without further prompts from the teacher. After creating the graph on the far right hand side of in Figure 4.41, the teacher suggested that Rory increase the size of the plot window. As he did this, he questioned "Why?" After discussing this with the teacher and viewing the resized plot window, Rory said, "That makes it easier." From that point, Rory dragged the horizontal axis to a continuous scale. He was then able to determine that "There's more girls between 160 and 199 than boys [referring to height] ... and between 120 and 159 they're equal, girls and boys."

Understanding data. Rory's apparent competence at using *TinkerPlots* belied his understanding of graphs and data. During the session he showed that he lacked an understanding of the difference between categorical and continuous data. When asked to create a graph that showed the relationship between two attributes, Rory dragged the attribute student number in to a plot window. When asked if the graph was going to be useful, he replied, "No, because I can't even understand it." He then dragged the attribute belly button height on to the vertical axis (Figure 4.42). Again, he was asked if the graph was going to be useful. Sensing that he did not have an appropriate graph, Rory replied, "But it's got belly button [height]." This indicated that he understood that he needed two attributes to make a covariation graph but did not understand the implications of choosing an attribute with categorical data that were not meaningfully related to measurement information.

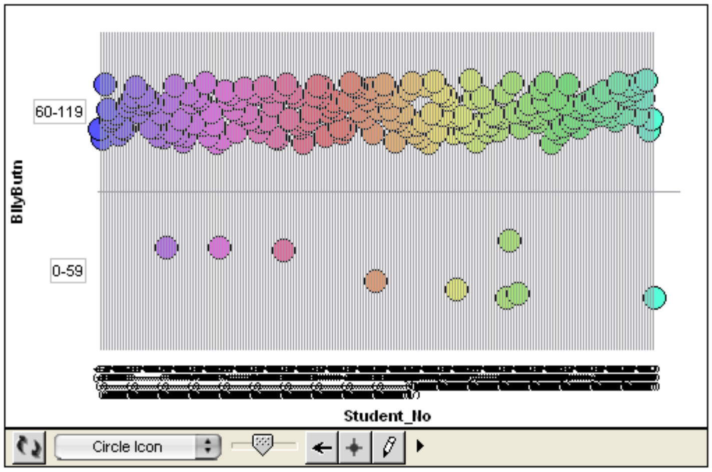


Figure 4.42. Rory's plot displaying belly button height and student number.

Thinking about data. Rory's explanations about what graphs showed focused on the number of students within a certain range or within a region on a graph. He linked the information to the context of the data and drew on his personal knowledge and experience of the context but did not use the relational nature of the covariation graphs to describe the relationships between two attributes. This was evidenced when he was asked if he could identify if a covariation graph showed there was a relationship between the students' height and belly button height. He replied, "Yes, there's a whole heap in there [points to the data cluster in the middle of the graph in Figure 4.43]. There's more than any [other] spot." The teacher then asked Rory to use the drawing tool to indicate on the graph what he was referring to. Rory used the tool to draw the diagonal line on the graph. He then described the relationship between the two attributes as,

> The smallest, smallest person's height is probably about 121 and then the biggest belly button, from the belly button down is 111, so there's … 10 cm between there so their legs is almost as big as a person.

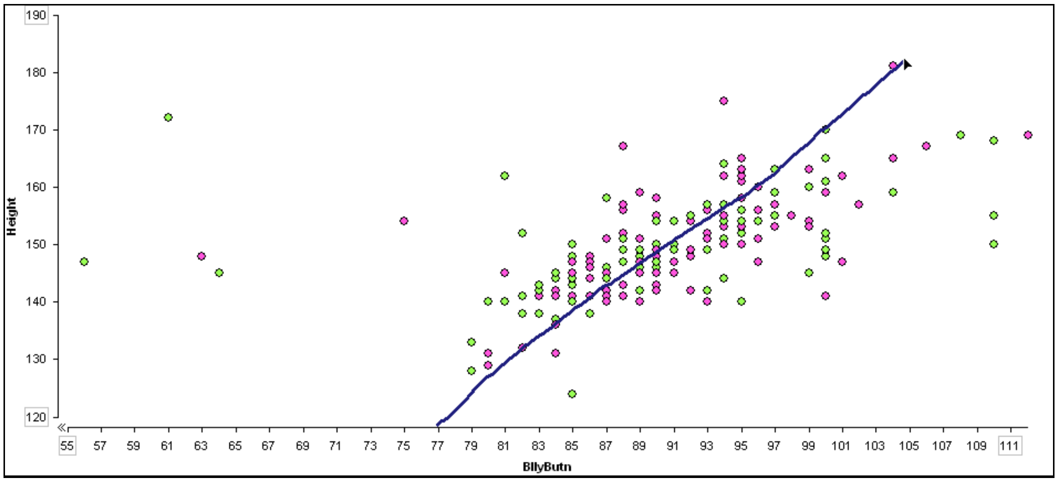


Figure 4.43. Rory's scatterplot with the trend drawn.

Although no data point existed on the graph that showed the relationship Rory described, he had made the connection that one person with a particular belly button would have long legs when compared to another person who was not very tall. After this comment the teacher continued to question Rory to see if he could discuss the relationship between the two attributes in more general terms. As a result, Rory stated unconfidently, "belly button height is umm, most people that are, like, small, probably have a small belly button height." He then declared, "But, look, this one's got 61 and he's 157." This showed that although he was able to make a generalised statement that was reasonable he still used particular data points to support or question his thinking. Rory used the same sort of reasoning when he was describing the relationship between foot length and height. His explanation included: "Umm, it all depends. Because some people might have, be real tall and they might have a, a like medium sized foot and not that big." He then went on to say about the relationship, "Not really, because the third highest ... if you've got a big foot then you're tall [to show a trend], because the third tallest person, he had a 10 cm foot."

Rory's interpretation of which of his graphs showed the strongest relationship between two attributes was the same as Mitchell. He had created two covariation graphs that looked at the relationship between belly button height and height, and between foot length and height. He also created an association graph for gender and height. Of these, Rory selected the gender and height association graph. He argued, "Well the hat plots, the hat plots look exactly the same! They are! And so's the average." Rory interpreted the sameness of these characteristics as showing a strong relationship between attributes. This was another occasion where Rory used specific values to justify his thinking.

Shaun

Introduction. Shaun was not at school when the other students completed their interviews. His interview took place two weeks later, after the second term school holidays. This may have contributed to the way he worked through the session. Although it was evident he put time into thinking carefully about what he was asked to do, his actions and responses were not as automatic and as fluid as the other students had demonstrated. This made him more purposeful in the way he worked but also caused him to experience moments of frustration when he was unable to answer questions or follow through on a prompt from the teacher. Despite these issues, Shaun was happy with his achievements. At the end of the session the teacher asked him if he felt he had learnt something by participating in the project. Shaun said he had learnt how to use *TinkerPlots* and then added, “Usually I’ve never got to make graphs. I’ve only had to answer questions and that.”

Generic knowledge. As with the other students, Shaun was able to access the main features of *TinkerPlots* to construct a variety of graphs. At the beginning of the session he showed that he understood the representation offered by the data cards by citing the attributes and some examples of values. This was the only time he referred to the cards. On other occasions when he wanted to determine the value of a data point he used a number of strategies to read the information from the graphs. These included reading from the scale of the axes, inserting reference lines, the mean, or the median. These features provided the information he was looking for and hence using the data cards was unnecessary.

Although Shaun added the mean and median to three graphs, he did not use them to interpret the data. He did, however, cite their values. When he added the median on a graph for the first time he was asked what the icon represented. He noted the name correctly but then struggled to find the words to describe what it showed. At first he took some thinking time and said, “Isn’t that the most common or something?” Realising the teacher did not agree with the description, he offered, “The most around ... [He smiled and shook his head].” With encouragement from the teacher and after a pause he added, “The middle or something.” His responses did not resonate with confidence and his uncertainty was evidenced in the way he worked down the list of what he knew about measures of centre until he got the right one, which happened to be the last possibility. Shaun was able to identify the mode correctly on a graph and describe it correctly, though on two occasions he referred to it as the “peak” of a graph. Questioning from the teacher revealed that he was talking about the mode.

Being creative with data. Shaun worked hard during the session, producing seven graphs. He created two covariation graphs, one scatterplot, one stacked dot plot, and two horizontally split stacked dot plots. He inserted hat plots on all the stacked dot plots but did not use them to look at the covariation graphs. The other graph he created was a scatterplot with the attribute height on both axes. As mentioned previously, he inserted measures of centre on some of the graphs but only applied them to the stacked dot plots. Of the measures of centre used, Shaun only used the mode to interpret the data.

Although all of the students removed outliers from the graphs, Shaun was the only one who commented on the changes to the graphs that resulted. As he was deleting outliers from a stacked dot plot he commented, “It’s made the graph … It’s just shown us the middle … like where all the others are. It’s made the graph smaller.... No, I mean bigger. The numbers are bigger … it’s hard to explain.” He qualified this by saying, “Because the other ones are gone it’s taken out all the numbers in between those … like all the numbers in between the outliers and like the smallest one,” indicating that he understood removing the outliers had changed the end points of the axis to accommodate the remaining data.

Shaun had developed an intuitive sense about the spread of data. When he was comparing the belly button height of students sorted by gender in a stacked dot plot, he said, “It’s roughly the same [for males and females]. They’ve got the same amount of people roughly about 50% and roughly about the same … outside the 50%.” At this point he added a hat plot to the graph. He then commented as he pointed to regions on the graph, “It wasn’t what I expected but it shows it a little bit more. I thought that it, like this, would be out to round here and out to there. Same with that.” Although Shaun’s approximation of the middle 50% of the data covered a broader range than what was determined by the hat plot, he used the hat plot to confirm his conjecture and recognised it provided good evidence to support his interpretation of the data.

Understanding data. Determining Shaun's understanding of the graphs he created required much probing and persistent questioning. Most of the time, his responses were short and succinct, and offered little insight into his understanding. This was evident at the beginning of the session when he created a graph with the attribute height on both axes, as shown in Figure 4.44. He described the graph as "Umm, a height graph." When asked to elaborate further he confirmed it was "just a height graph." To determine if Shaun understood what the graph represented, the teacher probed more deeply by asking if there was anything special about the graph. He replied saying "It goes up in a line." The discussion that ensued revealed that he understood that the direct correlation of the data in the graph was due to having the same attribute on both axes. Although it was his intention only to include the attribute height in the graph, he did not indicate why he chose to construct it as a scatterplot.

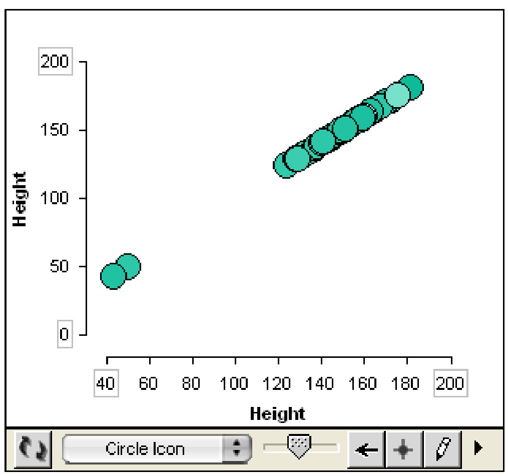


Figure 4.44. Shaun's scatterplot with the same attribute on both axes.

Thinking about data. Shaun demonstrated on a number of occasions that he understood that the data points on a covariation graph represented data for two attributes. At one stage, he had constructed a covariation graph and tried to stack the data. He realised immediately that this was not possible saying, "Because it can't … because it's got two … it's got to do two sets of information. It can't actually stack." Supporting this further was his explanation about the position of the data points in the graph in Figure 4.44, "Because I've got height there and height there [points at different parts of the graph on the screen]." Later on, when he was justifying why a graph only showed there was "sort of a relationship" between two attributes he commented, "Because there's a foot length of 14 and his height is around 155 so it should be like … if he's got 155 he should be up around here somewhere." His speculation about what the foot length needed to be for the data to show a relationship showed that he had made the connection between the individual values and the relationship between the attributes.

Like other students Shaun identified the relationship between two attributes. All of his comments about such relationships, however, were offered with a cautionary note. For example, "In some cases that there was and some cases there wasn't" and "Sort of a relationship. Yes… same as last time … belly button floor … yeah, I think so" and "Umm … most of the people the higher the foot they have the higher the belly button to the floor." This showed that he recognised the variation in the data that made the declaration about a relationship uncertain and questionable.

Natasha

Student introduction. Natasha started the session slowly. She struggled to recall how to use *TinkerPlots* and was reticent in providing answers to questions about the interpretation of graphs. She was uncomfortable revealing her thinking, often saying, "I don't know." On these occasions, additional questions or prompting from the teacher on how to move forward with a task resulted in Natasha providing more information but she still had difficulty articulating her knowledge and thinking. As the session progressed, Natasha became more comfortable working with *TinkerPlots* and required less assistance but remained restrained in her responses to questions about the data.

Generic knowledge. Like Jessica, Natasha had difficulty articulating her ideas about graphs and statistical ideas. She did not, however, display the same expertise with *TinkerPlots* that Jessica did. Although she faltered in the beginning, she knew how to create basic graphs and had no trouble accessing the features in the menus or using the drag-and-drop facility of the software when requested to do so. She read data values from the cards and the graphs and showed she understood the points on the graphs were a representation of the data in the cards.

During the session Natasha did not question what she was asked to do. She did not ask for an explanation of statistical terms like "trend" or "relationship" and gave reasonable answers to questions involving these key ideas. From these conversations it could be construed that Natasha understood the language that was used. Further interrogation of her thinking revealed that although she had a basic idea of the terms and ideas, her understanding was limited. An example is provided in the *Thinking about data* section.

Being creative with data. In the beginning, Natasha relied heavily on directions from the teacher to assist her to create graphs. As she worked, she made it apparent that she did not want to make errors. She sought reassurance and confirmation that she was doing correctly what she was asked to do. This was evidenced in the way she asked questions like, "OK. Umm … drag them to make a stack plot?" after opening a plot window and "Drag that out? [size of plot window] No? Stack it? … O-Kay!" when asked how the graph could be changed to make it easier to understand. Although Natasha lacked the confidence needed to proceed independently, she worked cooperatively and tried hard to complete the tasks.

The first graph Natasha constructed was a stacked dot plot with a hat plot for the attribute belly button height. She used the graph to determine the ranges of the data under the crown and under the brim of the hat plot but was unable to talk about the graph in terms of the spread of the data. She also created a covariation graph for the attributes belly button height and foot length and another for the attributes height and foot length. The final graph she constructed was a horizontally split stacked dot plot for the attributes gender and height. She added the count and hat plots to the graph and coloured the data by gender. Natasha did not access the mean or the median.

Understanding data. Natasha's understanding of what covariation graphs represented was limited. She knew how to drag two attributes into a plot window to create a covariation graph but only cited values of data points from one of the attributes when she described what information was in the graph. Her focus was on the attribute represented by the horizontal axis. She did, however, talk about the attributes collectively when describing the relationship between the two attributes. It was not evident that she used the graphs to make these determinations.

Although able to identify particular information from graphs, Natasha had difficulty drawing information together to explain her ideas. From the horizontally split stacked dot plot for the attributes gender and height, Natasha determined that the middle 50% of the males and females for the attribute height were the same, because the crowns for each gender covered the same range of data. She then went on to say that “Girls are more taller than boys. Well, in this case anyway.” When reminded that she had determined that the middles of the data for each gender were the same, she became confused and could not explain her thinking. It was evident that she understood the middle 50% of the data for each gender was the same but based her decision about girls being taller than boys on the individual data points in the upper 25% of the graph.

Thinking about data. Natasha drew on her understanding of the context to make statements about the relationship between two attributes. When asked to describe the relationship between the attributes height and foot length in a covariation graph she stated, “the smaller you are the smaller your foot length is.” In order to determine if Natasha used the graph to come to this conclusion, the teacher suggested that she use the drawing tool to show the trend that she identified. The circled region on the graph in Figure 4.45 was the addition Natasha made to the graph. Natasha’s initial response showed that she understood the basic sense of what she was asked about the relationship between two attributes but drawing the trend in the way that she did revealed that she did not understand fully how the graph contributed to her statement. It also revealed that she did not have the understanding of the term “trend” that was anticipated by the teacher.

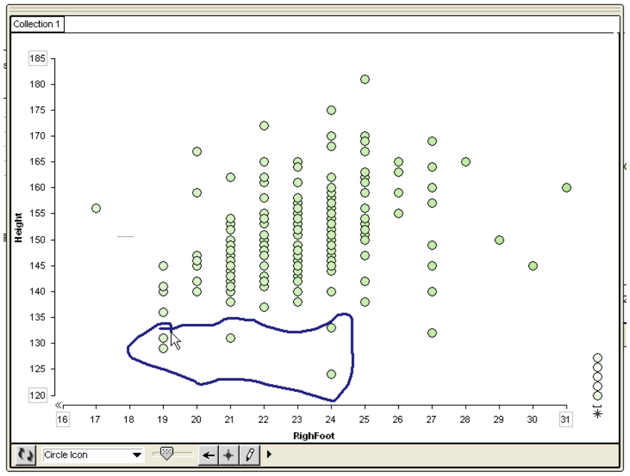


Figure 4.45. The circled region of Natasha's plot that showed where she considered there to be a relationship between height and foot length evident.

Johnty

Student introduction. Johnty was the type of student who worked well one day and not the next. There were occasions when he was fully committed to the learning process and made an effort to engage actively with the tasks and activities completed in the sessions. On other occasions and on the day of the final session, Johnty chose not to engage fully with the tasks. For the most part, Johnty did not work independently and required constant prompting from the teacher to create graphs and interpret them.

Generic knowledge. Johnty understood the way in which *TinkerPlots* represented the data on the case cards and understood how the data on the cards were represented in a graph yet he did not convey his understanding effectively when asked to describe a graph. Inconsistencies such as this typified the way Johnty worked with *TinkerPlots*. At the start of the session, he dragged a plot window onto the screen of the computer. He then stopped. The data in the plot window were randomly arranged and were highlighted for the attribute belly button. This confused Johnty and he could not explain the plot or describe the data represented. It was evident that he had expected a graph to appear on the screen and had forgotten that the randomised distribution of data points in the plot window was the first step to creating a graph. After trying the suggestions offered by the teacher Johnty created a stacked dot plot for the attribute belly button height. When he was asked which attribute was represented in the graph he said, "I think, umm … dunno exactly." The teacher then rephrased the question and asked what information was in the graph. Johnty replied by saying, “Umm, it’s the count [reading from the vertical axis].” Again the teacher attempted to get Johnty to describe what information that was in the graph and he said, “Nothing.” It wasn’t until the teacher asked Johnty to read the title on the horizontal axis that he acknowledged the graph was about the attribute belly button height. The apparent lack of knowledge about the graph and how to read it was contrary to the knowledge about *TinkerPlots* that he had articulated as he created the graph. Johnty read values for the attributes on the case cards and extracted information from a variety of cases. He also clicked on one of the data points in the graph and read the data from the case card. As the session progressed, Johnty demonstrated a greater level of competence with *TinkerPlots* and was more forthcoming with ideas.

Being creative with data. By the end of the session, Johnty was creating graphs with little intervention but was still limited in his application of the features in *TinkerPlots*. He did, however, add the mean to a covariation graph that showed the relationship between the attributes foot length and belly button height and a vertically split stacked dot plot for the attributes gender and height. He also added hat plots to these graphs and read the ranges of the data that were under the crowns of the hat plots.

Like a number of other students, Johnty applied the features of *TinkerPlots* to graphs with the expectation they would be useful. This was not always the case. As an example, he tried repeatedly to change the size of the plot window to make the data points spread out and not overlap. After the teacher suggested that reducing the size of the data icons may be useful, Johnty used that function on all the graphs he created. It was evident he was not satisfied with a graph until he had all the data icons separated and not overlapping. Another example was when Johnty tried to stack the data in a covariation graph.

There were times when Johnty experienced confusion over the graphical representation produced. As well as the example given in the *Generic knowledge* section, there was another occasion when he was trying to create a graph that showed the relationship between gender and height. Initially, he dragged the attribute gender on to the horizontal axis and the attribute foot length to the vertical axis. He then changed the horizontal axis to the attribute height. The data points in the graph remained highlighted for the attribute gender as shown in Figure 4.46. From this point, Johnty tried unsuccessfully to separate the data into two bins horizontally, one for males and one for females. He did not realise that the attribute gender was not on the vertical axis and did not know how to get to the representation he wanted. Johnty knew which graph type he was aiming for but was not experienced enough with *TinkerPlots* to trouble shoot effectively and change the graph to the representation anticipated.

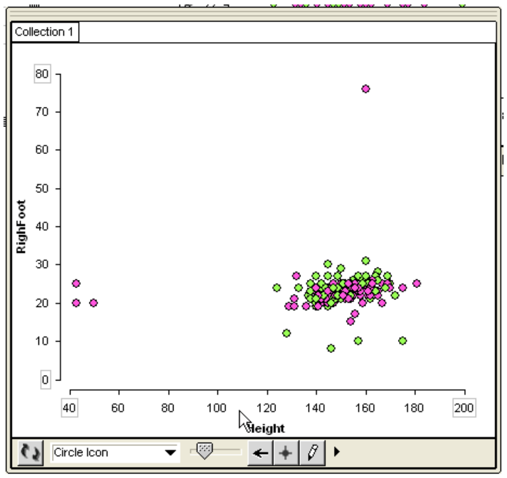


Figure 4.46. Johnty's scatterplot that he created instead of the intended association graph displaying gender and height.

Understanding data. Johnty's understanding of data was characterised by his ability to identify the individual messages in the data. When asked to determine the relationship between foot length and height from a covariation graph, Johnty said, "Dunno." He then went on to explain, "Not really 'cos … that one's tall but it's not the biggest foot." Adding to this he commented on the variation in the data where at least one data point showed there was a relationship between the attributes and another that confirmed his initial conclusion that there was not really a relationship. "'Cos that one, that one adds up. Like, they're pretty short and they've got a pretty small foot. And that one doesn't. It's only 140cm and has a 27cm foot." Johnty's explanation showed he used the variation in the data to draw his conclusion. It also showed he recognised that some of the data points fit with his expectations for a relationship to exist. Johnty used his knowledge of the context to establish how a relationship would be exemplified and then used individual data points on the graph to determine that his expectations for a relationship to exist were not universally realised.

Johnty also used individual characteristics of hat plots and the mean to make judgements about graphs. For a stacked dot plot with the attribute height on the horizontal axis and the vertical axis split into bins for the attribute gender, Johnty stated that the "Girls' height's a bit taller." When he was asked to explain further he said, "Ah, I'll put a hat plot on. [Short pause while he did this.] Pretty much, they're the same." But then he went on to say, after adding the mean to the graph, "Girls' average is a little bit bigger."

Thinking about data. The use of individual characteristics of graphs by Johnty to draw out the messages in the data hinted that he was moving towards being able to interpret data more comprehensively. He did, for example, make the shift from only using individual data points to interpreting data globally when he described the relationship between the attributes height and belly button height in a covariation graph. When asked to describe how the graph showed if there was a relationship he replied, "Mmm, it goes up like diagonally…. The taller you are, usually, the higher your belly button is." Johnty's explanation demonstrates that he used the graph effectively to draw his conclusion.

Johnty's interpretation of graphs was partly based on his knowledge of the context and partly on his direct interpretation of the graphs. The way he phrased "the taller you are, usually…" in the previous quote, implies that he used his knowledge of the context to support his thinking about the graph. He also used the context of the data to ascertain if a data point was an outlier. He did this by comparing the belly button height measures in the data set to a benchmark that he had established from his knowledge of the context. As he removed the cases with belly button height of 65cm and 66cm he declared, "They're usually a metre something." There were other occasions when Johnty's decisions about outliers were based on the position of data points on the graph. He removed any outliers that were isolated from the main cluster of data. He then used his knowledge of the context to make decisions about data points that were closer to the main cluster of data on a graph. Johnty was more particular about removing all data points that he considered to be outliers than the other students. He had established very firm ideas about what were acceptable measures for foot length, height, and belly button height for the age group of students in the data set and used that information to make his choices.

William

Introduction. William enjoyed working with computers and showed enthusiasm for the work completed in the sessions. On the day of the last session he was very tired. He yawned repeatedly as he worked and needed many prompts to focus on the tasks at hand. William indicated that he had stayed up late the night before playing computer games online. The enthusiasm for working with computers and his engagement with *TinkerPlots* was undermined by his fatigue after gaming online for hours.

Generic knowledge. William was a competent user of *TinkerPlots*. He did not require assistance to access the main features of *TinkerPlots* or to construct graphs. When asked to construct a graph he did so without hesitation. He worked fluidly within the *TinkerPlots* interface to drag attributes into a plot window, change the bin width on the axes, remove outliers, and insert the mean and then numerical value of the mean. William did not add hat plots to the graphs. As he created the graphs he explained clearly that the data were organised on case cards and that the data for individual cases were recorded on separate cards. William made the connection between the data on the cards and the data values in a graph by highlighting a data point and reading the values of the attribute highlighted from the relevant case card.

The language William used throughout the session when describing graphs and the characteristics of graphs was mostly everyday language. There was evidence, however, that he was moving towards making more formal statistical language part of his vocabulary. He referred to the mean as the average, identified the mode correctly, and described outliers as "Out, outgoers or something." William had also internalised some terms that were distinctive to *TinkerPlots*, such as cases, case cards, bins, and attributes. He used these terms appropriately as he worked without prompting.

Being creative with data. William was the only student who did not construct any graphs with a continuous scale for the attributes with continuous measurement data. All the graphs he constructed had the scale of the axes sorted in bins, even the covariation graphs. Once he had a graph constructed to his satisfaction he did not attempt to change the representation if it did not provide the information he was seeking. He preferred instead to construct a new graph and apply potential changes to it. An example of the graphs William constructed is in Figure 4.47.

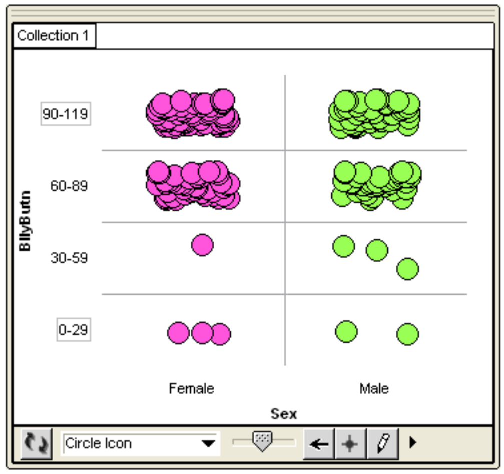


Figure 4.47. William's plot with belly button height on the vertical axis split into bins.

William was inquisitive about the data and wanted to find the mean and the mode of the data. He applied the mean to his graphs from the measures of centre icons in the *TinkerPlots* menu but was confused by the way in which *TinkerPlots* represented the mean when the data were organised in bins. William did not understand that the blue line across the width of the bin indicated the mean of the data was located within that particular bin. It became evident he expected a single value for the mean when he added the numerical value of the mean to the graph and commented "That's better." After the mean was added to the graphs, William did not use it to interpret the data. Although expressing the desire to determine the mode of a graph on a two occasions, once the mean was established, William did not analyse the graphs to ascertain the mode. One of the graphs William constructed with the mean added is in Figure 4.48.

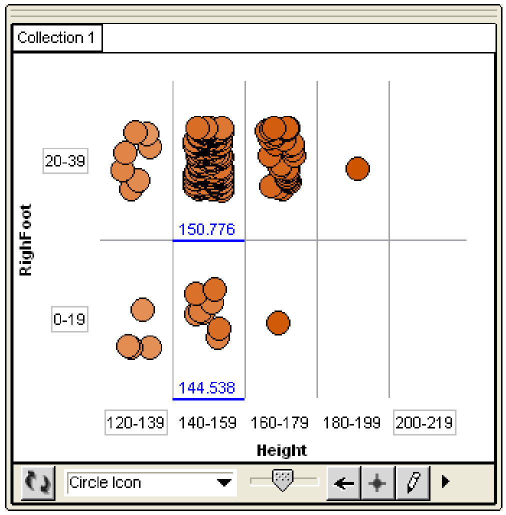


Figure 4.48. William's plot displaying foot length and height split into bins with the mean added.

Understanding data. For the most part, William worked with purpose to construct the graphs he wanted and understood the graphs he created. Comments like, "Most people's feet are 20 cm to 39 cm," showed that he realised the bins in a graph covered a range of data and when he said, "There's most of the tallest people … which means that they've also got the, the highest foot," showed that he understood the relational aspect of a covariation graph. Although he did not know how to interpret the way in which the mean was represented on a graph with bins, William identified correctly that the mean was related to the attribute on the horizontal axis as was the case in Figure 4.48.

Thinking about data. Although the context of the data was familiar to William and he was able to connect with the nature of the data recorded, he did not use his knowledge of the context to interpret the graphs. His judgements were based on the position and the density of data in a graph. When asked to describe the relationship between the attributes height and belly button height from the graph shown in Figure 4.49, William replied, "I think that the taller you are the bigger your belly button." He then indicated by pointing that he was talking about the region in the top right hand bin of the graph.

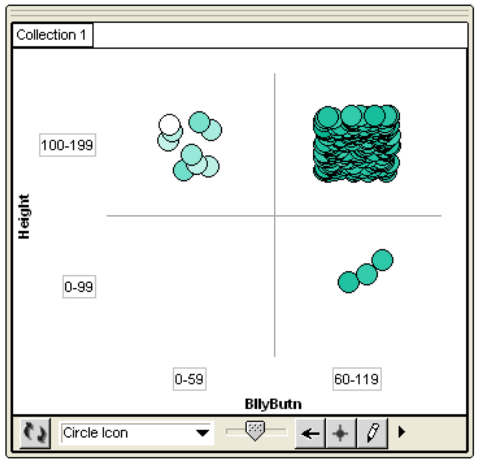


Figure 4.49. William's plot displaying height and belly button height.

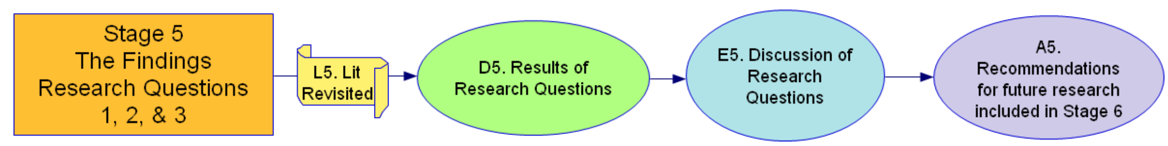
William generally used the position of data in the graphs to determine if the data were outliers. On several occasions he removed outliers because they were separated from the main clusters of data and did not make connections to the context of the data. He did not justify his decisions by referring to his knowledge of measures of foot length, height, or belly button height gained from the previous sessions or his personal experiences. There was, however, an occasion when he used the data available to make his decision saying, "It's not real, that one, [pointing] it ain't, because it's one of the tallest but it's umm, got one of the smallest feet."

Concluding Remarks

The twelve Student Profiles developed from the Student Interviews in this stage of the inquiry provide valuable information about the statistical thinking and reasoning of the students as well as their graph creation and interpretation skills using *TinkerPlots*. The Student Profiles provide a comprehensive view of each student's knowledge and understanding of graphs and graphing across all the dimensions of the *Model of Learning Behaviour in EDA Graphing Environments* (Figure 1.7). No student demonstrated understanding for all of the key behaviours in the dimensions but all students could engage with some of the key behaviours for each of the dimensions. The key behaviours of "Recognising appropriate use of different forms of graphs" (*Understanding data* dimension) and "Asking questions about the data" (*Thinking about data* dimension) were notably absent from the Student Profiles. The activities undertaken by the students in the Student Interviews did not provide the opportunity to evidence these behaviours. Activities that included extended statistical investigations that required students to pose questions and make decisions about how to present different data would be suited to evidencing these behaviours. Considering the analysis in this stage of the inquiry is the start of answering the research questions, which is done in Stage 6 of the thesis. The Student Profiles are utilised multiple times in Stage 5 to answer the research questions as further distilling of the results in the Student Profiles takes place.

Stage 5

The Findings



The purpose of Stage 5 of the inquiry is to present the results of the three research questions:

4. How can the learning behaviours of students as they engage with exploratory data analysis software be characterised through a framework that can then be used to explore and analyse students' understanding of covariation using *TinkerPlots*?
5.
6. How do students interact with the exploratory data analysis software, *TinkerPlots*, to represent data in a variety of forms when exploring questions about relationships within a data set?
7.
8. How do students develop an understanding of covariation in the exploratory data analysis software environment afforded by *TinkerPlots* and use these understandings to provide informal justification for their conclusions about the relationships identified?

In this stage of the thesis the framework, *Model of Learning Behaviour in EDA Graphing Environments* (Figure 1.7), is revisited in response to Research Question 1 to demonstrate the usefulness of the framework for characterising students' learning behaviours when using *TinkerPlots* to conduct data analysis. In order to answer Research Questions 2 and 3, further data analysis is carried out. Returning to the Student Interview data coded in Stage 4 to develop the Student Profiles extends the data analysis process and contributes to the soundness of the inquiry design adopted for this inquiry (Seeto & Herrington, 2006). Iterative data analysis allows data to be interrogated in multiple ways to provide insights about the different constructs and relationships relevant to the questions explored (Lincoln & Guba, 1985). The data analysis process is applied to the Student Interview data to answer Research Question 2 and Research Question 3.

Research Question 1 – Framework to Characterise Student Learning

The framework, *Model of Learning Behaviour in EDA Graphing Environments* (Figure 1.7), developed for this inquiry provided a standpoint from which to examine students' interaction with *TinkerPlots* as they engaged in data analysis activities. This model enabled an exploration of students' development of understanding of covariation from a perspective not previously possible.

Developed from a synthesis of the literature, the four dimensions of the *Model of Learning Behaviour*, *Generic knowledge, Being creative with data, Understanding data,* and *Thinking about data*, recognise the complex nature of creating and interpreting graphical representations of data. The constituent parts of each dimension were drawn from other models of graphing (Friel et al., 2001; Moritz, 2004; Pfannkuch & Wild, 2004; Shaughnessy, 2007) and take into consideration learners' behaviours and knowledge associated with using educational technology to create graphs (Alessi & Trollip, 2001; Kosslyn, 1989).

The framework also guided the development of the Sequence of Learning Experiences implemented in the inquiry and contributed to the development of the *Criteria for Evaluating Exemplary EDA Software* (Table 2.2). Aspects of each the four dimensions were evidenced in all of the twelve Student Profiles reported in Stage 4. Applying the framework to analyse the Student Interviews to construct the Student Profiles demonstrates its usefulness and capacity to document a comprehensive view of student behaviour and learning. The usefulness of the framework for analysing students' developing understanding of covariation is demonstrated in the results for Research Question 3, later in this stage. This section revisits the framework according to each of the dimensions, and discusses the interconnections between the dimensions.

Generic Knowledge

In order for students to demonstrate their understanding of EDA strategies and the statistical ideas they evoke, they need to be able to create graphs to provide evidence of their thinking. To do this, students need to know how to use the educational technologies available at the time and understand the graphical representations that the technologies generate. In addition, they need to be able to express themselves verbally, as student learning cannot be deduced comprehensively from only looking at the graphical representations created. These three notions are intrinsically bound within the learning environment afforded by the educational technology used, in this case *TinkerPlots*.

The dimension of *Generic knowledge* included: Speaking the language of data and graphs, Understanding how to use the features of software and technology environments, and Recognising the characteristics of data and graphs. This dimension was useful in that it identified that the students, at times, lacked the language to articulate their thinking and showed that students were making the connections between their internalised everyday language and conventional statistical language (Leinhardt, Zaslavsky, & Stein, 1990). As well as using everyday language when talking about statistical ideas and concepts, some of the students used the statistical terms appropriately. This was a reflection of the developing nature of their understanding and limited previous experiences in working creatively with data. Close interaction with the teacher, however, provided the opportunity for questions to be asked in order to clarify the students' responses and prompt further discussion to reveal and provoke their understanding.

Although some of the students required reminding about how to use some of the more complex features of *TinkerPlots*, all of the students in this inquiry were proficient at using the basic features to create stacked dot plots and scatterplots. More variation was evident in their understanding of the constituent parts of the graphical representations. Kimberley, for example, needed reminding that a continuous attribute had to be selected before a hat plot could be added to a graph. Jake also had difficulty understanding that categorical and continuous attributes influenced the type of graphical representation possible. It was evident that understanding the difference between the types of attributes was lacking for most of the students. Making the distinction between the attributes was not a focus of the sequence of learning experiences but the insights gleaned from the Student Profiles indicate it would have been beneficial to focus on the characteristics of data and this should be considered when designing learning activities for future inquiries.

Being Creative with Data

The term "transnumeration" encapsulates the behaviours included in the *Being creative with data* dimension: Describing data from graphs, Constructing different forms of graphs, Summarising data, and Reducing data to graphical representations. Transnumeration means "changing representations to engender understanding" (Pfannkuch & Wild, 2004, p. 18) and includes transforming data into multiple graphical representations, summarising data, and changing data representations to communicate meaning. The learning environment afforded by *TinkerPlots* provided the opportunity for the students to work with data in these ways.

Analysis of the Student Interviews according to the *Being creative with data* dimension revealed that although the students put graphs together via different pathways, they often applied as many different features of *TinkerPlots* onto the graphs as possible when trying to make sense of the data. In many cases, the added features were redundant and did not assist students to make decisions about the data. All of the students constructed a variety of graphs and most of them summarised the data using hat plots. Many of them used the mode and the reference lines to determine specific data values and added the mean and the median. Often, the students did not use the mean or the median to interpret the data and some of the students demonstrated a lack of understanding of those statistical concepts. Having the freedom to choose which features to apply and then making the decision about the relevance of particular data summaries or data representations supported the students' transition from constructing graphs to describing what a graph showed when questioned by the teacher.

Understanding Data

The graphical representations used and how the data are represented assist in identifying trends and making predictions from the graphs (Jones et al., 2004). This implies that having knowledge of different graphical representations and forms of data is important. Also important is knowledge of the data when the same data are represented in a variety of forms. The *Understanding data* dimension focuses on the behaviours associated with recognising the way data are displayed in different forms but also the connections among different forms of the same data.

The *Understanding data* dimension includes the behaviours: Answering questions about the data; Identifying the messages from the data; Understanding the relationship among tables, graphs, and data; Making sense of data and graphs; and Recognising appropriate use of different forms of graphs. Scrutinising the students' data analysis strategies with this dimension showed that the students were able to make the connections between the information about the data recorded on the data cards and the information displayed in graphical representations created from the data. Having the data cards and the graphical representations displayed simultaneously in *TinkerPlots* facilitated the students making the connection between the two forms of representation. This proved useful when students clicked on a data point in a plot, which brought the relevant data card to the front of the stack, and then gathered other information about that data point from the data card. This supported the students to elicit the messages from the data.

This dimension was useful for identifying the way in which the students used the graphical representations of make sense of the data. It showed that the students focused on either the local or global characteristics of the data as described by Ben-Zvi and Arcavi (2001). Natalie, for example, only referred to the value of one data point for one attribute when answering questions about the relationship between two attributes, whereas Kimberley, James, and William were able to describe the relationship by identifying the trend in the data and projecting what the trend indicated for both the attributes involved. In contrast, Blaire and Mitchell gave more attention to the location of clusters of data and outliers when making sense of the data and graphs.

Using the *Understanding data* dimension also brought to attention the students' incomplete understanding of the appropriate use of different forms of graphs. Although able to use a scatterplot appropriately for identifying a trend when exploring covariation in the data, many of the students also used split stacked plots for the same purpose. This suggested that some of the students had not made the distinction between the purposes of each graph type and had not made the connections between the different forms of graphs and the statistical concepts they embodied.

Thinking about Data

Graph comprehension is described as “deriving meaning from graphs,” whereas graph interpretation is described as “rearranging material and sorting information from the less important factors” (Friel et al., 2001, p. 129). Both of these features of using graphs are encompassed in the behaviours included in the *Thinking about data* dimension: Asking questions about the data, Recognising the limitations of the data, Interpreting data, Making causal inferences based on the data, and Looking for possible causes of variation. Like the other dimensions, this dimension is complex and involves students working with data in many different ways.

Using the *Thinking about data* dimension highlighted the important role that the context of the data played when the students interpreted the data. The students’ descriptions and interpretations of the graphs they constructed maintained strong connections to the context of the data. Predominantly, the decisions about outliers were based on the context of the data but some students also made decisions about outliers based on their position in the graph. Interestingly, Johnty used both strategies. He removed any outliers that were isolated from the main cluster of data without considering the context or the value of the attributes concerned. Following this, he used his knowledge of the context to make decisions about data points that were closer to the main cluster of data.

Although the students’ knowledge of the context assisted some of the students to articulate their reasoning about covariation and supported their thinking about the relationship displayed in a scatterplot, it was also problematic for other students. When not able to express how a relationship between two attributes was evidenced in a graph, some students used their established knowledge of the context to assert a relationship existed. Possible causes of variation offered and inferences made by the students were also connected to their knowledge of the context, whereas the students’ understanding of the limitations of the data was gleaned directly from the graphical representations and expressed without reference to the context of the data. The limitations were expressed in terms of the spread of the data, the position of outliers, and the position of clusters of data.

This inquiry has shown that students make connections to the context of data, are influenced by the context of the data, and use the context of the data in meaningful ways (Fitzallen & Watson, 2011; Langrall et al., 2011; Pfannkuch, 2011). The knowledge of the context of the data associated with the students' previous experiences coupled with the interpretation of the graphical representations "are essential for students to become critical statistical thinkers" (Fitzallen & Watson, 2011, p. 259). It is necessary, however, for students to distance themselves from the context of the data to establish a connection to the statistical aspects of the data. The students in this inquiry did not do this but their propensity to rely on the context of the data may be an indication of their incomplete knowledge of statistical concepts and lack of experience at employing graphs to make causal inferences and ask questions about the data.

Knowledge of the context and the influence that knowledge has on students' thinking and reasoning about graphs is not acknowledged explicitly in the *Model of Learning Behaviour in EDA Graphing Environments* (Figure 1.7). Further framework development could include this aspect of data analysis in the framework and may be an interesting research focus. The prevalence of context in the students' statements indicates that it should be included but investigation into which dimension it should be situated is warranted.

Interconnections Among the Dimensions

The *Model of Learning Behaviour in EDA Graphing Environments* was not offered as a developmental model. Neither was it intended that the dimensions of the *Model of Learning Behaviour* be considered completely distinct and disconnected aspects of data analysis. The results from this inquiry suggest that the behaviours in the framework depended on the complexity of the students' thinking and the complexity of the questions they were answering. At times, when determining the foot length of the tallest person say, the students only applied behaviours in the *Generic knowledge* dimension. When justifying a statement about a relationship identified in the data, some students employed behaviours from one dimension and then shifted fluidly to using behaviours from other dimensions to build their argument. When this occurred the students applied the behaviours from multiple dimensions to express and support their thinking. They were not restricted to using only behaviours from one dimension at a time. This showed that the students had made interconnections across dimensions. Although the interconnections are complex and important, by looking closely at students' data analysis strategies for each of the dimensions, as has been done in this inquiry, much can be learned by utilising the *Model of Learning Behaviour in EDA Graphing Environments* (Figure 1.7) when exploring the nature of student learning about statistical concepts with educational technologies, such as *TinkerPlots*.

Research Question 2 – Student Interaction with *TinkerPlots*

This inquiry identified that students employed three strategies when creating graphs and interpreting data. These strategies were derived largely from data coded for the dimensions *Generic knowledge* and *Being creative with data.* The identified strategies were: Snatch and Grab, Proceed and Falter, and Explore and Complete.

Thematic analysis, as described by Miles and Huberman (1994), was used to ascertain the strategies that exemplified the ways in which the students engaged with *TinkerPlots*, as they completed statistical activities. After the data from the audio data and the on-screen capture video data generated by *Captivate* (Adobe, 2007) were coded according to the four dimensions of the *Model of Learning Behaviour in EDA Graphing Environments* (Figure 1.7) to construct the Student Profiles. Another round of data analysis was conducted to answer Research Question 2: How do students interact with the exploratory data analysis software *TinkerPlots*. The data coded for the behaviours within the four dimensions of the *Model of Learning Behaviour* were sorted into two sub categories: one associated with how students interact with *TinkerPlots* to create data representations; the other was associated with how students interact with *TinkerPlots* to answer questions about the data. The identified strategies are discussed in turn.

Snatch and Grab

The Snatch and Grab strategy was used by students who had incomplete knowledge about the potential of *TinkerPlots* to support data analysis activities. They were not always self-directed in their selection of graph types and required prompts and suggestions from the teacher to extend their exploration of the data beyond the basic graph construction phase. These students tried to take advantage of *TinkerPlots* by adding multiple features to the graphs but made little sense of how the features contributed to a better understanding of what the graphs showed. The students often accessed the features in a random manner. Their aim was to access the features they were familiar with in the hope the features helped answer the questions. On occasions they were successful but more often the students could not see how the additional features contributed to the ability to answer the questions about the data.

Although Natalie's responses and descriptions were not very sophisticated, it was clear she had established a basic understanding of the utility and purpose of graphs. She was open to exploring the application of *TinkerPlots* and had a positive attitude about the software and its potential. As she worked she adopted a "let's see what happens" attitude and accessed multiple features of *TinkerPlots* to find a representation that assisted her to answer the questions about the data. It was evident that she drew on her previous experiences to interpret graphs but had difficulty as her knowledge and understanding was limited. Natalie knew many pieces of different information about graphs and graphing but she had not made significant connections between these pieces of information. This constrained her thinking, restricted her examination of graphs, and stymied her explanations as she grappled to put all her ideas together. Like other students, Natalie's interpretation of graphs focused on specific values and did not extend to consideration of relational characteristics.

Although not always self-directed in the construction of graphs, Natasha showed that she had an established knowledge of a variety of graph types and how to use them to extract specific information. Her graph interpretation skills, however, were under developed. Like a number of other students, Natasha had gained knowledge about graphs and data but had not made meaningful connections between what she knew and how to use it effectively. Sometimes her interpretations were incomplete but they were sufficient to establish that she had an emerging understanding of how graphs could be used to interpret data and tell a story. Natasha accessed many of the features of *TinkerPlots* as she created a variety of graphs but she often sought reassurance from the teacher that her choice was going to be productive. It became evident that Natasha based her selection of features on her knowledge of what was available rather than being a purposeful choice because she understood how the feature was going to facilitate the data interpretations.

Kimberley wanted the physical appearance of the graphs to fit a preconceived idea. She liked the data points to be separated and did not feel comfortable when they were all bunched together. When the graphs had clusters of data that were bunched, she had trouble "reading" the graphs. This prevented her from looking at aggregates of data to tell a story and caused her to focus on individual data points on the fringe of clusters. Although Kimberley was hampered by what she perceived as lack of clarity, she was not hampered by expecting a certain type of graph and was able to work flexibly within the constraints she imposed. Within the *TinkerPlots* environment she created different types of graphs and did not follow any particular routine when constructing the graphs. Kimberley used the features of *TinkerPlots* in much the same way with all the graphs she constructed. The tinkering she did by changing the size of icons, plot windows, and scales of axes, as well as inserting hat plots, the count, the mean, and reference lines, was her attempt to get the graphs to a point where they answered the questions she was exploring. Kimberley took advantage of the tools she had at her disposal within the *TinkerPlots* environment to help her think through what the graphs offered. She identified readily whether they were useful or not but tried them all. There were times, when this way of working was not very productive but at other times proved to be helpful.

Johnty required reminders about how to create graphs and access the features of TinkerPlots. As the session progressed he became more competent but was inconsistent in the way he worked. He was also inconsistent in the way he interpreted graphs. Sometimes his responses to questions only involved stating the specific value of data points. Other times, he offered responses that showed the complexity of his thinking when interpreting graphs. Johnty showed that he understood the characteristics of graphs and used the structure of graphs to convey his understanding. He also showed he had the potential to go beyond the use of individual characteristics to describe the data to interpreting the data from a global perspective. Like Kimberley, Johnty fiddled constantly with the graphs by changing the appearance of the size of the plot window and the data icons, as well as accessing all the available features of *TinkerPlots* without purpose and often inappropriately. This was evident when he tried to add hat plots to a graph without a continuous scale and tried to stack the data in a scatterplot.

Proceed and Falter

The students who used the Proceed and Falter strategy used *TinkerPlots* as a construction tool. They were purposeful in the way they went about creating graphs and approached the tasks given with confidence. These students understood which graph type would be useful to answer the question asked but often did not follow through with responses that indicated they understood why the representations were useful. At this stage, the students benefited from suggestions from the teacher to move on and explore the data further. Often these students relied on their prior learning of graphing and statistical concepts to answer questions. This constrained the students' thinking and prohibited them from being open to the story the data representations had to tell.

Jake was not very adventurous and did not experiment with the options *TinkerPlots* made available. His choices of graph types and features to apply were based on what he wanted to see happen. When a representation did not appear as expected, he became hesitant and did not know how to proceed. Although Jake could construct graphs easily, his lack of confidence restricted him from using *TinkerPlots* to its full potential. As a consequence, Jake treated the graphs he constructed like graphs printed on paper – static and fixed. Because his interpretation of graphs was influenced by his established expectations, he was unable to instigate changes as he could not envisage how the graphs could be different.

Blaire was also constrained by her established ideas but it was not the look of the graphs that restricted her thinking. Blaire's focus was on using the mean to make sense of the graphs. When considering the graphs created by Blaire, it became apparent that she used the mean to compare attributes and to describe the relationship between two attributes. Although she used the term "average," her application of the mean was inappropriate. When determining the average from a covariation graph, the language she used described the mode not the mean average. On another occasion she described the mean as being in the middle of a cluster of data. In some instances, she used the numerical value of the mean generated by *TinkerPlots* to compare groups. In these cases she used the mean as a single value and did not demonstrate that she understood the mean was representative of the whole group of data. When asked to explain what the "average" meant, Blaire read the data value from the graph and was unable to expand further.

Rory was another student who was confident enough to create a variety of graphs without assistance but often did not go further when asked to discuss and describe what the graphs revealed about the data. Rory often lacked initiative, which may be attributed to his lack of understanding of data, in general, and how they were represented in *TinkerPlots*, more specifically. He knew how to use *TinkerPlots* but chose not to use it in the way other students had done to move past an initial plot and explore further possibilities in the data. Although he had used *TinkerPlots* on many occasions, he was still discovering what it had to offer. Rory was not very adventurous and did not use the features of *TinkerPlots* to advantage. His understanding of how to use *TinkerPlots* to facilitate the interpretation of data was limited.

Jessica was knowledgeable about *TinkerPlots* and worked confidently and instinctively with the software to create graphs. Jessica knew what graph type would fulfil what she was asked to do but it was not evident that she knew how to use the graphs effectively to answer questions. This was also the case when she used the hat plot and the mean. Jessica knew these representations were useful but could not articulate why. As a fallback position, Jessica focused on individual data points in graphs and her intuitive knowledge of the context to make informal inferences about the data. When not focussing on individual values, she focused on the centre of the data. The features of *TinkerPlots*, such as the mean and the drawing tool, helped Jessica by giving her a point of focus to stimulate her thinking.

Although Shaun experienced lapses of memory about how to use *TinkerPlots* and the language needed to explain the graphs and their characteristics, he was still able to connect with the representations constructed with the software. He did not, however, always understand what some of the features contributed to the interpretation of the data and his explanations offered few insights into his thinking about the data and the way in which he used the features of *TinkerPlots* to make his decisions. Shaun recognised the changes to the graphs when deleting outliers, speculated about the variation in the data when determining a relationship between attributes, and anticipated where the crown of a hat plot would sit on a graph. This showed that he actively used the representations offered in *TinkerPlots* to interpret the data but there were occasions when his interpretations were flawed and hence revealed an incomplete understanding of some representations.

Explore and Complete

The students who used the Explore and Complete strategy used *TinkerPlots* effectively to support the interpretation of graphs. This approach, as the name implies, was characterised by a sense of exploration thus allowing the data to take centre stage and *TinkerPlots* to play a supporting role. These students did not pre-empt what the graphs would look like or predict the story the graphs had to tell before they were constructed. As a result, the responses to questions and the descriptions of graphs conveyed that the intuitions developed about graphing and data analysis were shaped by the data representations. Explanations about the data were unhampered by fixed ideas and preconceptions and often decisions were made quickly about what a graph showed and those decisions were in response to the representations *TinkerPlots* offered.

William was unusual in that he was the only student who chose not use a continuous scale to determine if there was a relationship between two numerical attributes. When he created the graphs he dragged the axes enough to extend the number of bins but not enough to create a continuous scale. He then added the mean but was unsure how to interpret the resulting representation. William was not deterred by this as he was able to interpret the graphs appropriately by using his understanding of the spread of the data and the relational nature of covariation graphs to make conclusions about the data. Although William displayed hesitation when confronted with representations he did not fully understand, it did not constrain his thinking and he was able to explore the data further using other representations. James worked confidently and was able to create a variety of graphs without assistance. Once he created the graphs he did not refer back to the case cards and focused his attention on the graphs to determine values of data points. James did not need the reassurance of reading values from the case cards, the mean, and reference lines as Kimberley and Blaire had done. Although he did apply these features to the graphs, he was much more confident and did not need the values generated by the features to confirm his thinking.

James' interpretation of graphs focused on the distribution of the data. His decisions were influenced by the physical structure of the graphs and the position of the data points within the plot window. James was very purposeful in the way he accessed the features of *TinkerPlots* and only made changes to graphs, such as adding a hat plot, when he wanted to know what a particular feature had to say about the data. His actions were exploratory in nature but were deliberate and not based on a trial-and-error strategy adopted by many of the other students.

Mitchell explored the story the data had to tell by moving back and forth between looking at clusters of data and individual values. He was able to integrate these two characteristics of graphs into his decision making process. This enabled him to make decisions about graphs from a general perspective and then use specific values to justify his decisions and identify variation. His interpretations of the data were based on the position of data points in a graph rather than their specific values. The connections Mitchell made to the context of the data were directly related to his interpretation of the graphs and were not influenced by his knowledge of and previous experience with the context. Unlike James, Mitchell did adopt a trial-and-error strategy when accessing the different features of *TinkerPlots* but knew immediately if the changes instigated were useful.

Discussion of Student Interaction with TinkerPlots

The three strategies adopted by the students in this study – Snatch and grab, Proceed and Falter, and Explore and Complete – show the different ways in which the students worked within the *TinkerPlots* environment and reveal their propensity to exploit the affordance of the software. A summary of the students according to each of the strategies adopted is in Table 5.1. The students who employed the Explore and Complete strategy had a comprehensive way of working with data. They successfully integrated their knowledge of graph creation and graph interpretation. Most of the time, their thinking was moving back and forth between these modes as they made sense of the data. The students who employed the Proceed and Falter strategy put their initial efforts into creating graphs and then stopped and interpreted the data as a separate activity. At times, teacher intervention was necessary for some of the students to transition successfully from the creation mode to the interpretation mode. The students who employed the Snatch and Grab strategy, often worked aimlessly when interpreting the data. They did this by repeatedly accessing multiple features of *TinkerPlots* without considering if the particular features would be useful or not. Like the other students, the students who used the Snatch and Grab strategy were competent at using *TinkerPlots* to create graphs and relied on the features of *TinkerPlots* to provide the inspiration needed to think about the data. As these students did not take time to think about the changes they made, many of the features accessed were redundant and not helpful. Also, little advantage was taken of the affordances of the features that may have been useful. Although, the intention of the analysis was not meant to be hierarchical the three strategies exhibited characteristics that could be considered developmental. Interrogating the way in which students interact with *TinkerPlots* may provide a precursor for determining their propensity to work holistically in that environment.

Table 5.1.
Students and Interaction Strategy Adopted When Using TinkerPlots

Snatch and Grab	Proceed and Falter	Explore and Complete
Natalie, Natasha, Kimberley, Johnty	Shaun, Blaire, Rory, Jessica, Jake	James, William, Mitchell

Effective and productive learning will not occur by just giving students access to statistical software. The teacher must play an active role in establishing and supporting the development of student thinking and reasoning (Behrens, 1997; Ben-Zvi & Friedlander, 1997; Cobb, McClain et al. 2003). As students' statistical understanding and knowledge of graphs becomes more sophisticated "the teacher's role changes from active instructor to fellow investigator" (Ben Zvi & Friedlander, 1997, p. 54). In the interim, understanding how students use interactive software would inform teachers about when to and in what ways to intervene in the learning process. Students who use the Explore and Complete strategy make connections between the graphical representations and the meaning they embody. When this occurs, the teacher can focus attention on what the data mean and inferences that can be made from the data. When students use the Proceed and Falter strategy, teacher intervention may need to focus attention on how the graphical representations answer the question being explored before attention is given to thinking further about the data. When students use the Snatch and Grab strategy, teacher intervention may need to focus on both the graphical representation and the question being explored to ensure the students develop an understanding of the alignment necessary between the graphical representations constructed and the question being explored.

Although the three strategies described are indicative of the way in which the students in this inquiry engaged with *TinkerPlots*, it appears more evidence is needed to determine if the three strategies are hierarchical or are characteristic of the different ways students engage with graphing software applications. Future research is needed to follow students to determine if the way they engage with graphing software is progressive as they develop an understanding of statistical concepts and become familiar with the features of the graphing software.

Research Question 3 – Development of Understanding of Covariation

The Student Profiles were analysed a second time to answer Research Question 3: How do students develop an understanding of covariation in the exploratory data analysis software environment afforded by *TinkerPlots* and use these understandings to provide informal justification for their conclusions about the relationships identified? Kosslyn (1989) contends that to analyse a graph it is necessary to understand the interrelated connections among the constituents of a graph. He asserts that understanding the interrelated connections fosters the interpretation of graphs on three levels. First, the individual elements and their organisation can be described. Second, understanding of the display can be determined by looking at the relations among the elements of the graph. Lastly, the analysis can extend to the interpretation of the symbols and lines that goes beyond the literal reading of the information. Influenced by the work of Kosslyn, Curcio (1989) considered school students' interpretation of graphs from three perspectives: reading data directly, reading "between" the data, and reading "beyond" the data. These phrases reflect to some extent the increasing demands of the levels suggested by Kosslyn (1989). In 2007, Shaughnessy added to the work of Curcio by suggesting the further need to read "behind" the data. The additional dimension suggested by Shaughnessy (2007) together with the reading "beyond" the data perspective offered by Curcio, reflects the complexity of graph interpretation at the third level suggested by Kosslyn.

The three levels of graph interpretation highlighted by Kosslyn (1989) and others were used to devise a framework for analysing the students' responses about covariation and the relationship seen in the data from the graphs they created. The framework devised was based on the SOLO model developed by Biggs and Collis (1982). It has three levels that describe the increasing complexity of thinking required when interpreting graphs. At the first level, termed uni-structural, responses employ single elements to describe the relationship seen in the data. These are isolated statements that are not linked together to draw conclusions. This aligns with the first level of the Kosslyn framework. At the second level, termed multi-structural, responses employ multiple elements to describe the relationship seen in the data. These are offered in the same excerpt of conversation as a string of ideas that are supporting evidence about the relationship seen but are not presented in an interrelated fashion. This aligns with the second level of the Kosslyn framework. At the third level, relational, responses employ multiple elements to describe and justify the relationships seen in the data and make explicit how the elements are interrelated, including elements beyond the literal reading of the data, such as the context. This aligns with the third level of the Kossyln framework.

To answer Research Question 3 the data from the Student Profiles coded for the behaviours within the four dimensions of the *Model of Learning Behaviour in EDA Graphing Environments* were coded again. The students' statements, descriptions, and justifications about the relationship seen in a graph were analysed to determine the complexity of the responses according the SOLO model developed. The data to answer Research Question 3 were dominated by but not restricted to the data coded for the dimensions *Understanding data* and *Thinking about data*.

Student Reasoning about Covariation

Jake's responses to questions about the relationship between two attributes were uni-structural. It was evident he appreciated that graphs could be used to establish if there was a relationship between two attributes and understood that the graphs had to display the data for two attributes to determine the nature of a relationship. Upon request, he constructed appropriate graphs to explore the possibility of a relationship existing but his interpretation of them was extremely limited. When asked to describe a covariation graph displaying foot length and height (Figure 4.35), Jake said, "Hmm … the … the foot's not really getting any bigger. It's staying kind of the same." To justify his conclusion he cited the values of the various data points and determined, "Umm … The difference between the tallest, the highest one and the mode is … 17, I think." Although Jake extracted valuable information from the graphs as isolated statements, he did not make any connections among the information, his generalised conclusions, and his understanding of the context. Most of the time, his judgements were based on the comparison of individual values and ranges of data in a graph.

Natasha based her judgements about the relationship between two attributes on her knowledge of the context. For each of the covariation graphs she created, she made generalised statements like, "Yeah. 'Cos the taller you are, like … there's usually got like a higher … higher belly button height, if that makes sense?" She was also able to make conjectures about the relationship but was limited to considering one data point at a time and did not make any collective statements to suggest her thinking was shifting beyond being uni-structural. As an example she said, "Umm … because like if their foot was, probably say about 10 centimetres, they'd probably have to have the foot of a 7 year old or something," when referring to the graph in Figure 4.45. Most of the time her statements reflected an understanding of the context from a personal perspective and she did not use the graphs directly to make her judgements.

Rory made many statements that indicated he had a good understanding that *TinkerPlots* could be used to look at the relationship between two attributes. During the discussion about the purpose of the session he volunteered, "Before we thought, when we thought last time we seen … if they've got a big foot that they're pretty tall. I want to try that." After constructing the graph he reported, "Yeah that the height and, well the big people have the big foot, sort of." When describing this and the other graphs he constructed, the evidence he offered to support his statements focused on either the value of specific data points or regions of the graphs, where there were clusters of data. This was evident when he described separate components of the graph in Figure 4.43. He did not, however, use this information to justify the trend he identified in the data. Although Rory made numerous statements about each graph that described appropriately the relationship evident, his statements were uni-structural. They concentrated on single aspects of the graphs and Rory did not bring his ideas together to talk about the related nature of the individual components described.

Johnty was able to make sound isolated statements about the graphs, the data, and the context of the data but did not draw these elements together effectively when making judgements about the relationship between two attributes. It was evident his understanding of covariation was under developed as used a covariation graph to compare two attributes (Figure 4.46). Johnty's statements about the covariation graphs were uni-structural and for the most part, relied on the comparison of individual data values or were based on the physical location of the data in the graphs. This was evident when he said a relationship existed between the attributes belly button height and height because the dots in the graph were "going up." When asked to expand on how the graph showed the relationship, Johnty did not use the information in the graph to justify his stance. Instead, he used his understanding of the context to explain, "The taller that you are, like, your belly button's not up as high." This statement revealed he understood that a person's belly button height could not be greater than that same person's height, which was the case for one of the cases in the data set. Further evidence of relying on the context occurred when he repeated "The taller you are, usually, the higher your belly button is," without referring to the graph or the data. This indicated that he knew there was a relationship between the two attributes from his knowledge of the context only as he did not used the graphs to support his conclusions. Each of these single statements was in response to a probing question from the teacher. None of Johnty's responses indicated that he had made connections between his successive statements.

The responses Natalie offered were predominantly uni-structural. Most of the time her references to the relationship between two attributes were isolated statements based on values of the attributes for one data point and did not include generalised statements about the group of students. The generalised statements that she did make were based on her knowledge of the context and did not make direct reference to information in the graphs. Her understanding that the value of one attribute may influence the value of another attribute was evident when she commented that a person with a belly button height of a couple of centimetres would be "*really* short." She also identified appropriately that "the smaller you are the smaller your foot length is." There were, however, instances when Natalie's responses were beginning to be multi-structural. This occurred when she was making a decision about an outlier. She said, "Umm, it's an outlier because of, like, you wouldn't think it would be 85 because she's 124 and I'm taller than her and mine is only 65," when referring to Graph 4 in Figure 4.40. This indicates that she was bringing together her personal knowledge of the context and information from the data to make her decision. On another occasion she said, "Umm, the boys have got bigger feet than what the girls do but the girls have got more people on, with the same height and the same foot length." These examples show that Natalie's thinking is grounded firmly in considering the value of specific data points or the comparative size of data points. Although the statements indicate she identified multiple aspects of the graphs that were useful, her descriptions and subsequent discussion of the information that followed did not indicate that she was bringing the information together to talk about the global relationship between two attributes in a meaningful way.

Shaun demonstrated that his comments about covariation graphs were multi-structural. He was able to think about multiple aspects of the graphs when making his conclusions but did not bring the information together effectively to demonstrate that his descriptions were relational. When asked if the covariation graphs showed a relationship between two attributes, Shaun made claims that were offered very cautiously even when a strong relationship was evident (Figure 4.44). His answers included, "sort of" and "there's a bit of a relationship." Although Shaun identified variation within the graphs that did not fit with his generalised conclusions, he was unable to make confident statements about the relationship in the graphs. In one exchange with the teacher, Shaun commented on the relationship evident in a graph by saying, "Umm, sort of … because in like this bit [points to a region of the graph] it's like the higher the height the bigger the belly button to the floor." He then went on to say that this applied most of the time and added, "there's only about 5 that aren't." During another exchange, Shaun identified that a "clump" in a graph showed "most of the people, the higher the foot they have, the higher the belly button to the floor." He also identified specific data points that were outside the main clump and looked at the values of the attributes for those data points. He did not, however, indicate how the additional information influenced his generalised statement.

Most of the time, Blaire made statements about the graphs that were multi-structural. She identified appropriately the relationship between two attributes in the covariation graphs and made connections to her existing ideas about the relationships from her personal knowledge of the context. She also identified aspects of the graphs that verified her initial conclusions. She did not, however, bring the information together effectively to make sense of the data. This was evident when she was asked about the relationship seen in a covariation graph displaying the attributes belly button height and height. She immediately moved on to say, "Ok, as your … of course you would already know that as you get taller your belly button height would get higher." Blaire then went on to declare that the mean belly button height would be in the middle of a cluster of data in the graph. Although adding the mean to the graph corroborated her assumption about the mean, she could not go further to use the mean to justify her generalised statements about the relationship shown in the graph.

Jessica's responses were multi-structural. When analysing a covariation graph she identified the relationship between the two attributes when she said, "That yeah, as you grow, your foot length usually grows too." When asked to describe how the graph showed this she went on to say, "Well like there … the largest person, their foot length is only about 25 cm whereas this person here, they have the largest foot length but they're not the tallest." Although Jessica's additional comment contradicted her global statement about the nature of the relationship identified, it does show that she was able to describe multiple aspects of the graph that could be used to justify or refute her conclusion. She did not, however, go on to use this extra information purposefully to do this. Jessica was making sense of the data by considering different aspects of the graph (for example see Figure 4.28) but she did not tie the information together successfully to make responses that were relational.

When analysing covariation graphs, Kimberley made uni-structural statements about the graphs she was analysing. The uni-structural statements included, "Mmm … well the higher the person is the longer the foot length … seems." She then backed up such statements with other uni-structural statements like "No [relationship], because this one here's the highest, and with about 185 but he's only got a length of 25 cm for the foot." This statement referred to an individual data point on a scatterplot (Figure 4.20). Kimberley read the values of the data point from the case card and pointed to its position on the graph. As Kimberley worked, she made multiple statements about the data that showed she could extract a variety of information from the graphs. The statements showed she had an appreciation that there was a relationship between the two attributes in a graph but her isolated statements were presented as a list of individual characteristics without connections being made between the statements or the multiple aspects identified.

James' statements about the data showed he had made good connections between the context of the data and the graphical representations constructed. He made generalised statements about the relationships between two attributes in the graphs, identified multiple aspects of the graphs to support his conclusions, identified regions of variation that did not fit with his conclusions, and made connections to his knowledge of the context to make sense of his conclusions. The way in which James put all the information together to justify and explain his reasoning shows that his thinking was relational. As an example, after stating "that well, like … when they were getting taller, their belly button height was getting bigger," James went on to explain, "And like from around here. [points to a region of the scatterplot] I think them ones are like taller, but their belly buttons are only 50 and 60. They'd have to be very short, like to have a belly button that small. Like these ones are all what you'd expect from like, that age [points to a cluster of data on the scatterplot]. Whereas like these ones out here … like aren't" [points to outliers on the scatterplot] (Figure 4.25).

The responses offered by William about covariation graphs were relational. He could identify and describe the relationship between two attributes and used the graphs effectively to support his thinking. At the beginning of the session he described the relationship between two attributes when he stated, "The bigger, the bigger you are the bigger your foot length is." He made this statement before he started to construct any graphs and based the statement on his prior knowledge of the context. During the session he showed that he could use the graphs and the data to describe the relationship between two attributes, which aligned with the statement expressed from the context only. William did this when he made direct connections between the data and the graphical representations. Initially, he used individual data points to support his statements within the graphical representations. For example, he said, "It tells us that that one's the biggest now and it's also the biggest foot length" as he pointed to an individual data point (Figure 4.48). On other occasions, he extended his explanations to include descriptions of clusters of data combined with descriptions of particular data points within a graph. He made generalised statements about the clusters of data and then used individual data points to support his generalised statement or show examples when the data varied from the generalisation expressed in his statement. Although William did not bring his personal knowledge of the context again into his comments about the graphs, it was evident that he understood how the graphs and the relationships displayed were linked to the context.

Mitchell made good connections between his knowledge of the context and the multiple characteristics of graphs to make statements about the data that were relational. On one occasion he determined that the relationship evident in a covariation graph that displayed the attributes height and foot length was the same as what was identified from another graph during a previous session (Graphs 2 & 4 in Figure 4.38). He said, "Yeah well … the … pretty much the same as before. The longer the foot length the higher you're going to go." Using the drawing tool, Mitchell added a trend line to the graph that showed the region on the graph that he was referring to. Following this example, Mitchell constructed a covariation graph for the attributes height and belly button height and commented, "Yeah I'd say, well, yeah, there's a couple of people that are off it a bit more but it's pretty much the same as foot length and height. The taller your belly button is the higher you are." This statement indicates that Mitchell had identified that the second graph displayed greater variation than the first graph but he also recognised that the trend in the second graph was similar to that in the first graph. This example shows that Mitchell was able to think about the data from a global perspective, identify the variation evident, and make comparisons to another graph about a different set of attributes.

Mitchell also used *TinkerPlots* in a relational manner. Often, he made conclusions about graphs quickly and articulated his thinking clearly. He then used the features of *TinkerPlots* to confirm his thoughts. He added hat plots, the mean, and reference lines to see if they added information that may be useful. Most of the time, they supported his initial responses and contributed to his confidence in his initial statements. The process he went through started with a conclusion, which he checked by using a feature of *TinkerPlots* such as the mean. Mitchell repeated this process using other features of *TinkerPlots*. As he worked he considered the new information and integrated it with his previous statements about the graph.

Summary of Student Reasoning about Covariation

It is generally assumed that young children's ability to evaluate relationships in data and reflect on supporting evidence is undeveloped (Koerber, Sodian, Thoermer, & Nett, 2005). The evidence from this inquiry suggests that students in the upper primary years are able to evaluate the relationship between two attributes and develop an understanding of covariation. The degree to which they explain and express that understanding varies and was evidenced at three levels of understanding. The varying levels of understanding of covariation indicate that some students were only able to draw on their intuitions developed through everyday experiences, whereas other students used a "mixture of everyday and deeply understood formal knowledge" as foreshadowed by Leinhardt et al. 1990, p. 24). Table 5.2 shows the names of the students who demonstrated learning outcomes at each of the levels of the SOLO framework.

Table 5.2.
Students' Achievement for Covariation According to the Levels of the SOLO Framework

Uni-structural	Multi-structural	Relational
Jake, Natasha, Rory, Johnty, Natalie, Kimberley	Shaun, Blaire, Jessica	James, William, Mitchell

At the beginning level, uni-structural, the students relied on their knowledge of the context of the data or individual elements of graphs to make their conclusions. They were able to offer information from the graphs to support their statements but often their justifications and examples were offered piece-meal without making any connections among the different constituent parts of the graphs and without making connections between the graphs and their knowledge of the context. The students could also identify a trend in a graph and use it to determine the relationship between two attributes but their explanations about how they used the trend to determine the relationship were incomplete. Often their statements involved a declaration that there was or was not a trend evident but little or no justification or reasoning was offered to explain how they made the judgement.

At the next level, multi-structural, the students considered multiple elements of graphs but did not elaborate on the connections between the graphs and their knowledge of the context. They used multiple constituent parts of their graphs to explain their thinking and to describe the covariation identified but their responses did not show they had an appreciation of the way in which different elements of the graphs were related and how they could be used to support conjectures and conclusions about the relationship identified. They were able to identify regions of the graphs that varied from the trend but were unable to explain how the variation informed their overall conclusions.

At the highest level, the ability of the students to bring together effectively the elements of the graphs and their knowledge of the context increases. The students at the relational level had developed more sophisticated ways of describing the relationship between two attributes. They did this from both global and local perspectives. They were able to identify a trend in a graph, identify the variation in the graph that did not meet their expectations for a relationship to exist, and use multiple constituent parts of graphs to support their reasoning. They interlaced their explanations with information about specific data points, the variation, and the trend evident, as well as made connections between the information in the graphs and their knowledge of the context to justify their conclusions.

Discussion of Student Reasoning about Covariation

This inquiry demonstrates how important it is for students to understand how graphs are constructed and the relationship between the different elements of graphs to be able to evaluate a graph to describe and determine covariation. The students operating at the highest level of the thinking for covariation were able to make the connections between the constituent parts of the graph, the context of the data, and the message in the data characterised by the trend. These students were able to express their thinking verbally and talked about their graphs from both a generalised and local perspective. This shows that it is important that teaching and learning activities are flexible enough to allow students to develop strategies for moving between thinking about what can be seen in the data from the generalised perspective to thinking about how they identified and formulated the relationship (Ben-Zvi & Acavi, 2001). Although being able to identify and express the relationship between two attributes as the trend seen in the data is important, it is equally important that students can describe how the more localised perspective of the spread of a cluster of data informs or challenges their generalised view. Placing learning experiences within interactive technological learning environments, such as *TinkerPlots*, and getting students to activate their thoughts and reasoning hones the focus in on what the graphs show and how they can be used by the students to evidence their reasoning. This takes away the need to learn specific graph conventions, which can be the focus of learning experiences at other times. The view that graph interpretation and graph creation can be dealt with separately for Year 8 students was proposed by Cobb, McClain et al. (2003). The findings in this thesis assert that the introduction of the concept of covariation can be adopted for upper primary students.

Rangecroft (1991a) suggested a sequence for the introduction of different graph types that started with lower primary students constructing bar graphs. She then offered a sequence that sees a glut of different graph types introduced at the beginning of high school with the introduction of scatterplots, stem-and-leaf plots, line graphs, pie charts, and box plots, across the senior years of schooling (1991b). The suggestions from Rangecroft are mirrored in curriculum documents and text books (for examples see Maths Quest 7-12 series at http://au.wiley.com) and to some extent follow the historical developments in graph construction (Watson & Fitzallen, 2010). This study showed that students in Years 5/6 can describe covariation and reason about covariation and generalise about a trend evident in the data. Considering that students in this study are able to work with covariation it may be pertinent to consider bringing the initial introduction of scatterplots into the upper primary years of schooling. This would provide students with the opportunity to establish the notion of covariation, how it is characterised in graphs, and how to reason about covariation well before the introduction of the statistical analysis of correlation and the translation of a trend into an algebraic expression in Year 10.

Postscript – Student Understanding of Association

The main focus of this inquiry was to investigate students' understanding of covariation when using *TinkerPlots*, as well as the way in which students engaged with *TinkerPlots* as a construction tool and as a thinking tool. In doing so, other aspects of exploratory data analysis behaviours were uncovered. This can be attributed to the design based research methodology adopted for this inquiry. In keeping with the methodology, an iterative approach was taken to analyse the data. A dominant theme that emerged from the data was the students' understanding of association. Including insights into students' understanding of association in this section demonstrates the utility of the methodology for uncovering insights that were unanticipated and shows how the methodology adopted for the inquiry can be used to build a more comprehensive understanding of a particular phenomenon. This section demonstrates the relevance of association to the main theme of covariation and as a postscript, is offered as a discussion point for the development of future research. Although students' understanding of association was not the main theme of the inquiry it is included in this section to sign-post that research into this aspect of students' statistical reasoning needs to be explored further as it deserves a deep separate examination that cannot be done within this inquiry.

Association is related to but distinctly different from covariation. Covariation graphs, such as scatterplots, have numerical attributes on both axes of a graph (Moritz, 2004). Association graphs have numerical data on one axis and categorical data on the other axis (Batanero et al., 1996). Whereas covariation allows for the relationship between two attributes to be explored, association graphs facilitate a comparison of categories or multiple data sets for the one numerical attribute. Association for two categorical variables can also be displayed in two-way tables (Batanero et al.), but that representation was not explored in this inquiry.

Graphs that display directly the comparison of groups are created easily in *TinkerPlots*. The representation suited to this purpose is the split stacked dot plot. Figure 5.1 has an example created by Johnty when he was exploring the relationship between gender and height.

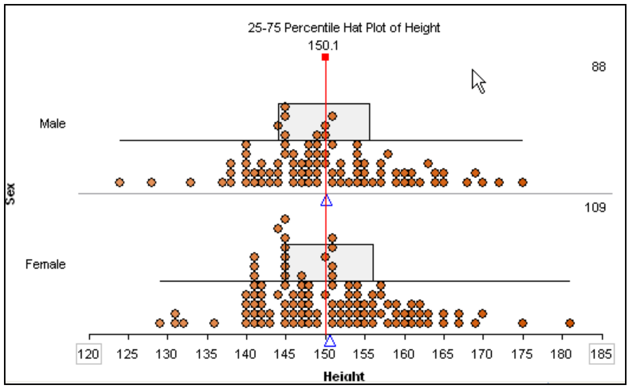


Figure 5.1. Split stacked dot plot displaying the association between gender and height.

The final task in the interview protocol required the students to look at all the graphs created during the session and determine which graph of those created showed the strongest relationship between two attributes. The students were asked to choose the graph that they were most confident showed a strong relationship. Of the 10 students that completed this task, seven chose graphs that showed an association between two attributes rather than graphs that displayed covariation. This indicated that some of the students had not established fully the purpose of the different graph types and the meaning they embodied.

From the selection of graphs in Figure 4.38, Mitchell selected Graph 1 and asserted that the graph selected showed the strongest relationship between height and gender. Graph 1 is reproduced in Figure 5.2. Mitchell determined that the males and females heights were the same by comparing the two groups using the mean, the position of the crown of the hat, and the spread of the data. He equated the closeness of these characteristics of the data to infer there was a strong relationship evident in the graph. By convention, interpretation of this graph would show there was no connection between the gender and the height of the students.

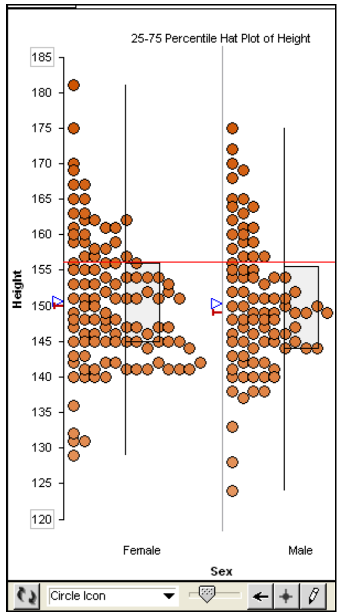


Figure 5.2. Mitchell's association plot displaying the relationship between height and gender.

Johnty, Rory, and Blaire also chose graphs that displayed the association between gender and height to be the one that showed the strongest relationship between two attributes. Blaire used the mean to make her decision. As the mean height was the same for both genders, Blaire said she was confident it was the best graph. Johnty used the mean as well to make his conclusion but also used the crown of the hat to support his decision. As Rory made his choice he said, "The hat plots [are] exactly the same. So is the average."

Jake and Natalie also used the crowns of the hats to justify their choices of a split stacked dot plot that displayed gender and foot length. Adding to Natalie's confidence was the spread of the data. She said, "It is not as spread out [pointing to a split stacked dot plot]. This one is not as neat and you don't know how many dots there are [pointing to a scatterplot]." Shaun selected a graph with gender and belly button height and said, "They've got the same amount of people like under the [hat] roughly about 50% and roughly about the same … outside the 50%."

William, Natasha, and James selected scatterplots that displayed the relationship between belly button height and height to be the ones that showed the strongest relationship between two attributes. In all three cases, the decision was based on the trend evident in the data. As William pointed to the scatterplot with height and belly button height, he said:

"Umm … probably this one, because umm, it shows us that the strongest, was umm, it shows you that the height, depending on where it is, chances are that that's where the bigger belly button height is probably." Jessica and Kimberley did not make a contribution to this discussion as they ran out of time and did not complete the final question on the interview protocol.

Discussion of Student Understanding of Association

Although all of the students constructed split stacked dot plots, the plots varied somewhat. The students individualised their graphs by accessing different features of *TinkerPlots*, such as the mean, hat plots or reference lines. Regardless of the features added and the scale of the axes selected for the split stacked dot plots, each of the students indicated that it was the similarity or closeness of the information from the *TinkerPlots* features in each section of the graphs that influenced their response. As can be seen in the split stacked plot displayed in Figure 5.2, the mean for the males and females is the same and the range of the crown of the hat is also very similar for both genders. When making the decision about the relationship between two attributes from this type of graph, the students gave little attention to the distribution of the data across the outer reaches of the hat brim. On occasions, the overall range of the data for each category was considered but students only mentioned the range of the data when there was a large difference in the ranges of each of the categories.

Konold (2002) and Cobb et al. (2003) suggest that splitting a graph into slices similar to the display in Figure 5.3 provides an alternative to the conventional scatterplot for determining the relationship between two attributes. The graph displays the attribute height split into bins or slices as well as the mean added for each slice of data. Analysis in the change in the mean across the range of the data indicates a trend of increased heights associated with increased belly button height. Figure 5.4 is another variation of this graph type but the slices in the data are horizontal. The graph type displayed in Figures 5.3 and 5.4 is appropriate when the two attributes are numerical. A number of successive categories can be compared visually to make a judgement about the relationship between the two attributes. A problem arises when one attribute is categorical (unordered) and the categorical data are limited to two categories as displayed for gender in Figure 5.1. When this is the case the data in the two categories can be compared but they cannot be used to determine if there is a relationship between the two attributes in the same way as occurs with scatterplots. It is possible, however, to use representations similar to Figures 5.1 and 5.2 to determine if two groups are the same or if there is a difference between two groups.

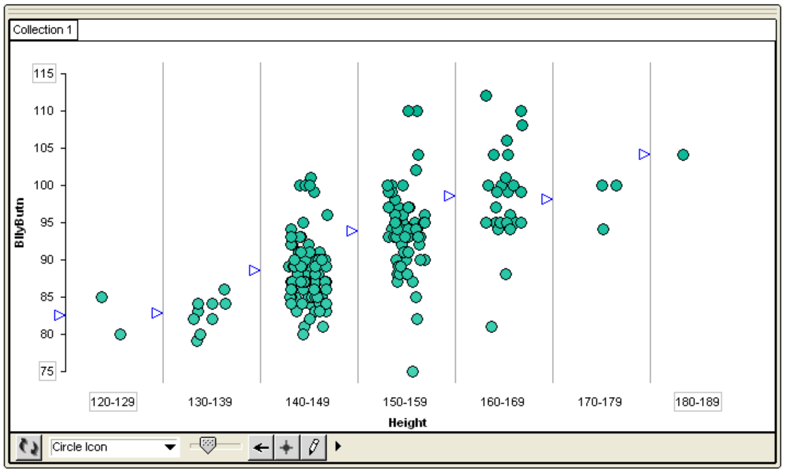


Figure 5.3. Scatterplot with horizontal axis split into bins and the mean inserted.

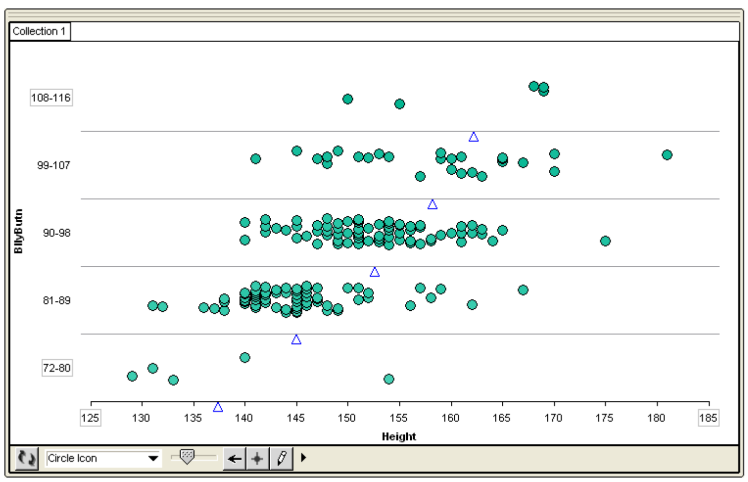


Figure 5.4. Scatterplot with vertical axis split into bins and the mean inserted.

The students in this inquiry used association graphs to make decisions about the relationship between one numerical attribute and one categorical attribute. This indicated that they did not realise the limitations of the data, nor did they appear to acknowledge that they needed to use different thinking strategies when they analysed different graph types. When they analysed the split stacked dot plots that displayed the association between two attributes as in Figure 5.2, they compared the data using relative frequencies as foreshadowed by Batanero et al. (1996). When they analysed scatterplots they used relational thinking to determine if there was a trend evident (Mortiz, 2004; Zeiffler & Garfield, 2009). Understanding the different thinking provoked by the different representations has the potential of empowering teachers to guide student learning more effectively towards an understanding of the particular learning outcomes targeted. Teachers will then be in a position to choose the most appropriate data sets and graphical representations whether the learning outcome is to be understanding of covariation or association.

The results about students' understanding of association presented in this section are limited. Although the examples given are indicative of the behaviours exhibited by the students, a comprehensive analysis of the data was not undertaken as learning about association was not a primary focus of the Sequence of Learning Experiences developed for this inquiry. In hindsight, it would have been advantageous to include association as a focus in order for the students to develop an understanding of the application of the different graph types. It was, however, encouraging that the students had developed the confidence to try to transfer their knowledge of covariation to a variety of graphical representations.

Concluding Remarks

A purpose of graphs in this inquiry is to provide visual representations for conducting exploratory data analysis using data gathered from a context that was meaningful to the students. Three main elements contribute to the data analysis process – the image of the graph, the characteristics of the graph, and the context of the data. The overall image of a graph provides one representation that can be used to make sense of the data. It is a picture that is viewed from a global perspective and is often used to identify a trend in the data. A graph also has many characteristics that can be used when thinking about the data. The characteristics, such as the mode, scale of an axis, or the variation in the spread of the data can be extracted directly from the graph or from calculations performed by graphing software, such as *TinkerPlots*. In these instances, the software is a tool that is used to gain access to the characteristics of the graph. Another factor that influences conclusions that are made about graphs is the context of the data. An understanding of the context may be gained from personal experiences of the context or from other information about the data. Interpretation of the context, however, is achieved by employing EDA strategies.

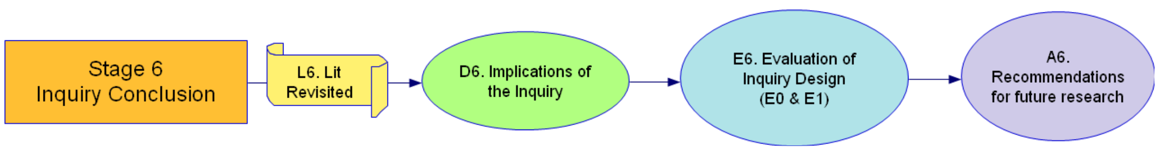
This inquiry explored the connections and interconnections among the elements of the data analysis process when using *TinkerPlots* from three perspectives, which were included in the three research questions. First, it applied a theory/praxis framework, *Model of Learning Behaviour in EDA Graphing Environments,* developed for this inquiry to construct Student Profiles that characterise students according to the four dimensions of the framework. This thesis acknowledges that future development may include aspects of sense making, statistical reasoning, and discourse as a way to deepen understandings of students the development of students statistical thinking and reasoning (Watson, 2006).

In this inquiry, Student Profiles were used as data to determine the strategies the students used when engaging with *TinkerPlots* and their development of understanding of covariation. The insights gleaned from the three perspectives from this inquiry, collectively, provide a comprehensive view of student learning about covariation in the learning environment afforded by *TinkerPlots*. Although interrogated separately, the three perspectives are closely linked and intrinsically entwined. On a broader scale, the three perspectives cover the issues of theory, pedagogy, and content as they relate to student learning about the statistical concepts within data analysis software environments.

Also gleaned from the study were findings about students' understanding of association (Batanero et al., 1996; Moritz, 2004; Zieffler & Garfield, 2009). The findings highlight the difficulty students have when applying their knowledge about one statistical concept that relies on particular graphical representations to other graphical representations that do not embody the same meaning. These findings also highlight the need for student learning about statistical concepts and data analysis skills to be developed over time to ensure they establish a comprehensive knowledge of graph creation and graph interpretation strategies.

Stage 6

Inquiry Conclusion



The aim of this inquiry was to explore student learning about the statistical concept of covariation in the technological learning environment afforded by the software package, *TinkerPlots.* The purpose of the inquiry was to extend understanding of the factors that influence student learning when working with software packages and to contribute to understanding how students reason about covariation. Twelve students worked through a sequence of learning that provided the opportunity to gain insights into the way the students used *TinkerPlots* to create graphs and then reason about data as they interpreted the graphs. Central to the inquiry was the conceptual framework, *Model for Learning in EDA Graphing Environments*, developed in Stage 1 of the inquiry. The four dimensions – *Generic knowledge*, *Being creative with data*, *Understanding data*, and *Thinking about data* – provided the lens through which the students' prior knowledge about graph creation and interpretation, their interaction with *TinkerPlots*, and their learning about covariation was examined. This inquiry meets recommendations from Ben-Zvi (2004) and Shaughnessy (2007) who suggest that pedagogical practice about the ways in which students' reasoning about data analysis develops should be examined.

Two objectives underpinned the inquiry:

- First, a theoretical objective, to develop a conceptual framework that characterises learning in exploratory data analysis (EDA) graphing environments that also aligns with and extends current research about graphing and data analysis.
- Second, a practical objective, to contribute to an understanding of the complex nature of the teaching and learning of data analysis skills, particularly covariation, within the EDA software environment afforded by *TinkerPlots*.

This stage of the thesis summarises each stage of the inquiry to indicate the way in which each stage contributed to the objectives of the inquiry and goes on to consider the implications of the inquiry for the future curriculum development, pedagogical practice, and the application of technology in the teaching and learning of statistical concepts. Following those sections an evaluation of the inquiry design is presented. The limitations of the inquiry are addressed throughout the evaluation of the inquiry section.

Summary of the Inquiry

Stage 0 – Inquiry Commencement

Enabling the inquiry by selecting the framing methodology and developing the inquiry design was implemented in Stage 0. In order to accommodate the coupling of the theory building aspects of Inquiry Objective 1 and the practical focus of Inquiry Objective 2 a methodology that was qualitative, flexible, and adaptable was needed. Educational design research methodology was deemed suitable as it provided an avenue for theory building (Akker et al., 2006; Sandoval, 2004) while concurrently investigating student learning in technological learning environments (Seeto & Herrington, 2006). The four phases of educational design research that align with the integrative learning design (ILD) framework developed by Banna-Ritland (2003) were adapted to suit the short term nature and context of this inquiry and then used to guide the inquiry design. The phases developed for this inquiry were:

- Analysis of practical problems;
- Development of solutions with a theoretical framework;

- Evaluation of solutions; and
- Application of solutions and reflection on implementation (Table 0.1).

The four phases were used to construct the inquiry design that evolved through seven stages, with each stage utilising the phases of the ILD framework relevant to that stage (Figure 0.2). The selection of the methodology and the construction of the inquiry design set up the inquiry to meet Inquiry Objective 1 and work towards meeting Inquiry Objective 2.

Stage 1 – Development of Model of Learning Behaviour

The main focus of Stage 1 of the inquiry was the development of the *Model of Learning Behaviour in EDA Graphing Environments* (Figure 1.7). It was determined that a new conceptual framework was needed as the existing models of graphing did not adequately take into account the way students work when creating graphs and the potential of technological graphing environments currently available. Interrogation of the literature on models of graphing revealed that a number of the models were similar and were based on the work of Curcio (1989) (cited in Shaughnessy et al., 1996). The dimensions of statistical thinking offered by Pfannkuch and Wild (2004) provided a comprehensive model of statistical thinking that was too broad a perspective for this inquiry but included behaviours associated with constructing graphs. Moritz (2004) also offered a model that encompassed the translation process involved when reasoning about covariation but again, the model he offered was not relevant in its entirety and like the other models did not consider technological learning environments directly. To address the shortcomings of the models interrogated, the *Model of Learning Behaviour in EDA Graphing Environments* (Figure 1.7) was developed. Relevant elements from the existing theoretical frameworks were incorporated into the new model. The development of this conceptual framework addressed directly Inquiry Objective 1 as it was aligned with and was built upon from current research about graphing and data analysis.

To meet the second inquiry objective, this stage identified the characteristics of graphs and sought to understand the development of graphs and graphing from an historical perspective. The main developments since the 1600s were explored briefly, with more attention given to the EDA strategies developed in recent times. EDA strategies developed by Tukey in 1977 broadened the view of graphing and provided alternative techniques for conducting data analysis (Hartwig & Dearing, 1979). This stage then looked at the literature on covariation to develop an understanding of how it is conceptualised and how it is represented in graphs. An anomaly that arose from the literature was the way in which covariation was described. Often it was referred to in conjunction with the term "association" and clear distinctions between covariation and association were not made. This inquiry found that the research that focused on middle years students' understanding of covariation was limited; however, the following main themes emerged from the studies.

1. Students often focus on individual data points rather than look more broadly at the global trend in the data.
2. Students' prior knowledge about the context can enhance or inhibit students' reasoning about data.
3. Students' intuitions about covariation provide a springboard for extending their thinking further.
4. Students' understanding of covariation should be built upon an initial understanding of variation and distribution.

Stage 2 – Development of Model of Learning Behaviour

The software package, *TinkerPlots,* was provided for this study by the publishers Key Curriculum Press. Although it was purported to be a good innovation (Konold & Miller, 2005), this stage of the inquiry sought to determine what was good about the software (Ferdig, 2006). An examination of the literature revealed that existing frameworks had been developed to evaluate web-based learning objects (Handal et al., 2006). Although used for that purpose the existing frameworks did not take into consideration the interactive nature of some new technologies nor did they accommodate the learning environment landscape-type software such as *TinkerPlots* offers. To evaluate *TinkerPlots* comprehensively, it was determined that a new set of criteria needed to be developed. Relevant elements from existing evaluation frameworks and the *Model of Learning Behaviour in EDA Graphing Environments* (Figure 1.7) were drawn together construct the *Criteria for Evaluating Exemplary EDA Software* (Table 2.2). Similar to the *Model of Learning Behaviour in EDA*

Graphing Environments, the *Criteria for Evaluating Exemplary EDA Software* contributed to Inquiry Objective 1.

The application of the *Criteria for Evaluating Exemplary EDA Software* (Table 2.2) to evaluate *TinkerPlots* in this stage of the inquiry contributed towards Inquiry Objective 2. The evaluation of *TinkerPlots* found that the software package provided an easy to use, dynamic, and interactive learning environment for students to engage in when analysing data and creating graphical representations. The successful application of the criteria to evaluate *TinkerPlots* also confirmed the utility of the set of criteria as an evaluation instrument.

This stage of the inquiry explored the literature to determine the way in which *TinkerPlots* was developed, used in classrooms, and applied in research studies. The applicability of *TinkerPlots* for the purpose of developing an understanding of covariation was exemplified in the insights gleaned about how learners use and interact with *TinkerPlots.* The literature review also provided valuable information about the type of questions and the context of statistical investigations students and teachers explored with *TinkerPlots*. Overall, this information contributed to Inquiry Objective 2 as it demonstrated the flexibility of *TinkerPlots* to accommodate different users' learning and facilitate the learning of a variety of statistical concepts in many and varied contexts.

Stage 3 – Establishment of Student Prior Learning

Developing and implementing a student survey for the purpose of identifying prior knowledge about graphing and understanding of covariation of students participating in the inquiry was the focus of Stage 3. Determining student prior learning is a fundamental aspect of a student-centred pedagogy (Woolfolk & Margetts, 2007; Van de Walle, 2007). It supports teachers to develop appropriate learning experiences for students to ensure student learning builds on and exploits what students know already. It potentially identifies student misconceptions and gaps in knowledge. The evaluation of student responses on the survey contributed directly to Inquiry Objective 2. From the student survey it was found that most of the students could construct a column graph and read values from a pictograph. The majority of the time they recognised easily one attribute from examples of bivariate data and scatterplots but had difficulty making connections to the second attribute. There were indications that their experiences with graphing software was extremely limited as were their experiences with identifying trends in the data and making informal inferences based on the data. Many students did not use the data and graphs to inform their decisions; they relied on their knowledge of the context of the data to make judgements about and draw conclusions from the data. These results reflect those gleaned from the studies that developed and utilised the items originally (e.g. Moritz, 2003a, 2003b; Watson & Kelly, 2003). Finding that the group of students participating in the inquiry were likely a typical sample from the Student Survey indicated that insights gleaned from the students would be useful when thinking about how students in Year 5/6, in general, reason about covariation.

Stage 4 – Sequence of Learning and Outcomes

Stage 4 contributed to both inquiry objectives concurrently. Basing the development of the sequence of learning on theoretical frameworks and exemplars of learning activities followed by its implementation in this inquiry exemplified the way in which theory intersects with praxis (Kemmis, 2010). The development of a sequence of learning using theoretical frameworks from the literature demonstrated the advantages of intentionally using a synthesised theoretical position to guide teaching practices. This was also reflected in the use of the *Model of Learning Behaviour in EDA Graphing Environments* (Figure 1.7) to evaluate the student work to develop student profiles. The student profiles provided rich descriptive data that were then used to answer the three research questions in Stage 5.

Stage 5 – The Findings

The purpose of Stage 5 was primarily to answer the three research questions. Research Question 1 focused on the development and utility of the *Model of Learning Behaviour in EDA Graphing Environments*. The model was needed as previous models for graphing did not take into consideration the contribution students make to the graph creation process, nor did they take into consideration the interactive nature of graphing software packages available today. The *Model of Learning Behaviour in EDA Graphing Environments* was used repeatedly throughout the inquiry and was found to be robust and reliable as a framework for describing students' thinking, evaluating students' work, and progressing the theoretical understanding of student learning in EDA graphing environments. It is anticipated that the framework would contribute to assisting teachers and researchers alike.

Research Question 2 focused on students' interaction with *TinkerPlots* as they developed an understanding of the software package in the first instance, and then used *TinkerPlots* to further their understanding of graph creation and graph interpretation, particularly an understanding of covariation. The students demonstrated that they interacted with *TinkerPlots* in different ways as they created a variety of graph types, analysed data to identify trends in graphs, and described relationships between attributes. The three dominant strategies employed were: *snatch and grab*, *proceed and falter*, and *explore and complete*. Further research is required to determine if the characteristics of computer interaction described from this inquiry are related to the cognitive development of students.

Covariation was the big statistical idea that was the focus of Research Question 3. The Year 5/6 students in this inquiry reasoned about covariation in various ways and demonstrated an understanding of covariation at three levels. At the univariate level, students demonstrated that they had intuitive notions about covariation. They were able to identify the relationship between two attributes by recognising a trend in a graph but did not elaborate further or try to justify their conclusions. At the multivariate level, students were able to use multiple elements of graphs and their personal knowledge of the context to describe the relationship identified in a graph and to describe variation in the data. They did not, however, indicate that they had made connections among the ideas that they expressed about the relationship identified. At the relational level, students not only were able to use multiple elements from the graphs and their personal knowledge of the context but also were able to make connections among the elements to explain and justify their conclusions that a relationship between two attributes was evident. These results signify that it is appropriate for students in the upper years of primary school to explore covariation and use their understanding of covariation to draw conclusions about statistical questions.

Implications of the Inquiry

This section of the thesis discusses the implications of the inquiry in relation to the new Australian curriculum, *The Australian Curriculum – Mathematics* (ACARA, 2011) and makes recommendations for future curriculum development. The findings of the inquiry also indicate there are implications for teachers' pedagogical practice. This inquiry acknowledges that the teaching and learning of statistical concepts and data analysis using EDA software are intrinsically intertwined and influenced by the knowledge and confidence of the teacher. Hence, the implications associated with classroom practice are discussed in terms of teachers' technological pedagogical content knowledge (TPCK) (Mishra & Koehler, 2006; Niess, 2005; Zhao, Pugh, Sheldon & Byers, 2002). Following on from this, the implications for the choice of software to conduct EDA activities are presented. An emphasis is placed on the importance of basing the selection of a software package on its ability to engender understanding of graph creation and graph interpretation skills. This inquiry puts forward a case for the selection to be *TinkerPlots*.

Implications for Curriculum Development – A Case for the Big Ideas of Statistics

Technology has the potential to improve student learning about statistical concepts. To ensure teachers are empowered to change their pedagogical practices accordingly, curriculum frameworks need to be re-visited to ensure they reflect the relevant content and incorporate technology in ways that assist teachers to envisage new practices (Chance et al., 2007). Garfield and Ben-Zvi (2004) argue that learning experiences should focus on the big ideas of statistics, such as association, distribution, covariation, and inference. This has implications for curriculum frameworks written for the Australian context. To be helpful, curriculum frameworks need to use specific statistical language to assist teachers to make the connections between the content in the curriculum and the broader statistical concepts with which they are associated. In the U.S., the *GAISE Report* (Franklin et al., 2005) was written to assist teachers to interpret the NCTM's *Principles and Standards for School Mathematics* (2000) as it was identified that some teachers did not "see the statistics curriculum strand for grades K-12 as a cohesive and coherent curriculum strand" (Franklin et al., 2005, p. 5). The *GAISE Report* positions the use of technology, covariation, and other statistical concepts explicitly within a developmental curriculum framework that "provides a conceptual structure for statistics education that gives a coherent picture of the overall curriculum" (p. 5).

This inquiry based the choice of content to be explored on curriculum frameworks that were relevant when the inquiry began. These were sourced from Australia (AEC, 1991, 1994; DoET, 2007) and the U.S. (Franklin et al., 2005; NCTM, 2000). In Australia in 2011, a new national mathematics curriculum, *The Australian Curriculum – Mathematics* (ACM) (ACARA, 2011) was introduced. One of the aims of introducing a national curriculum was to provide greater consistency in the delivery of the curriculum across all the states of Australia. Previously, individual states were responsible for the curriculum; however, this was seen as problematic as the delivery of the content in the curriculum did not always align across the states. Although there is merit in having a national curriculum, the new mathematics curriculum has resulted in major changes to the delivery of statistical content in Tasmania, and Australia more broadly (Watson & Fitzallen, 2010).

The results from this inquiry indicate that students in Year 5/6 are able to reason about covariation; yet, the ACM (ACARA, 2011) does not introduce covariation in the curriculum until Year 10. The ACM includes the content strand Statistics and Probability. Within that content strand, the sub-strand Data Representation and Interpretation, sequences graph creation, graph interpretation, and data analysis across all the compulsory years of schooling (Foundation Year – Year 10). The word covariation is not used in the curriculum document but is implied in Year 10 when it is recommended that students "Use scatter plots to investigate and comment on relationships between two continuous variables" (p. 46) and "Investigate and describe bivariate data where the independent variable is time" (p. 46).

On one hand, the language used in the ACM (ACARA, 2011) is non-threatening and would not intimidate teachers who lacked confidence in teaching statistical concepts but on the other hand, the ACM is not helpful for teachers who want to provide students with holistic learning experiences across the years that incorporate all aspects of working and thinking statistically (Watson, 2009). The application of graphs described in the ACM is expressed in general terms and is primarily used to "identify and investigate issues." The skills listed to do this include: "construct," "describe," and "interpret" graphs and data. The ACM does not, however, link the application of different graph types with the big ideas of statistics that they each promote.

The ACM (2011) is explicit about when students should use bivariate, continuous, and categorical data. It is also explicit about the graph types to be employed in particular years. The staging of the introduction of graph types aligns with the suggestions from

Rangecroft (1991a, 1991b) and mirrors the sequence of graph developments, historically. The historical developments of graph types were largely dependent on the technologies available for data collection and data analysis at those times (Beniger & Robyn, 1978; Friendly, 2007). It is, therefore, logical that the complexity of graph types increased over time as technologies became more powerful. By introducing the graph types in much the same order as the historical developments, the ACM is recognising that different graph types are more complex to create but is not acknowledging that the technologies used to create the graph types originally are now available ubiquitously.

New educational technologies, such as *TinkerPlots*, make complex graph types accessible earlier in the learning progression than they are currently staged. There is hence, no longer the need to introduce them in the curriculum in the same order they were historically developed. Leaving the introduction of scatterplots until Year 10 excludes the opportunity to build on the capacity of younger students, as this thesis demonstrated for Year 5/6 students, to develop intuitive ideas and foundation knowledge about covariation. Including covariation earlier in the curriculum and staging its development across all the compulsory years of schooling would provide the opportunity for students to develop intuitions in the first instance, and then develop a more comprehensive understanding of the underpinning ideas of covariation before encountering the more formal application of covariation and correlation in Year 10 and beyond.

Competence in using information and communication technology (ICT) is included in the ACM (ACARA, 2011) as one of the seven General Capabilities. The intention is that students develop and use ICT and the other general capabilities "across learning areas and in their lives outside school" in order "to succeed in life and work in the twenty-first century" (p. 8). The descriptor for the ICT competence is general and broad. Specific reference is, however, given to the application of "complex graphical and CAS calculators [to] be used to make calculations, draw graphs and interpret data in ways that have previously not been possible" (p. 9). Other than a brief mention of the use of spreadsheets, no further guidance is given for the application of digital technologies that facilitate graph creation and data analysis. Graphing software is only referred to in the ACM in relation to the graphing of linear and non-linear functions. In relation to the content strand of Statistics and Probability, and the sub-strand Data Representation and Interpretation in particular, the term "technologies" only appears in four year level descriptors – Year 3 (ACMSP069), Year 4 (ACMSP096), Year 5 (ACMSP119), & Year 10A (ACMSP279).

The ACM (ACARA, 2011) focuses on the construction of a variety of graph types and makes cursory references to the interpretation of graphs. It lacks direction for the use of technologies and makes no connections between particular graph types and the statistical thinking they evoke. As a result, the focus in classrooms is on students constructing graphs by hand and the importance of the mechanics of graphing supersedes the thinking about the statistical ideas that are amplified by the use of technology (Chance et al., 2007; Cobb, 2007). As evidenced in this inquiry, students in primary school rarely get beyond the construction of bar charts. This was determined when the students' prior knowledge of graphs and graphing was explored in Stage 3. It was also evidenced in this inquiry that primary students are able to use EDA graphing software effectively to construct graphs and then use the technology to support and justify their thinking about and interpretation of data. This inquiry demonstrated through the Student Survey (Stage 3) results that most of the students could perform the basic fundamental tasks associated with creating and interpreting graphs, albeit predominantly associated with bar charts and column graphs. This was not unexpected as current curriculum documents used at the schools designate that students in Years 5/6 construct and interpret these types of graphs. The results of the inquiry presented in Stage 5 of the thesis indicate that some students were able to extend beyond developing procedural knowledge of creating and interpreting graphs with *TinkerPlots* and engage in higher-order thinking about the graphs and the different graph types used. It is, therefore, plausible that more can be demanded from students in the upper primary years of schooling. In order to encourage the development of classroom activities that provide the opportunity for students to extend their learning beyond the construction and application of bar charts and column graphs, curriculum documents need to be more explicit about using a variety of graph types earlier on in the curriculum and include opportunities for students to engage in more meaningful statistical investigations and learning than is currently expected.

Future curriculum development in statistics education needs to come to terms with the way in which EDA software changes educational learning environments and influences student thinking. The ACM (ACARA, 2011) is a passive framework that does not promote the use of technologies. It recognises they are available but does not provide information that positions learning within those technological learning environments that is useful for teachers. It is understandable and right that the ACM does not name any particular software package but providing indications of the type of learning expected for a particular statistical concept from using technologies would assist teachers to select the appropriate technologies that provide those opportunities. The ACM proclaims that its goal for the use of ICT is to support students to attain learning needed to be active citizens in the twenty-first century. This is congruent with the purpose of supporting students' development of understanding of statistical concepts (Watson, 2006).

It is time curriculum frameworks, such as the ACM acknowledged how students learn and develop statistical thinking and reasoning when engaged with twenty-first century technologies. This inquiry proposes that covariation could be incorporated into the curriculum prior to Year 10. In order to do this the language in the curriculum would need to change to include formal statistical terminology to support teachers to make the connections among the graph types used, the statistical concept they embody, and the statistical thinking they provoke. These issues pose considerable challenges for curriculum writers. Curriculum writers not only have to consider when covariation and the other big ideas of statistical thinking and reasoning should be introduced in the curriculum but also have to consider how these underpinning ideas can be developed and sustained over the full extent of the compulsory years of schooling.

Implications for Classroom Practice – A Case for TPCK

Technological pedagogical content knowledge (TPCK), in broad terms, is the integration of knowledge of subject matter with knowledge of technology and knowledge of teaching and learning (Mishra & Koehler, 2006; Niess, 2005). Mishra and Koehler argue that TPCK advances the notion of pedagogical content knowledge (PCK) introduced by Shulman (1986) to include the learning domain of technology.

PCK involves "an understanding of how particular topics, problems, or issues are organized, represented, and adapted to the diverse interests and abilities of learners, and presented for instruction" (Shulman, 1987, p. 8). PCK is exemplified when pedagogy and content are completely intertwined. This occurs when teachers' exhibit

> knowledge of models of key ideas, understanding [of] what makes topics easy or hard to learn, knowledge of common student conceptions, … knowledge of deep and complete explanations, the ability to select appropriate representations to convey key ideas, and an awareness of why confusions and misconceptions are likely to occur. (Chick, Baker, Pham, & Cheng, 2006, p. 297)

The TPCK framework described by Mishra and Koehler (2006) suggests that teachers' knowledge of technology overlaps with the pedagogy and content aspects of PCK. They contend that their framework "emphasizes the connections, interactions, affordances, and constraints between and among content, pedagogy, and technology … and emphasizes the complex interplay of these three bodies of knowledge" (p. 1025).

It could also be argued that PCK acknowledges the role of technology as a resource and there is no need to extend the term to TPCK. Chick et al.'s (2006) elements of PCK include descriptors associated with teaching strategies, student thinking, cognitive demands of tasks, appropriate and detailed representations of concepts, knowledge of resources, and purpose of content knowledge. All of these elements are relevant whether using a number line for understanding place value or *TinkerPlots* for understanding covariation. When using technology, it is the resource that is different, not the fundamental conceptualisation of teaching and learning that PCK acknowledges. The difficulty with taking such a standpoint is that many teachers see technologies as being external to their everyday practice and not central to the teaching and learning of particular concepts (Mishra & Koehler, 2006). It is not until technologies become transparent in pedagogical practice that PCK can stand on its own when using technologies. As an interim, the stance of the TPCK framework can assist teachers to draw together their knowledge of best practice and integrate it with their knowledge of technologies.

Determining whether TPCK is an extension of or a subset of PCK is open for interpretation and beyond the scope of this inquiry. This does not exclude the need to unpack what educational technologies bring to learning experiences and how they conceptualise and represent the mathematical ideas they embody, in order for teachers to use them effectively to guide, support, and develop student learning of particular mathematical content. To design effective learning experiences with EDA graphing software for students, teachers need to have extensive knowledge about the specific learning environment afforded by the particular software, know how that software influences students' thinking and know how to use the software effectively to develop students' understanding of statistical concepts, as well as have a knowledge of the content to be developed. This inquiry has exemplified the possibilities for using EDA software when such an approach is taken.

The choice of which statistical software packages best suit the facilitation of particular learning outcomes complicates the work of teachers. This inquiry offers the set of *Criteria for Evaluating EDA Software* (Table 2.2) as a framework for interrogating the affordances of statistical software. Applying the criteria would enable teachers to choose the software most appropriate for the learning required and provide the necessary arguments to justify expenditure on software not freely available. Most importantly, applying the criteria would invite teachers to bring together their knowledge of content, pedagogy, and technology as a means for establishing or consolidating their TPCK. Well established TPCK would empower teachers to shift their application of technology from a just-in-time approach that relies on assistance from students to work out how to use the software, to adopt a more considered approach that makes the most of innovative technologies like *TinkerPlots*.

Implications for the Choice of Software – A Case for TinkerPlots

The choice of which software students use should be based on the ability of the software to construct the graphical representations required and support students' thinking and reasoning when analysing data as well as their development of conceptual understanding of statistical concepts and ideas. The influence of *TinkerPlots* on students' thinking highlighted in this inquiry indicates that the students' learning was bound to the context of the learning environment. This has implications for the choice of software to be used by students.

Most commercial statistical packages, educational data analysis tools, web- or computer-based applets, and simulation software can be described as route-type or landscape-type software. Route-type software "fit in a particular learning trajectory; and landscape software … provide an open landscape in which teachers and students may freely explore data" (Garfield & Ben-Zvi, 2004, p. 402). Bakker (2004) warns that the use of landscape software may give students too much choice and is distracting. He also advocates that

> The choice of one or the other program also has to do with the aim. If we want to guide a class as a whole towards understanding specific notions and graphs, then a route-type tool might prove more suitable. When aiming at genuine data analysis with multivariate data sets from the start, a tool like Tinkerplots is more appropriate. (Bakker, 2002, p.5)

Some software such as commercial statistical software packages, provide a learning environment that may be considered a middle ground between route-type and landscape-type software. “They offer a variety of simultaneous representations that are easily manipulated and modified, as well as simulation of different distributions” (Garfield & Ben-Zvi, 2004, p. 402) but do not offer the freedom of landscape software as graph types are selected from a set range of graphs and once selected cannot be manipulated or changed. To change an attribute in a scatterplot, for example, a new graph must be constructed. The effective use of this type of software relies on the students knowing which graph type best suits their purposes before creating the graphs and having an understanding of the difference between categorical and continuous data.

The data analysis tool available on the U.K. CensusAtSchool website (http://www.censusatschool.org.uk/) is an example of a web-based statistical software applet that is similar to commercial software packages. As can be seen in Figure 6.1 a selection of graphs to choose from is provided. The list on the left indicates the variables in the data set. The continuous attributes are coded yellow, whereas the categorical data are coded pink. Each graph type has a colour coded bar to indicate the type of variable to add to each graph type. To create a graph, one or more of the variables are dragged onto the graph type chosen and the software generates the completed graph. The colour coding of attributes is useful as it provides a hint for selecting the graph type and which attributes to use for that graph type. The software, however, is restrictive as it generates particular graphs at a time and does not allow the transition from one graph type to another to occur as is possible with *TinkerPlots*. The U.K. CensusAtSchool graph selection is superior to other similar commercial statistical software packages as it will not generate 3-D plots and starplots unless three continuous attributes are selected.

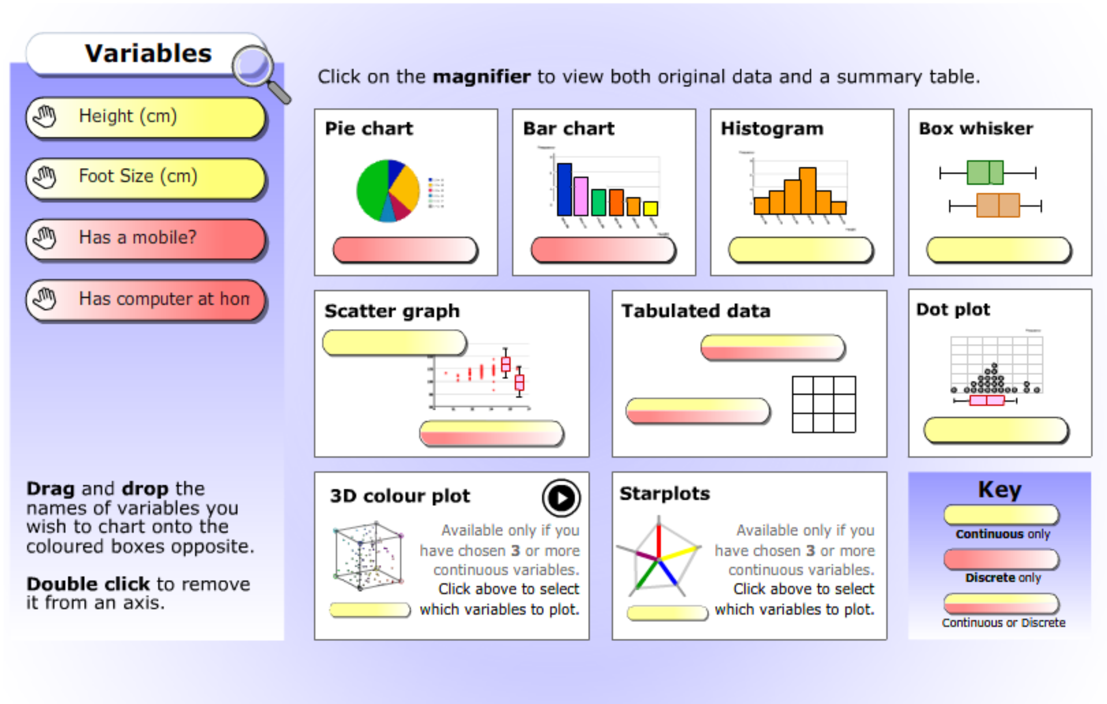


Figure 6.1. U.K. CensusAtSchool data analysis tool interface.

This inquiry used *TinkerPlots*, which is landscape-type software. The inquiry showcased *TinkerPlots* through its evaluation using the *Criteria for Evaluating EDA Software* (Table 2.2), its application as a learning tool for students, and its application as a research tool. The selection of a landscape tool for the inquiry was opportune. Providing the chance for students to work freely within the *TinkerPlots* learning environment, gave the students the freedom to construct graphs that were meaningful to them. As well as establishing the different levels of development of covariation exhibited by the students and their characteristics as computer users as they interacted with *TinkerPlots*, the inquiry also revealed anomalies in students' thinking when reasoning about graphical representations that display association. The ability of *TinkerPlots* to display bivariate data in a number of ways made it possible to see that the students applied thinking developed when reasoning about the covariation of two measurement attributes to reason about displays with one measurement attribute and one categorical attribute. At times, the students attempted to transfer their knowledge of one graph type to another and in doing so showed that their understanding about the difference between the two graph types was incomplete. The finding about how students used the information in association graphs to describe a relationship between two attributes was unanticipated and somewhat surprising. The finding was realised because *TinkerPlots* provided the means of exploring data in a number of different ways. This highlights the pivotal role that the technological environment afforded by *TinkerPlots* played in facilitating and revealing the students' statistical thinking and reasoning skills. This was possible due to the application of the *Model of Learning Behaviour* (Figure 1.7) throughout the inquiry. The inquiry also draws attention to the need for teachers to understand fully all of the representations possible from particular software not just the ones students are expected to choose according to established conventions.

Directions for Future Research

This inquiry has shown that Year 5/6 students have intuitive ideas about covariation that have the potential to be developed further. To facilitate the development of understanding of covariation, teaching strategies need to shift students' initial focus on individual points on a graph to a global focus looking at the graph to draw conclusions. This would ensure that students have a well established understanding of covariation when they encounter the application of formal statistical procedures to determine the correlation coefficient as a measure of the strength of covariation in post Year 10 statistics courses. Future research could focus on older students' understanding as they make the transition from identifying and describing covariation informally from graphs to the formal use of statistical tests for covariation.

TinkerPlots was used in this inquiry to record and display data collected by the students. Familiarity with the context of the data supported the students but it was also found to be problematic, at times. Following on from this inquiry, exploring students' understanding of covariation with data that were not so closely related to the students' knowledge of the context would be worthwhile. Choosing data from contexts less familiar to students would prevent the students from making claims that were based solely on their knowledge of the context without interrogating the data. Less familiar contexts would position the students so that they had to draw conclusions from the data and the graphical representations created as they would not have enough knowledge of the context of the data to make decisions without the data. In addition, exploring students' understanding of covariation within learning environments offered through other software packages would be of value. Future research could also focus on other ways of generating data and graphical representations. For example, the recently released version of *TinkerPlots* (Konold & Miller, 2011) includes a “Sampler,” which could be used to generate and record data. Using the Sampler to generate data would provide a transition point from using data that students can relate to on a personal level, as with data from familiar contexts, to abstract data that could be generated by *TinkerPlots*.

The data used for the Student Interviews in this inquiry were generated from the Australian CensusAtSchool data base (www.abs.gov.au). The data set was not cleaned to remove outliers and inappropriate data values before the students used it. Scatterplots generated from two of the measurement attributes hence did not have neat and tidy trends. This was intentional. The purpose was to have graphical representations that challenged students’ thinking rather than present students with neat positive trends that were obvious. Future research could use less messy data to make trends and relationships obvious with the intention of exploring students’ interpretation of the trend when the scale of the axis or axes of a scatterplot changes. One of the observations made in this inquiry was that some students relied on the visual impression of the slope of a trend to draw conclusions. They did not always check the scale of the axes to confirm or refute conclusions drawn from the visual interpretation. This could also be explored using scatterplots that displayed negative covariation.

Central to this inquiry was the *Model of Learning Behaviour in EDA Graphing Environments*. It was utilised as a theoretical framework for guiding the inquiry and as a practical research tool for analysing data. It proved to be very useful from both the theoretical and practical perspectives. Future research could focus on the application of the framework to determine if it is sufficiently general in its current form or requires adaptation to accommodate research exploring other statistical concepts. Similarly, future research could focus on the application of the *Criteria for Evaluating Exemplary EDA Software* with other statistical graphing software with the aim of confirming its utility or making recommendations for improvement.

This inquiry acknowledged the role of the researcher as teacher in the research process but did not focus on the work of the teacher and kept the role of the teacher in the background, as in other research (e.g., Bakker, 2004). One reason for this decision was that the learning environment was not a typical classroom setting. Although the learning environment was authentic in that it replicated a learning situation where a teacher may give students assistance when working at a computer, it lacked the social interaction and discourse that typifies regular classrooms. Follow-up research projects could focus on the role of teachers while implementing learning opportunities with *TinkerPlots* as part of their every day work.

Evaluation of the Inquiry

The desire to produce valuable, credible and scientifically based research was a motivating factor that underpinned this inquiry. As the inquiry draws to a close revisiting this notion is appropriate. To ensure the inquiry could be considered valid and rigorous, suggestions by the NRC (2002) were acknowledged by ensuring the inquiry was aligned with a conceptual framework, in order to ensure the research questions were investigated directly whilst taking into account the context of the inquiry. It was designed purposefully to allow direct interrogation of student thinking and reasoning about covariation while using *TinkerPlots*. The application of educational design research methodology (Akker et al., 2006) allowed the implementation of the inquiry to produce research findings that have the potential to contribute to theory building about student learning in EDA graphing environments and inform teaching practices. Educational design research was selected as it was iterative and cyclical in nature (Barab & Squire, 2004; Seeto & Herrington, 2006; Shavelson et al., 2003). It should also be recognised that a fundamental tenet of educational design research is that ongoing evaluation is an essential part of the research process (Kelly, 2006; Phillips, 2006). This inquiry has achieved that by not only including an evaluation of student learning before and after the implementation of the sequence of learning experiences but also an evaluation of *TinkerPlots* as an appropriate learning tool, and the Student Survey as an informative assessment tool. It is fitting, therefore, that the final cycle in this inquiry concludes with an evaluation of the inquiry design. The six principles of scientific inquiry offered by the NRC are used as a frame for the evaluation. The six principles are:

1. Pose significant questions that can be investigated empirically.
2. Link research to relevant theory.
3. Use methods that permit direct investigation of the question.
4. Provide a coherent and explicit chain of reasoning.
5. Replicate and generalise across studies.
6. Disclose research to encourage professional scrutiny and critique.

Implicitly bound in the quality of any research are limitations and issues of credibility and dependability. Discussion of these issues as they relate to this inquiry is incorporated into the evaluation of the inquiry according to the NRC six principles of scientific inquiry in this section.

Scientific Principle 1 - Pose significant questions that can be investigated empirically

Research questions serve as boundaries for an inquiry without constraining it (Wiersma, 1995) and help frame and refine the inquiry (Fontana & Frey, 2003). This is achieved by establishing a deep and thorough understanding of the relevant research literature (O'Donnell, 2004). The research questions posed for this inquiry were determined after a comprehensive review of the literature. Briefly, there is very little research available on the ways in which students use technology to develop an understanding of statistical concepts and more specifically, covariation. In addition, little is understood about the ways in which features of graphing software promote and enhance understanding of those concepts. Over the past fifteen years there has been extensive research conducted in the area of statistical literacy, thinking and reasoning but few studies have concentrated on the context of educational technology environments or the statistical concept of covariation (Cobb, McClain et al., 2003; Moritz, 2004; Ross & Cousins, 1993; Shaughnessy, 2007). Exploration of these issues was the main focus of this inquiry.

To attend to the gaps in the literature about the understanding of student learning, the research questions for this inquiry focused on both the theoretical and practical aspects of students' reasoning about covariation. To answer the research questions, a synergy was established between the theoretical and practical aspects of the research questions. The theory building aspect of Research Question 1, with the development of the *Model of Learning Behaviour in EDA Graphing Environments* (Figure 1.7), was used to set up an inquiry process that contributed to the practical requirements for answering Research Questions 2 and 3. In turn, the outcomes of Research Questions 2 and 3 provided the evidence that confirmed the *Model of Learning Behaviour's* utility as a theoretical/praxis framework.

Scientific Principle 2 - Link research to relevant theory

This inquiry explores the intersection of theory and praxis. Theoretical perspectives were used to guide the inquiry design, underpin methodological decisions, plan the learning sequence, and evaluate student learning. The enactment of the inquiry positioned the researcher as a teacher in the learning environment to explore student learning, which embedded the inquiry within praxis (Kemmis, 2010). Research into praxis has two main purposes: "(1) to guide the development of educational praxis, and (2) to guide the development of education itself" (p. 9). Although the inquiry provided the learning environment of *TinkerPlots*, the software did not control or constrain the inquiry findings. The researcher performed the work of a teacher and facilitated student learning, thereby giving the findings credibility (Creswell, 2005) and dependability (Merriam, 1998).

This inquiry explicitly linked the theoretical frameworks with relevant theories from the perspectives of the use of both technology and statistics. Existing research in the area of statistics education has resulted in the development of a number of models of statistical thinking and reasoning about graphs (Friel et al., 2001; Moritz, 2004; Pfannkuch & Wild, 2004; Shaughnessy, 2007). These models are extremely useful in describing students' conceptual development of statistics. They were not, however, developed with the intention of explaining learning within technological learning environments. This inquiry used those existing models to develop a framework, the *Model of Learning Behaviour in EDA Graphing Environments* (Figure 1.7), which was utilised to characterise students' thinking when using and interacting with *TinkerPlots* and develop a framework for the evaluation of graphing software, *Criteria for Evaluating Exemplary EDA Software* (Table 2.2). The integration of previous frameworks into a new Model of Learning Behaviour provided a more complete framework than could be provided by the individual frameworks.

Scientific Principle 3 - Use methods that permit direct investigation of the questions

The methods and procedures selected to investigate the research questions in this inquiry were chosen as they provided the best evidence for answering the research questions. The intention was to do justice to the complexity of the learning environment offered by *TinkerPlots* and the learning it afforded (Janesick, 2003). Such a stance reflects what Janesick (2003) calls "crystallization." She contends that crystallization deepens understanding of complex topics by exploring the same phenomena from different perspectives. This increases the internal validity of a qualitative inquiry by gathering rich, thick descriptive data as advocated by Shavelson et al. (2003). Brown (1992), however, expressed concerns over data selection as a possible limitation of educational design research. Of key importance is that the data provide the evidence needed to answer the research questions (O'Donnell, 2004). In Stage 4 of the inquiry the students used *TinkerPlots* to explore data sets and complete learning activities that gave them the opportunity to develop their understanding of statistical concepts, particularly covariation. While working at the computer the students' interaction with *TinkerPlots* was recorded as a video on the computer using the screen capture software, *Captivate* (Adobe, 2007). Although only one form of data collection was used, the complexity of that data provided multiple perspectives of the learning environment and the students' thinking and reasoning. *Captivate* provided video data that evidenced the interactions between the students and *TinkerPlots* and recorded the artifacts in the form of graphs created by the students. The audio data evidenced the students' "thinking aloud" moments and interactions between the students and the teacher. This provided multiple perspectives from which to view the learning environment and the students thinking.

Scientific Principle 4 - Provide a coherent and explicit chain of reasoning

The research methodology devised for this inquiry linked the research questions with the objectives of the inquiry (Figure 0.1) and linked the Stages of Inquiry (Figure 0.2) with the structure of this thesis. These connections provided a foundation from which a coherent and explicit chain of reasoning emerges through the stages of this thesis. This reflects the iterative and cyclical nature of the educational design methodology adopted for this inquiry. It also serves as evidence of the enactment of the inquiry design.

Scientific Principle 5 - Replicate and generalise across studies

This inquiry recognises that the qualitative approach taken to collect, analyse, interpret, and present the data may be considered a limitation. The findings from adopting a qualitative research approach may not be generalisable in the traditional sense as statistical tests cannot be conducted to determine the significance of the findings (Barab & Squire, 2004). It can, however, be argued that the results are generalisable from a logical perspective. The results from the Student Survey indicated that the students in this inquiry were a typical sample, albeit a small sample (n=12). Therefore, the findings about students' reasoning about covariation when using *TinkerPlots* gleaned from the students in this inquiry are likely to represent the reasoning of other students given the same opportunities. Transferability of the findings, however, warrants further investigation.

The results can, however, inform the way in which a larger scale research project could be developed to explore students' conceptual development of statistical concepts. It is contended that the methodology and inquiry design could be replicated. However, elements of the research may be useful on a broader scale to inform teaching praxis. The frameworks, *Model of in Learning in EDA Graphing Environments* and *Criteria for Evaluating Exemplary EDA Software* as well as the Student Survey, Sequence of Learning Experiences, and assessment protocols developed as part of this inquiry can be applied outside the domain of research praxis and provide useful artifacts for teachers to use, thus broadening the applicability of the research.

Scientific Principle 6 - Disclose research to encourage professional scrutiny and critique

Peer review and public dissemination of research results are paramount to the research process. They can provide valuable feedback on what is written as well as the how it is written. They also provide a form of quality assurance that fosters the development of scholarly research. This researcher has taken advantage of the peer review process through the writing research papers, journal articles, and presenting at faculty "in-house" seminars.

Aspects of the inquiry such as the methodology were published to a broader audience and were open to critique by professionals in the field of ICT and statistics education. The peer review process provided valuable feedback and comment on the research approach (Fitzallen & Brown, 2007), the design and evaluation of the Student Survey (Fitzallen, 2008), and the data analysis process (Fitzallen, 2011).

The development of the *Model of Learning Behaviour in EDA Graphing Environments* (Figure 1.7) that underpins this inquiry was first published in 2006 (Fitzallen, 2006). Since then, publication of its application to other aspects of the inquiry has occurred. Namely, the application of the *Model of Learning Behaviour* to develop criteria for evaluating EDA software and apply the criteria to evaluate *TinkerPlots* (Fitzallen & Brown, 2006b; Fitzallen, 2007); to develop and validate the Student Survey (Fitzallen, 2008); and to analyse the data in relation to identifying students' development of understanding of covariation (Fitzallen, 2011). In all instances, the *Model of Learning Behaviour* was not criticised. Modifications made to the original model (Fitzallen, 2006) arose from its application in the inquiry not as a consequence of the peer review process.

The following refereed conference papers and journal articles were generated from this inquiry to date.

Fitzallen, N. (2006). A model of students' statistical thinking and reasoning about graphs in an ICT environment. In P. Grootenboer, R. Zevenbergen & M. Chinnappan (Eds.), *Identities, Cultures and Learning Spaces* (Proceedings of 29th annual conference of the Mathematics Education Research Group of Australasia, Canberra, pp. 203-210). Sydney: MERGA.

Fitzallen, N. (2007). Evaluating data analysis software. *Australian Primary Mathematics Classroom, 12*(1), 23-28.

Fitzallen, N., & Brown, N. (2006b). Evaluating data-analysis software: Exploring opportunities for developing statistical thinking and reasoning. In N. Armstrong & C. Sherwood (Eds.) *IT's up here for thinking. Proceedings of the Australian Computers in Education Conference, Cairns,* October 2-4, 2006. [CDROM]

Fitzallen, N., & Brown, N. (2007). Evidence-based research in practice. In P. J. Jeffrey (Ed.), *Engaging Pedagogies. Proceedings of the Australian Association for Research in Education 2006 International Education Research Conference, Adelaide, Nov 26-30, 2006.* [CDROM]

Fitzallen, N. (2008). Validation of an assessment instrument developed for eliciting student prior learning in graphing and data analysis. In M. Goos, R. Brown, & K. Makar (Eds.),

Navigating currents and charting directions (Proceedings of the 31st annual conference of the Mathematics Education Research Group of Australasia, Brisbane, pp. 203-209). Sydney: MERGA.
Fitzallen, N. (2011). Interpreting graphs: Students developing an understanding of covariation. *Researching across the boundaries*. *Australian Association for Research in Education 2011 International Education Research Conference, Hobart, Nov 27-Dec 1, 2011*. (Accepted for publication through the peer review process but not presented.)

Concluding Remarks

This thesis documents the journey of one researcher striving to provide insights about student learning that have the potential to improve the development of understanding of covariation for students. It maintains links between theory and praxis, puts students' learning at the centre of the inquiry, and engages with an educational technology, *TinkerPlots*, that is currently changing and influencing teaching and learning environments in statistics education. In terms of students' development of understanding of covariation, the inquiry identified that young students are able to reason about covariation and display three levels of reasoning according to a SOLO hierarchy (Biggs & Collis, 1982).

The inquiry approach captured the students' interaction with *TinkerPlots* while working through a sequence of learning, which incorporated student engagement with the fundamental notions of covariation. The thesis-developed framework, *Model of Learning* (Figure 1.7), supported the development of the sequence of learning and assisted in the analysis of students' development of understanding of covariation. Most importantly, it enabled students' developing understanding of covariation to be characterised in ways that were not possible previously.

The application of *TinkerPlots* in this inquiry goes beyond current practices in relation to immersion, integration and support for knowledge acquisition within EDA software environments. Through the evaluation of *TinkerPlots* and its subsequent implementation in the inquiry, it was identified that *TinkerPlots* provides a powerful learning environment for supporting students' understanding of covariation. This confirms the premise that *TinkerPlots* was developed from an understanding of students' underlying cognitive processes (Konold, 2007).

The way in which students engaged with *TinkerPlots* was also investigated. The results suggest that students adopt three different strategies when accessing the features of *TinkerPlots* while creating and interpreting graphs. These strategies are: *Snatch and Grab*, *Proceed and Falter*, and *Explore and Complete*. The results also suggest that there is a hierarchical relationship among the three strategies but further research is required to before that assertion can be made.

Although the rate at which new educational technologies are developed out paces the rate at which effective and valid research can be conducted, continuing efforts are needed to ensure teachers are informed about how the new technologies address students' learning needs and facilitate the learning outcomes desired. Knowledge of the benefits of educational technologies will empower teachers to choose educational technologies selectively and purposefully to exploit the potential they have to improve student learning outcomes. This thesis goes some way towards supporting teachers by providing information that is directly related to students' development of understanding of covariation, the way students engage with *TinkerPlots* as part of the development process, and the influence *TinkerPlots* has on the learning experience.

References

Adobe Systems Incorporated. (2007). *Adobe Captivate 3.* [Computer software] San Jose, CA: Adobe Systems Incorporated.

Akker, J. van den, Gravemeijer, K., McKenney, S., & Nieveen, N. (Eds.). (2006). Introducing educational design research. *Educational design research* (pp. 3-7). New York: Routledge.

Alessi, S., & Trollip, S. (2001). *Multimedia for learning: Methods and development* (3rd ed.). Boston: Allyn & Bacon.

Angeli, C., & Valanides, N. (2009). Epistemological and methodological issues for the conceptualization, development, and assessment of ICT–TPCK: Advances in technological pedagogical content knowledge (TPCK). *Computers and Education, 52,* 154-168. Retrieved August 10, 2011, from http://teaching.cycu.edu.tw/pdf/2009_TPCK.pdf

Australian Curriculum, Assessment and Reporting Authority. (2011). *The Australian Curriculum – Mathematics. Version 1.2, March 8, 2011*. Retrieved May 23, 2011, from http://www.australiancurriculum.edu.au/Mathem*atics/Rationale*

Australian Education Council. (1991). *A national statement on mathematics for Australian schools.* Carlton, Vic.: Curriculum Corporation.

Australian Education Council. (1994). *Mathematics - A curriculum profile for Australian schools*. Carlton, Vic.: Curriculum Corporation.

Bakker, A. (2002). Route-type and landscape-type software for learning statistical data analysis. In B. Phillips (Ed.), *Developing a Statistically Literate Society*. *Proceedings of the Sixth International Conference on Teaching Statistics (ICOTS6, July, 2002) Cape Town, South Africa.* Voorburg, The Netherlands: International Statistical Institute. Retrieved March 10, 2010, from http://www.fi.uu.nl/publicaties/literatuur/4762.pdf

Bakker, A. (2004). *Design research in statistics education: On symbolizing and computer tools*. Utrecht: CD-β Press, Centre for Science and Mathematics Education.

Bakker, A., & Fredrickson, A. (2005). Comparing distributions and growing samples by hand and with a computer tool. In W. J. Masalski (Ed.), *Technology-supported mathematics learning environments: Sixty-seventh yearbook of the National Council of Teachers of Mathematics* (pp. 75-91). Reston, VA: National Council of Teachers of Mathematics.

Bakker, A., & Gravemeijer, K. P. E. (2004). Learning to reason about distribution. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 147-168). Dordrecht: Kluwer Academic Publishers.

Bannan-Ritland, B. (2003). The role of design in research: The integrative learning design framework. *Educational Researcher*, *32*(1), 21-25.

Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *The Journal of Learning Sciences*, *13*(1), 1-14.

Batanero, C., Estepa, A., Godino, J. D., & Green, D. R. (1996). Intuitive strategies and preconceptions about association in contingency tables. *Journal for Research in Mathematics Education, 27*(2), 151-169.

Behrens, J. T. (1997). Toward a theory and practice of using interactive graphics in statistics education. In J. B. Garfield & G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics. Proceedings of the 1996 IASE Round Table Conference University of Granada, Spain* (pp. 111-121). Voorburg, The Netherlands: International Statistical Institute.

Beniger, J. R., & Robyn, D. L. (1978). Quantitative graphics in statistics. *The American Statistician, 32*(1), 1-11.

Ben-Zvi, D. (2000). Toward understanding the role of technological tools in statistical learning. *Mathematical Thinking and Learning, 2*(1-2), 127-155.

Ben-Zvi, D. (2004). Reasoning about data analysis. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 121-146). Dordrecht: Kluwer Academic Publishers.

Ben-Zvi, D. (2006). Scaffolding students' informal inference and argumentation. In A. Rossman & B. Chance (Eds.), *Working Cooperatively in Statistics Education. Proceedings of the Seventh International Conference on Teaching Statistics*. *Salvador, Bahia: Brazil.*

Voorburg, The Netherlands: International Statistical Institute. Retrieved July, 1, 2008 from
http://www.redeabe.org.br/ICOTS7/Proceedings/index.html

Ben-Zvi, D., & Arcavi, A. (2001). Junior high school students' construction of global views of data and data representations. *Educational Studies in Mathematics, 45,* 35-65.

Ben-Zvi, D., & Friedlander, A. (1997). Statistical thinking in technological environment. In J. B. Garfield & G. Burrill (Eds.) *Research on the role of technology in teaching and learning statistics. Proceedings of the 1996 IASE Round Table Conference University of Granada, Spain* (pp. 45-55). Voorburg, The Netherlands: International Statistical Institute.

Ben-Zvi, D., Garfield, J., & Zeiffler, A. (2006). Research in the statistics classroom: Learning from teaching experiments. In G. F. Burrill & P. C. Elliot (Eds.), *Thinking and reasoning with data and chance: Sixty-eighth yearbook* (pp. 467-481). Reston VA: National Council of Teachers of Mathematics.

Bertin, J. (1981). *Graphics and graphic information-processing*. Translated by W. J. Berg. Berlin: De Gruyter.

Bertin, J. (1983). *Semiology of graphics: Diagrams, networks, maps*. Translated by W. J. Berg. Madison, WI: University of Wisconsin Press. (Original work published 1967)

Biehler, R. (1997). Software for learning and for doing statistics. *International Statistical Review, 65*(2), 167-189.

Biggs, J. B., & Collis, K. F. (1982). *Evaluating the quality of learning. The SOLO taxonomy*. New York: Academic Press.

Black, P., & Wiliam, D. (1998). Inside the black box: Raising the standards through classroom assessment. *Phi Delta Kappan, 80*(2), 139-148.

Bloom, B. S., Madaus, G. F., & Hastings, J. T. (1981). *Evaluation to improve learning*. New York, NY: McGraw-Hill.

Boaz, A., & Ashby, D. (2003). *Fit for purpose? Assessing research quality for evidence based policy and practice.* ESRC UK Centre for Evidence Based Policy and Practice: Working Paper 11. Retrieved September 3, 2011, from
http://www.kcl.ac.uk/content/1/c6/03/46/04/wp11.pdf

Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *The Journal of the Learning Sciences*, *2*(2), 141-178.

Brown, N., Watson, J., & Wright, S. (2011). My favourite ratio: An inquiry about Pi. *Australian Mathematics Teacher*, *67*(1), 14-26.

Burkhardt, H., & Schoenfeld, A. H. (2003). Improving educational research: Toward a more useful, more influential, and better-funded enterprise. *Educational Researcher*, *32*(9), 3-14.

Camden, M. (2009). Engaging learners in statistical education: Some ‘whys’ and ‘hows.’ In *Teaching and modelling in Statistics.* The proceedings of the third annual Applied Statistics Education and Research Collaboration Research Conference, December 7-8, Newcastle, Australia. Retrieved August 5, 2011, from http://www.uow.edu.au/content/groups/public/@web/@inf/@math/documents/doc/uow074434.pdf

Card, S. K., Mackinlay, J., & Shneiderman, B. (1999). *Readings in information visualization: Using vision to think.* San Francisco: Morgan Kaufmann Publishers.

Chance, B., Ben-Zvi, D., Garfield, J., & Medina, E. (2007). The role of technology in improving student learning of statistics. *Technology Innovations in Statistics Education 1*(1), Article 2. Retrieved September 10, 2011, from http://escolarship.org/uc/item/8sd2t4rr

Chick, H., Baker, M., Pham, T., & Cheng, H. (2006). Aspects of teachers’ pedagogical content knowledge for decimals. In J. Novotná, H. Moraová, M. Krátká, & N. Stehlíková (Eds.), *Proceedings 30th Conference of the International Group for the Psychology of Mathematics Education, Vol. 2,* (pp. 297-304). Prague: PME. Retrieved September 10, 2011, from http://www.mtedu.tmue.edu.tw/leeys_teaching /PME/PME30(2006)/PME30-Volume2.pdf

Chick, H., & Watson, J. (2001). Data representations and interpretation by primary school students working in groups. *Mathematics Education Research Journal*, *13*(2), 91-111.

Cleveland, W. S. (1993). *The visualisation of data.* Summit, NJ: Hobart Press.

Cobb, G. W. (2007). The introductory statistics course: A ptolemaic curriculum? *Technology Innovations in Statistics Education 1*(1), Article 1. Retrieved September 10, 2011, from http://escolarship.org/uc/item/6hb3k0nz

Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher, 32*(1), 9-13.

Cobb, P., & Gravemeijer, K. P. E. (2004). Learning to reason about distribution. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 147-168). Dordrecht: Kluwer Academic Publishers.

Cobb, P., McClain, K., & Gravemeijer, K. (2003). Learning about statistical covariation. *Cognition and Instruction, 21*(1), 1-78.

Cobb, P., & Steffe, L. P. (1983). The constructivist researcher as teacher and model builder. *Journal for Research in Mathematics Education, 14,* 83–94.

Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *The Journal of the Learning Sciences, 13*(1), 15-42.

Connor, D., Davies, N., & Holmes, P. (2006). Using real data and technology to develop statistical thinking. In G. F. Burrill & P. C. Elliot (Eds.), *Thinking and reasoning with data and chance: Sixty-eighth yearbook* (pp. 185-194). Reston VA: National Council of Teachers of Mathematics.

Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches*. Thousand Oaks, CA: Sage Publications, Inc.

Creswell, J. W. (2005). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Upper Saddle River, NJ: Pearson Education, Inc.

Curcio, F. R. (1989). *Developing graph comprehension: Elementary and middle school activities.* Reston, VA: National Council of Teachers of Mathematics.

delMas, R. C. (1997). A framework for the evaluation of software for teaching statistical concepts. In J. B. Garfield & G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics. Proceedings of the 1996 IASE Round Table Conference University of Granada, Spain (pp. 83-98).* Voorburg, The Netherlands: International Statistical Institute.

Department of Education, Tasmania. (2006). *Guidelines for conducting research in Tasmanian government schools and colleges.* Retrieved December 12, 2006, from http://www.education.tas.gov.au/dept/reports/research

Department of Education, Tasmania. (2007). *The Tasmanian curriculum: Mathematics-numeracy. K-10 syllabus and support materials.* Hobart, Tas.: Author.

Design-Based Research Collective. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher, 32*(1), 5–8.

Diaconis, P. (1983). Theories in data analysis. In D. C. Hoaglin, F. Mosteller, & J. W. Tukey (Eds.), *Understanding robust and exploratory data analysis* (pp. 1-36). New York, NY: John Wiley & Sons, Inc.

Dunkels, A. (1988). EDA in the primary classroom - Graphing and concept formation combined. In R. Davidson & J. Swift (Eds.), *Proceedings of the Second International Conference on Teaching Statistics* (pp. 61-66). Victoria, BC: University of Victoria.

Earl, L. (2003). *Assessment as learning: Using classroom assessment to maximize student learning*. Thousand Oaks, CA: Corwin Press.

Eisenhart, M., & Towne, L. (2003). Contestation and change in national policy on “scientifically based” education research. *Educational Researcher, 32*(7), 31-38.

Emerson, J. D., & Hoaglin, D. C. (1983). Stem-and-leaf displays. In D. C. Hoaglin, F. Mosteller, & J. W. Tukey (Eds.), *Understanding robust and exploratory data analysis* (pp. 7-32). New York: John Wiley & Sons, Inc.

Emerson, J. D., & Stoto, M. A. (1983). Boxplots and batch comparison. In D. C. Hoaglin, F. Mosteller, & J. W. Tukey (Eds.), *Understanding robust and exploratory data analysis* (pp. 97-125). New York: John Wiley & Sons, Inc.

Emerson, J. D., & Strenio, J. (1983). Boxplots and batch comparison. In D. C. Hoaglin, F. Mosteller, & J. W. Tukey (Eds.), *Understanding robust and exploratory data analysis* (pp. 58-96). New York: John Wiley & Sons, Inc.

Erickson, F., & Gutierrez, K. (2002). Culture, rigor, and science in educational research. *Educational Researcher*, *31*(8), 21-24.

Ernie, K. (1996). Technology reviews: Prob Sim. *Mathematics Teacher, 89*(4), 359-360.

Ferdig, R. E. (2006). Assessing technologies for teaching and learning: Understanding the importance of technological pedagogical content knowledge. *British Journal of Educational Technology, 37*(5), 749-760.

Feuer, M. (2005). *Looking at the future of education research and policy.* An interview with Burton and Inglis lecture series keynote speaker. Retrieved May 29, 2006, from http://www.gsc.harvard.edu/news/features/feuer02012005.html

Feuer, M. J., Towne, L., & Shavelson, R. J. (2002). Scientific culture and educational research. *Educational Researcher, 31*(8), 4-14.

Fienberg, S. E. (1979). Graphical methods in statistics. *The American Statistician, 33*(4), 165-178.

Finger, G., Russell, G., Jamieson-Proctor, R., & Russell, N. (2007). *Transforming learning with ICT: Making IT happen.* Frenchs Forest, NSW: Pearson Education Australia.

Fitzallen, N. (2005a). Profiling teachers' integration of ICT into professional practice. In P. J. Jeffrey (Ed.), *Doing the Public Good: Positioning Education Research. Proceedings of the Australian Association for Research in Education 2004 International Education Research Conference, Melbourne, Nov 28-Dec 2, 2004.* [CDROM]

Fitzallen, N. (2005b). Integrating ICT into professional practice: A case study of four mathematics teachers. In P. Clarkson, A. Downton, D. Gronn, M. Horne, A. McDonough, R. Pierce, & A. Roche (Eds.), *Building Connections: Research, Theory and Practice* (Proceedings of 28th annual conference of the Mathematics Education Research Group of Australasia, Melbourne, pp. 353-360). Sydney: MERGA.

Fitzallen, N. (2006). A model of students' statistical thinking and reasoning about graphs in an ICT environment. In P. Grootenboer, R. Zevenbergen & M. Chinnappan (Eds.), *Identities, Cultures and Learning Spaces* (Proceedings of 29th annual conference of the Mathematics Education Research Group of Australasia, Canberra, pp. 203-210). Sydney: MERGA.

Fitzallen, N. (2007). Evaluating data analysis software. *Australian Primary Mathematics Classroom, 12*(1), 23-28.

Fitzallen, N. (2008). Validation of an assessment instrument developed for eliciting student prior learning in graphing and data analysis. In M. Goos, R. Brown, & K. Makar (Eds.), *Navigating currents and charting directions* (Proceedings of the 31st annual conference of the Mathematics Education Research Group of Australasia, Brisbane, pp. 203-209). Sydney: MERGA.

Fitzallen, N., & Brown, N. (2006a). What profiling tells us about ICT and professional practice. In P. J. Jeffrey (Ed.), *Creative Dissent: Constructive Solutions. Proceedings of the Australian Association for Research in Education 2005 International Education Research Conference, Parramatta, Nov 27 – Dec 1, 2005.* [CDROM]

Fitzallen, N., & Brown, N. (2006b). Evaluating data-analysis software: Exploring opportunities for developing statistical thinking and reasoning. In N. Armstrong & C. Sherwood (Eds.), *IT's up here for thinking. Proceedings of the Australian Computers in Education Conference, Cairns,* October 2-4, 2006. [CDROM]

Fitzallen, N., & Brown, N. (2007). Evidence-based research in practice. In P. J. Jeffrey (Ed.), *Engaging Pedagogies. Proceedings of the Australian Association for Research in Education 2006 International Education Research Conference, Adelaide, Nov 26-30, 2006.* [CDROM]

Fitzallen, N., Brown, N., Booth, S., & Howells, K. (2008). Transforming professional practice: What profiling tells us. In P. J. Jeffrey (Ed.), *Research Impacts: Proving or Improving? Proceedings of the Australian Association for Research in Education 2007 International Education Research Conference, Fremantle, Nov 25-29, 2007.*

Fitzallen, N., & Watson, J. (2010). Developing statistical reasoning facilitated by TinkerPlots. In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8, July, 2010), Ljubljana, Slovenia.* Voorburg, The Netherlands: International Statistical Institute. Retrieved December 16, 2010, from http://www.stat.auckland.ac.nz/~iase/publications/icots8/ICOTS8_8A4_FITZALLEN.pdf

Fitzallen, N., & Watson, J. (2011). Graph creation and interpretation: Putting skills and context together. In J. Clark, B. Kissane, J. Mousley, T. Spencer, & S. Thorton (Eds.), *Mathematics: Traditions and [New] Practices.* (Proceedings of the joint conference of the Australian Association of Mathematics Teachers and Mathematics Education Research Group of Australasia, Alice Springs, pp. 253-260). Sydney: MERGA.

Fontana, A., & Frey, J. (2003). The interview: From structured questions to negotiated text. In N. K. Denzin & Y. S. Lincoln (Eds.), *Collecting and interpreting qualitative materials* (pp. 61-106). Thousand Oaks, CA: Sage Publications, Inc.

Fraivillig, J. L., Murphy, L. A., & Fuson, K. C. (1999). Advancing children's mathematical thinking in everyday mathematics classrooms. *Journal for Research in Mathematics Education, 30*(2), 148-170.

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2005). *Guidelines for assessment and instruction in Statistics Education (GAISE) Report: A Pre-K-12 Curriculum Framework.* Retrieved March 30, 2011, from http://www.amstat.org/education/gaise/

Freebody, P., & Muspratt, S. (2007). *Uses and effects of The Le@rning Federation's learning objects: An experimental and observational study.* Retrieved June 19, 2010 from http://www.thelearningfederation.edu.au/verve/_resources/tlf_report_final.pdf

Friel, S. N., Curcio, F. R., & Bright, G. W. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education, 32*, 124-158.

Friel, S. N., O'Connor, W., & Mamer, J. D. (2006). More than "meanmedianmode" and bar graphs: What's needed to have a statistical conversation? In G. F. Burrill & P. C. Elliot (Eds.), *Thinking and reasoning with data and chance: Sixty-eighth yearbook* (pp. 117-138). Reston VA: National Council of Teachers of Mathematics.

Friendly, M. (2007). A brief history of data visualization. In C. Chen, W. Hardle, & A. Unwin (Eds.), *Handbook of computational statistics: Data visualization, Vol III* (pp. 15-57). Heidelberg: Springer. Retrieved November 18, 2009, from http://www.math.yorku.ca/SCS/Gallery/milestone/

Friendly, M. (2009). *Milestones in the history of thematic cartography, statistical graphics, and data visualization.* Retrieved November 18, 2009, from http://www.math.yorku.ca/SCS/Gallery/milestone/

Funkhouser, H. G. (1937). Historical development of the graphical representations of statistical data. *Osiris, 3*, 269-404.

Garfield, J., & Ben-Zvi, D. (2004). Research on statistical literacy, reasoning, and thinking: Issues, challenges, and implications. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of* *developing statistical literacy, reasoning and thinking* (pp. 397-409). Dordrecht: Kluwer Academic Publishers.

Garman, B. (1996). Technology reviews: DataScope. *Mathematics Teacher, 89*(4), 359.

Goos, M., & Cretchley, P. (2003). Computers, multimedia, and the internet in mathematics education. In B. Perry, G. Anthony, & C. Diezmann (Eds.), *Research in mathematics education in Australasia 2000-2003* (pp. 151-174). Sydney: MERGA.

Gould, R., Kreuter, F., & Palmer, C. (2006). Towards statistical thinking: Making real data real. In A. Rossman & B. Chance (Eds.), *Working Cooperatively in Statistics Education. Proceedings of the Seventh International Conference on Teaching Statistics (ICOTS7, July, 2006), Salvador, Bahia, Brazil).* Voorburg, The Netherlands: International Statistical Institute. Retrieved July, 1, 2008 from http://www.redeabe.org.br/ICOTS7/Proceedings/index.html

Goyne, J. S., McDonough, S. K., & Padgett, D. D. (2000). Guidelines for evaluating educational software. *The Clearing House, 73*(6), 345-348.

Greenfield, P. M. (1984). A theory of the teacher in the learning activities of everyday life. *Everyday cognition: Its development in social context* (pp. 117-138). Bridgewater, NJ: Replica Books.

Greenwood, D. J., & Levin, M. (2003). Reconstructing the relationships between universities and society through action research. In N. K. Denzin & Y. S. Lincoln (Eds.), *The landscape of qualitative research: Theories and issues* (pp. 131-166). Thousand Oaks, CA: Sage Publications.

Groth (2006). Engaging students in authentic data analysis. In G. F. Burrill & P. C. Elliot (Eds.), *Thinking and reasoning with data and chance: Sixty-eighth yearbook* (pp. 41-48). Reston VA: National Council of Teachers of Mathematics.

Hall, J. (2008). Using Census At School and TinkerPlots to support Ontario elementary teachers statistics teaching and learning. In C. Batanero, G. Burrill, C. Reading, & A. Rossman (Eds.), *Joint ICMI/IASE Study: Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education. Proceedings of the ICMI Study 18 and 2008 IASE Round Table Conference.* Retrieved August 8, 2011, from http://www.ugr.es/~icmi/iase_study/Files/Topic6/T6P4_Hall.pdf

Hammerman, J., & Rubin, A. (2002). Visualizing a statistical world. *Hands On! 25*(2). Retrieved August, 11, 2005, from http://www.terc.edu.handsonissues/fo2/stats_world.html

Hammerman, J. K., & Rubin, A. (2003). *Reasoning in the presence of variability*. Draft paper presented at the Third International Research Forum on Statistical Reasoning, Thinking and Literacy. Lincoln, Nebraska, United States of America. Retrieved June 22, 2010, from http://cehs.unl.edu/srtl/pdf/hammerman_rubin.pdf

Hammerman, J. K., & Rubin, A. (2004). Strategies for managing statistical complexity with new software tools. *Statistics Education Research Journal, 3*(2), 17-41.

Handal, B., Handal, P., & Herrington, T. (2006). Evaluating online mathematics resources: A practical approach for teachers. *Australian Primary Mathematics Teacher, 11*(2), 8-14.

Hartwig, F., & Dearing, B. E. (1979). Exploratory data analysis. *Sage University paper Series on Quantitative Methods, Vol. 16*. Newbury Park, CA: Sage Publications Inc.

Harradine, A., & Konold, C. (2006). How representational medium affects the data displays students make. In A. Rossman & B. Chance (Eds.), *Working Cooperatively in Statistics Education. Proceedings of the Seventh International Conference on Teaching Statistics (ICOTS7, July, 2006), Salvador, Bahia, Brazil).* Voorburg, The Netherlands: International Statistical Institute. Retrieved July 1, 2008, from http://www.redeabe.org.br/ICOTS7/Proceedings/index.html

Heuvel-Panhuizen, M. van den (2009). *Assessment and realistic mathematics education.* Utrecht: Utrecht University. Retrieved August 8, 2011, from http://igitur-archive.library.uu.nl/dissertations/2005-0301-003023/c1.pdf

Hoadley, C. M. (2004). Methodological alignment in design-based research. *Educational Psychologist, 39*(4), 203-212.

Hyslop-Margison, E.J., & Strobel, J. (2007). Constructivism and education: Misunderstandings and pedagogical implications. *The Teacher Educator 43*(1), 72-86.

Johnson, N. L., & Kotz, S. (1972). *Continuous multivariate distributions.* New York: Wiley.

Jones, G. A., Langrall, C. W., Mooney, E. S., & Thornton, C. A. (2004). Models of development in statistical reasoning. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 97-118). Dordrecht: Kluwer Academic Publishers.

Jones, G. A., Langrall, C. W., Thornton, C. A., & Mogill, A. T. (1997). A framework for assessing and nurturing young children's thinking in probability. *Educational Studies in Mathematics, 32*(2), 101-125.

Jones, G. A., Thornton, C. A., Langrall, C. W., Mooney, E. S., Perry, B., & Putt, I. J. (2000). A framework for characterizing children's statistical thinking. *Mathematical Thinking and Learning*, *2*(4), 269-307.

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, *33*(7), 14-26.

Kalantzis, M. (2006). The Radford Address: Elements of a science of education. In P. J. Jeffrey (Ed.), *Proceedings of* the *Australian Association for Research in Education 2005 International Education Research Conference, Parramatta, Nov 27-Dec 1.* [CDROM]

Kelly, A. E. (2003). Research as design. *Educational Researcher, 32*(1), 3-5.

Kelly, A. (2006). Quality criteria for design research: Evidence and commitments. In J. van den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.), *Educational design research* (pp. 107-118). New York: Routledge.

Kemmis, S. (2010). Research for praxis: Knowing doing. *Pedagogy, Culture and Society, 18*(1), 9-27. Retrieved September 3, 2011, from http://www.tandfonline.com/doi/pdf/10.1080/14681360903556756

Key Curriculum Press. (2005). *Fathom*: *Dynamic data software.* [Computer software] Emeryville, CA: Key Curriculum Press.

Key Curriculum Press. (2007). *TinkerPlots workshop guide.* Emeryville, CA: Author.

Kidman, G. C., & Nason, R. (2000). When a visual representation is not worth a thousand words. *Technology in Mathematics Education Conference,* December 11-14. Auckland: University of Auckland and Auckland University of Technology.

Koerber, S., Sodian, B., Thoermer, C., & Nett, U. (2005). Scientific reasoning in young children: Preschoolers' ability to evaluate covariation evidence. *Swiss Journal of Psychology*, *64*(3), 141-152.

Konold, C. (2002). Teaching concepts rather than conventions. *New England Journal of Mathematics, 34*(2), 69-81.

Konold, C. (2007). Designing a data analysis tool for learners. In M. Lovett & P. Shah (Eds.), *Thinking with data* (pp. 267-291). New York: Taylor and Francis.

Konold, C., & Higgins, T. L. (2003). Reasoning about data. In J. Kilpatrick, W. D. Martin, & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 193-215). Reston, VA: National Council of Mathematics Teachers.

Konold, C., & Khalil, K. (2003). *If u can graff these numbers — 2, 15, 6 — your stat literit.* Paper presented at the Annual Meeting of the American Educational Research Association. Chicago, IL. Retrieved May 15, 2011, from http://www.causeweb.org/artist/articles/Konold_AERA.pdf

Konold, C., & Miller, C. (1994). *ProbSim: A probability simulation program* [Computer software]. Santa Barbara, CA: Intellimation Library for the Macintosh.

Konold, C., & Miller, C. D. (2005). *TinkerPlots: Dynamic data exploration.* [Computer software] Emeryville, CA: Key Curriculum Press.

Konold, C., & Miller, C. D. (2011). *TinkerPlots: Dynamic data exploration.* Version 2 [Computer software] Emeryville, CA: Key Curriculum Press.

Konold, C., Robinson, A., Khalil, K., Pollatsek, A., Well, A., Wing, R., & Mayr, S. (2002). Students' use of modal clumps to summarize data. In B. Phillips (Ed.), *Developing a Statistically Literate Society*. *Proceedings of the Sixth International Conference on Teaching Statistics (ICOTS6, July, 2002) Cape Town, South Africa.* Voorburg, The Netherlands: International Statistical Institute. Retrieved March 10, 2010, from http://srri.umass.edu/files/konold-2002sum.pdf

Kosslyn, S. M. (1989). Understanding charts and graphs. *Applied Cognitive Psychology, 3,* 185-226.

Kvale, S. (1996). *InterViews: An introduction to qualitative research interviewing*. Thousand Oaks, CA: Sage Publications, Inc.

Langrall, C., Nisbet, S., Mooney, E., & Jansem, S. (2011). The role of context expertise when comparing data. *Mathematical Thinking and Learning, 12*(1-2), 47-67.

Lave, J. (1988). *Cognition in Practice: Mind, mathematics and culture in everyday life.* Cambridge: Cambridge University Press.

Leinhardt, G., Zaslavsky, O., & Stein, M. K. (1990). Graphs, and graphing: Tasks, learning, and teaching. *Review of Educational Research, 60*(1), 1-64.

Lincoln, Y. S., & Guba, E. G. (2000). Paradigmatic controversies, contradictions, and emerging confluences. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed., pp. 163-188). Thousand Oaks, CA: Sage.

Lovett, M. C., & Greenhouse, J. B. (2000). Applying cognitive theory to statistics instruction. *The American Statistician, 54*(3), 196-206.

Mackenzie, N., & Knipe, S. (2006). Research dilemmas: Paradigms, methods and methodology. *Issues in Educational Research*, *16*(2), 193-205. Retrieved May 30, 2011, from http://www.iier.org.au/iier16/mackenzie.html

MacKinnon, D. (2009). Constructing and exploring Pascal's triangle in TinkerPlots. *Mathematics Teacher, 102*(8), 628-632.

Makar, K., & Rubin, A. (2009). A framework for thinking about informal statistical inference. *Statistics Education Research Journal, 8*(1), 82-105. Retrieved May 22, 2012 from http//:stat.auckland.ac.nz/serj

Marsh, C. (2004). *Becoming a teacher: Understanding skills and issues* (3rd ed.). Frenchs Forest, NSW: Pearson Education Australia.

McGehee, J., & Griffith, L. K. (2004). Technology enhances student learning across the curriculum. *Mathematics Teaching in Middle School, 9*(6), 344-349.

McKenney, S., Nieveen, N., & Akker, J. van den (2006). Design research from a curriculum perspective. In J. van den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.), *Educational design research* (pp. 67-90). New York: Routledge.

Merriam, S. B. (1998). *Qualitative research and case study application in education*. San Francisco: Jossey-Bass.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis*: *An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage Publications, Inc.

Ministerial Council of Education, Employment, Training, and Youth Affairs. (2005). *Pedagogy strategy: Learning in an online world*. Carlton, Vic: Curriculum Corporation.

Miriam, D., Edwards, R., & Alldred, P. (2001). Children and school-based research: 'Informed consent' or 'educated consent'? *British Educational Research Journal, 27*(3), 347-365.

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A new framework for teacher knowledge. *Teachers College Record, 108*(6), 1017–1054.

Monteiro, C., Asseker, A., Carvalho, L., & Campos, T. (2010). Student teachers developing their knowledge about data handling using TinkerPlots. In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society*. *Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8, July, 2010), Ljubljana, Slovenia*. Voorburg, The Netherlands: International Statistical Institute.

Mooney, E. S. (2002). A framework for characterizing middle school students' statistical thinking. *Mathematical Thinking and Learning*, *4*(1), 23-63.

Moore, D. S. (1997). New Pedagogy and new content: The case for statistics. *International Statistical Review, 65*(2), 123-137.

Moritz, J. B. (2003a). Interpreting a scattergraph displaying counterintuitive covariation. In L. Bragg, C. Campbell, G. Herbert, & J. Mousley (Eds.), *Mathematics Education Research: Innovation, Networking, Opportunity.* (Proceedings of the 26th annual conference of the Mathematics Education Research Group of Australasia, pp. 523-530). Melbourne, Vic: Deakin University.

Moritz, J. B. (2003b). Constructing coordinate graphs: Representing corresponding ordered values with variation in two-dimensional space. *Mathematics Education Research Journal*, *15*(3), 226-251.

Moritz, J. B. (2004). Reasoning about covariation. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 227-255). Dordrecht: Kluwer Academic Publishers.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics.* Reston, VA: Author.

National Health and Medical Research Council. (2007). *National statement on ethical conduct in human research.* Canberra: Commonwealth of Australia. Retrieved March 11, 2008, from http://www.nhmrc.gov.au/guidelines/publications/e72

National Research Council. (2002). Scientific research in education. In R. J. Shavelson & L. Towne (Eds.), *Committee on Scientific Principles for Education Research.* Washington, DC: National Academy Press. Retrieved May 20, 2006, from www.nap.edu/catalog/10236.html

Niess, M. L. (2005). Preparing teachers to teach science and mathematics with technology: Developing a technology pedagogical content knowledge. *Teaching and Teacher Education, 21*, 509-523.

O'Donnell, A. M. (2004). A commentary on design research. *Educational Psychologist, 39*(4), 255-260. Retrieved June 8, 2011, from http://dx.doi.org/10.1207/s15326985ep3904_7

O'Keefe, J. (1997). The human scatterplot. *Mathematics Teaching in the Middle School, 3*(3), 208-209.

Olson, D. R. (2004). The triumph of hope over experience in the search for "What Works": A response to Slavin. *Educational Researcher, 33*(1), 24-26.

Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage Publications Inc.

Paparistodemou, E., & Meletiou-Mavrotheris, M. (2008). Developing young students' informal inference skills in data analysis. *Statistics Education Research Journal, 7*(2), 83-106.

Pegg, J., & Tall, D. (2005). The fundamental cycle of concept construction underlying various theoretical frameworks. *International Reviews on Mathematical Education, 37*(6), *468-475.* Retrieved July 7, 2011, from http://dx.doi.org/10.1007/BF02655855

Perkins, D., & Blythe, T. (1994). Putting understanding up front. *Educational Leadership, 51*(5), 4-7.

Perso, T. (2006). Issues concerning the teaching and learning of mathematics and numeracy. *Australian Mathematics Teacher, 62*(1), 20-27.

Pfannkuch, M. (2011). The role of context in developing informal statistical inferential reasoning: A classroom study. *Mathematical Thinking and Learning, 12*(1-2), 27-46.

Pfannkuch, M., & Wild, C. (2004). Towards an understanding of statistical thinking. In D. Ben-Zvi, & J. Garfield, (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 17-46). Dordrecht, Netherlands: Kluwer Academic Publishers.

Phillips, D. C. (2006). Assessing the quality of design research proposals: Some philosophical perspective. In J. van den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.), *Educational design research* (pp. 93-99). New York: Routledge.

Playfair, W. (1805). *An inquiry into the permanent causes of the decline and fall of powerful and wealthy nations.* London: Greenland & Norris. Retrieved November 13, 2009, from http://books.google.com

Ramsden, P. (2003) *Learning to teach in higher education* (2nd ed.). London: Routledge Falmer.

Rangecroft, M. (1991a). Graphwork – Developing a progression. Part 1 – The early stages. *Teaching Statistics*, *13*(2), 44-46.

Rangecroft, M. (1991b). Graphwork – Developing a progression. Part 2 – A diversity of graphs. *Teaching Statistics*, *13*(3), 90-92.

Reading, C., & Reid, J. (2006). An emerging hierarchy of reasoning about distribution: From a variation perspective. *Statistic Education Research Journal*, *5*(2), 46–68.

Reeves, T. (2006). Design research from a technology perspective. In J. van den Akker, K. Gravemeijer, S. McKenney & N. Nieveen (Eds.), *Educational design research* (pp. 52-66). New York: Routledge.

Richardson, L. (2003). Writing: A method of inquiry. In N. K. Denzin & Y. S. Lincoln (Eds.), *Collecting and interpreting qualitative materials* (pp. 499-541). Thousand Oaks, CA: Sage Publications, Inc.

Roblyer, M. D. (2005). Educational technology research that makes a difference: Series Introduction. *Contemporary Issues in Technology and Teacher Education, 5*(2), 192-201. Retrieved February 25, 2010, from http://www.citejournal.org/articles/v5i2seminal1.pdf

Roschelle, J. (1995). *Learning in interactive environments: Prior knowledge and new experience.* Retrieved March 23, 2011, from http://www.exploratorium.edu/ifi/resources/museumeducation/priorknowledge.html

Ross, J. A., & Cousins, J. B. (1993). Patterns of student growth in reasoning about correlational problems. *Journal of Educational Psychology, 85*(1), 49-65.

Roth, W-M., Pozzer-Ardenghi, L., & Han, J. Y. (2005). *Critical graphicacy: Understanding visual representation practices in school science.* Dordrecht, Netherlands: Springer.

Rubin, A. (2002). Interactive visualization of statistical relationships: What do we gain? In B. Phillips (Ed.), *Developing a Statistically Literate Society*. *Proceedings of the Sixth International Conference on Teaching Statistics (ICOTS6, July, 2002) Cape Town, South*

Africa. Voorburg, The Netherlands: International Statistical Institute. Retrieved July 1, 2008, from http://www.stat.auckland.ac.nz/~iase/publications/1/7f6_rubi.pdf

Rubin, A. (2005). Math that matters. *Hands On! 28*(1), 3-7.

Rubin, A. (2007). Much has changed; little has changed: Revisiting the role of technology in Statistics Education 1992-2007. *Technology Innovations in Statistics Education, 1*(1), 1-33. Retrieved January 5, 2009, from http://escholarship.org/uc/item/833239sw

Rubin, A., & Hammerman, J. K. (2006). Understanding data through new software representations. In G. F. Burrill & P. C. Elliot (Eds.), *Thinking and reasoning with data and chance: Sixty-eighth yearbook* (pp. 241-256). Reston VA: National Council of Teachers of Mathematics.

Rubin, A., Hammerman, J. K. L., & Konold, C. (2006). Exploring informal inference with interactive visualisation software. In A. Rossman & B. Chance (Eds.), *Working Cooperatively in Statistics Education. Proceedings of the Seventh International Conference on Teaching Statistics, (ICOTS7, July 2-7, 2006). Salvador, Bahia: Brazil.* Voorburg, The Netherlands: International Statistical Institute. Retrieved July, 1, 2011 from http://www.stat.auckland.ac.nz/~iase/publications/17/2D3_RUBI.pdf

Ryan, G. W., & Bernard, H. R. (2003). Data management and analysis methods. In N. K. Denzin & Y. S. Lincoln (Eds.), *Collecting and interpreting qualitative materials* (pp. 259-309). Thousand Oaks, CA: Sage Publications, Inc.

Salomon, G., & Ben Zvi, D. (2006). The difficult marriage between education and technology: Is the marriage doomed? In L. Verschaffel, F. Dochy, M. Boekaerts & S. Vosniadou (Eds.), *Instructional psychology: Past, present and future trends (Essays in honor of Erik De Corte)*(pp. 209-222). Amsterdam: Elsevier.

Sandelowski, M. (2000). Focus on research methods: Whatever happened to qualitative description? *Research in Nursing & Health*, *23*, 334-340. Retrieved August 15, 2010, from http://pages.cpsc.ucalgary.ca/~sillito/cpsc-601.23/reading/sandelowski-2000.pdf

Sandoval, W. A. (2004). Developing theory by refining conjectures embodied in educational designs. *Educational Psychologists*, *39*(4), 213-223.

Scardamalia, M., & Bereiter, C. (2006). Knowledge building: Theory, pedagogy, and technology. In K. Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences* (pp. 97-118). New

York, NY: Cambridge University Press. Retrieved July 13, 2011, from http://ikit.org/fulltext/2006_KBTheory.pdf

Schwartz, D. L., Sears, D., & Chang, J. (2007). Reconsidering prior knowledge. In M. Lovett & P. Shah (Eds.), *Thinking with data* (pp. 319-344). New York: Taylor and Francis.

Seeto, D., & Herrington, J. (2006). Design-based research and the learning designer. *Proceedings of the 23rd annual ASCILITE conference: Who's learning? Whose technology?* Sydney, pp. 741-745. Retrieved September 8, 2010, from www.ascilite.org.au/conferences/sydney06/proceeding/.../p177.pdf

Shaughnessy, J. M. (2007). Research on statistics learning and reasoning. In F. K. Lester, Jr. (Ed.), *Second handbook on research on mathematics teaching and learning* (pp. 957-1009). Charlotte, NC: Information Age Publishing.

Shaughnessy, J. M., Garfield, J., & Greer, B. (1996). Data handling. In A. J. Bishop, K. Clements, C. Keitel, J. Kilpatrick, & C. Laborde (Eds.), *International handbook of mathematics education* (pp. 205-237). Dordrecht, The Netherlands: Kluwer.

Shavelson, R. J., Phillips, D. C., Towne, L., & Feuer, M. J. (2003). On the science of education design studies. *Educational Researcher, 32*(1), 25-28.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4-14.

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57*(1), 1 - 22.

Siegel, H. (2006). Epistemological diversity and education research: Much ado about nothing much. *Educational Researcher, 35*(2), 3-12.

Silverman, D. (2003). Analyzing talk and text. In N. K. Denzin & Y. S. Lincoln, *Collecting and interpreting qualitative materials* (pp. 340-362). Thousand Oaks, CA: Sage Publications, Inc.

Slavin, R. E. (2002). Evidence-based education policies: Transforming educational practice and research. *Educational Researcher, 31*(7), 15-21.

Slavin, R. E. (2005). *Show me the evidence: Effective programs for elementary and secondary schools.* Baltimore: John Hopkins University. Retrieved May 6, 2006, from www.cddre.org/Resources/ShowMeTheEvidence092705.pdf

Slavin, R. E. (2008). Perspectives on evidence-based in education. What works? Issues in sythesizing educational program evaluations. *Educational Researcher, 37*(1), 5-14. Retrieved March 13, 2011, from http://www.bestevidence.org.uk/assets/what_works_2008_ER.pdf

Slayton, J., & Llosa, L. (2005). The use of qualitative methods in large-scale evaluation: Improving the quality of the evaluation and the meaningfulness of the findings. *Teachers College Record*, *107*(12), 2543-2565.

Sloane, F. (2006). Normal and design sciences in education: Why both are necessary. In J. van den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.), *Educational design research* (pp. 19-44). New York: Routledge. Retrieved July 13, 2011, from: http://www.fi.uu.nl/publicaties/literatuur/EducationalDesignResearch.pdf

Smith, J. P. III, diSessa, A. A., & Roschelle, J. (1993-1994). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *The Journal of the Learning Sciences, 3*(2), 115-163.

Smith, A. F., & Prentice, D. A. (1993). Exploratory data analysis. In K. Gideon & C. Lewis (Eds.), *The handbook for data analysis in the behavioural sciences: Statistical issues* (pp. 349-390). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Squires, D., & McDougall, A. (1996). Software evaluation: A situated approach. *Journal of Computer Assisted Learning*, *12*, 146-161.

Squires, D., & Preece, J. (1996). Usability and learning: Evaluating the potential of educational software. *Computers and Education*, *27*(1), 15-22.

Steffe, L. P. (1991). The constructivist teaching experiment: Illustrations and implications. In E. von Glasersfeld (Ed.), *Radical constructivism in mathematics education* (pp. 177-194). Dordrecht, The Netherlands: Kluwer Academic Publishers. Retrieved August 20, 2011, from http://math.coe.uga.edu/olive/EMAT8990FYDS08/The%20constructivist%20teaching%20experiment.pdf

Steffe, L. P., & Thompson, P. W. (2000). Teaching experiment methodology: Underlying principles and essential elements. In R. Lesh & A. E. Kelly (Eds.), *Research design in mathematics and science education* (pp. 267-307). Hillsdale, NJ: Erlbaum. Retrieved August 10, 2011, from http://pat-thompson.net/PDFversions/2000TchExp.pdf

Stephan, M. L. (2009). What are you worth? *Mathematics Teaching in the Middle School*, *15*(1), 16-23.

St. Pierre, E. A. (2002). "Science" rejects postmodernism. *Educational Researcher, 31*(8), 25-27.

Tufte, E. R. (1983). *The visual display of quantitative information*. Cheshire, CT: Graphics Press.

Tufte, E. R. (1990). *Envisioning information*. Cheshire, CT: Graphics Press.

Tufte, E. R. (1997). *Visual explanations: Images and quantities, evidence and narrative.* Cheshire, CT: Graphics Press.

Tukey, J. W. (1977). *Exploratory data analysis.* Reading, MA: Addison-Wesley Publishing Company.

Tukey, J. W. (1980). We need both exploratory and confirmatory. *The American Statistician*, *34*(1), 23-25.

Van de Walle, J. (2007). *Elementary and middle school mathematics: A developmental approach* (6th ed.). Boston: Pearson Education Inc.

Velleman, P. F., & Hoaglin, D. C. (1981). *Applications, Basics, and Computing of Exploratory Data Analysis.* Retrieved December 15, 2010, from http://ecommons.library.cornell.edu/handle/1813/78

Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes.* Cambridge: Harvard University Press.

Wainer, H., & Velleman, P. F. (2001). Statistical graphics: Mapping the pathways of science. *Annual Review of Psychology*, *52*, 305-335.

Walford, G. (2001). *Doing qualitative educational research: A personal guide to the research process*. London: Continuum.

Watson, J. M. (2002). Doing research in statistics education: More than just data. In B. Phillips (Ed.), *Developing a Statistically Literate Society*. *Proceedings of the Sixth International Conference on Teaching Statistics (ICOTS6, July, 2002) Cape Town, South Africa.* Voorburg, The Netherlands: International Statistical Institute. Retrieved July 1, 2008, from http://www.stat.auckland.ac.nz/~iase/publications/1/03_wa.pdf

Watson, J. M. (2006). *Statistical literacy at school: Growth and goals*. Mahwah, NJ: Lawrence Erlbaum.

Watson, J. M. (2007). Use of technology for creating cognitive conflict in mathematics. In J. Sigafoos & V. A. Green (Eds.), *Technology and teaching* (pp. 65-74). New York: Nova Science Publishers, Inc.

Watson, J. (2008a). Eye colour and reaction time: An opportunity for critical statistical reasoning. *Australian Mathematics Teacher, 64*(3), 30-39.

Watson, J. (2008b). Exploring beginning inference with novice grade 7 students. *Statistics Education Research Journal, 7*(2), 59-82.

Watson, J. M. (2009). The development of statistical understanding at the elementary school level. *Mediterranean Journal of Mathematics Education, 8*(1), 89-109.

Watson, J., & Beswick, K. (2009). Which is bigger 250 tonnes or 17%: A tale of salt. In C. Hurst, M. Kemp, B. Kissane, L. Sparrow, & T. Spencer (Eds.), *Mathematics: It's mine* (Proceedings of the 22nd Biennial Conference of the Australian Association of Mathematics Teachers, Inc., Fremantle, pp. 175-184). Adelaide: AAMT, Inc.

Watson, J., Beswick, K., & Brown, N. (2011). *Providing the Mathematical Foundation for an Innovative Australia within Reform Based Learning Environments. ARC Final Report, LP0560543, June, 2011*. Hobart, Tas: Faculty of Education, University of Tasmania.

Watson, J., Beswick, K., Brown, N., Callingham, R., Muir, T., & Wright, S. (2011). *Digging into Australian data with TinkerPlots*. Melbourne, Vic: Objective Learning Materials.

Watson, J., & Callingham, R. (2003). Statistical literacy: A complex hierarchical construct. *Statistics Education Research Journal, 2*(2), 3-46.

Watson, J., & Donne, J. (2008). Building informal inference in grade 7. In M. Goos, R. Brown, & K. Makar (Eds.), *Navigating currents: Charting directions* (Proceedings of the 31st Annual Conference of the Mathematics Education Research Group of Australasia, Brisbane, pp. 555-562). Sydney, NSW: MERGA.

Watson, J., & Donne, J. (2009). TinkerPlots as a research tool to explore student understanding. *Technology Innovations in Statistics Education, 3*(1)1-35. Retrieved February 11, 2011, from http://escholarship.org/uc/item/8dp5t34t

Watson, J. M., & Fitzallen, N. E. (2010). *Development of graph understanding in the mathematics curriculum. Report for the NSW Department of Education and Training.* Sydney: State of New South Wales through the Department of Education and Training. Retrieved September 7, 2010, from http://www.curriculumsupport.education.nsw.gov.au/primary/mathematics/assets/pdf/dev_graph_undstdmaths.pdf

Watson, J. M., Fitzallen, N. E., Wilson, K. G., & Creed, J. F. (2008). The representational value of hats. *Mathematics Teaching in the Middle School 14*(1), 4-10.

Watson, J. M., & Kelly, B. A. (2003). Inference from a pictograph: Statistical literacy in action. In L. Bragg, C. Campbell, G. Herbert, & J. Mousley (Eds.), *Mathematics education research: Innovation, networking, opportunity* (Proceedings of the 26th Annual Conference of the Mathematics Education Research Group of Australasia, Geelong, pp. 720-727). Sydney, NSW: MERGA.

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

Watson, J., Skalicky, J., Fitzallen, N., & Wright, S. (2009). Licorice production and manufacturing: All-sorts of practical applications for statistics. *Australian Primary Mathematics Classroom, 14*(3), 4-13.

Watson, J., & Wright, S. (2008). Building informal inference with TinkerPlots in a measurement context. *Australian Mathematics Teacher*, *64*(4), 31-40.

Wiersma, W. (1995). *Research methods in education* (6th ed.). Needham Heights, MA: Allyn and Bacon.

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

Wiliam, D., & Black, P. (1996). Meanings and consequences: A basis for distinguishing formative and summative functions of assessment? *British Educational Research Journal*, *22*(5), 537-548.

Woolfolk, A., & Margetts, K. (2007). *Educational psychology*. Frenchs Forest, NSW: Pearson Education Australia.

Zhao, Y., Pugh, K., Sheldon, S., & Byers, J. L. (2002). Conditions for classroom technology innovations. *Teachers College Record*, *104*(3), 482-515.

Zieffler, A. S., & Garfield, J. B. (2009). Modeling the growth of students' covariational reasoning during an introductory statistics course. S*tatistics Education Research Journal, 8*(1), 7-31.