Examining Constructs of Statistical Variability Using Mobile Data Points

by

George L. Ekol

M.Sc. (Math.), Makerere University, 2000
B.Sc. (Math. and Physics), Makerere University, 1990

Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

in the
Mathematics Education Program
Faculty of Education

SIMON FRASER UNIVERSITY

Fall 2013

Approval

Name:	**George L. Ekol**
Degree:	**Doctor of Philosophy (Mathematics Education)**
Title of Thesis:	**Examining Constructs of Statistical Variability Using Mobile Data Points**
Examining Committee:	**Chair:** Stephen Campbell Associate Professor

Nathalie Sinclair
Senior Supervisor
Associate Professor

Rina Zazkis
Supervisor
Professor

Richard Lockhart
Supervisor
Professor
Department of Statistics & Actuarial Science

David Pimm
Internal Examiner
Professor

Robert delMas
External Examiner
Associate Professor
Department of Educational Psychology
University of Minnesota

Date Defended/Approved: December 06, 2013

Partial Copyright Licence

Ethics Statement

The author, whose name appears on the title page of this work, has obtained, for the research described in this work, either:

a. human research ethics approval from the Simon Fraser University Office of Research Ethics,

or

b. advance approval of the animal care protocol from the University Animal Care Committee of Simon Fraser University;

or has conducted the research

c. as a co-investigator, collaborator or research assistant in a research project approved in advance,

or

d. as a member of a course approved in advance for minimal risk human research, by the Office of Research Ethics.

A copy of the approval letter has been filed at the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for approval and letter of approval are filed with the relevant offices. Inquiries may be directed to those authorities.

Simon Fraser University Library
Burnaby, British Columbia, Canada

update Spring 2010

Abstract

Statistical variability is considered by researchers and educators as the very foundation of statistics and without variability, there would be no use for statistics. Research studies, however, show that while students are good at calculating the formal measures of variability such as range, interquartile range and standard deviation, many are challenged by what these measures mean. I assumed that post–secondary students' difficulties with the formal measures of variability are partly imposed by the predominantly static environments in which they learn those concepts. Thus, I designed two dynamic mathematics sketches (DMS) using *The* Geometer's *Sketchpad* and explored how first year university statistics students think about variability, focussing on their constructs of *distribution*, *mean* and *standard deviation*. Five students were clinically interviewed, firstly without using the DMS; secondly, while using the DMS in a computer-based environment; and lastly after interacting with the DMS.

I used one-on-one, task-based interviews and collected data following foundations of statistical thinking theoretical perspectives. Analysis of video transcripts and screen shots also relied on the foundations of statistical thinking, focusing on students' considerations of variability including aggregate reasoning with data; and also on semiotic mediation theoretical perspective. Data analysis revealed that before using the DMS, my participants were more likely to think about measures of variability in terms of procedures and calculations. However, during and after their interactions with the DMS, participants showed a difference in that they were more likely to link the changes in data distribution with change in standard deviation and the mean, and to discuss the functional connections in their own words. The findings seem to suggest that applying dynamic computing tools could provide students with deeper understanding of the meaning of statistical variability, and thus could help them build stronger foundation for understanding more challenging concepts in statistics and mathematics. The study also sheds light on the contributions of dynamic, physical and tactile learning tools in mediating meanings of statistical concepts. I also propose a multi-variation reasoning framework based on participants' interactions with the DMS in the computer-based environment. More contributions to, and implications for, university/college statistics research and curriculum are discussed.

Keywords: Statistical variability; dynamic mathematics sketches ; *Sketchpad*; computer-based environment ; considerations of variability; semiotic mediation; University /college statistics

To Priscilla, Gabriel, Martha, Dan and Rosemary.

Acknowledgements

First and foremost, I thank very much Dr. Nathalie Sinclair, my senior supervisor, for her immeasurable support in this research study. Her encouragement, professional advice, very careful attention to the authenticity of research information, timely feedback, and an open door throughout my doctoral program enabled me to complete this project. I also thank Dr. Sinclair for supporting me with research funding, which enabled me to attend and present my research at conferences. Her funding support not only enabled me to keep pace with research in statistics/mathematics education, but it also allowed me to join and actively participate in academic and professional communities locally and internationally.

Next, I thank my other two supervisory committee members: Dr. Rina Zazkis and Dr. Richard Lockhart for their invaluable feedback on my research. Their careful reading of my work and helpful feedback both challenged and encouraged me to dig deep and provide more insight into the study. Furthermore, I thank Dr. Lockhart, the Chair of the Department of Statistics and Actuarial Science for offering me multiple opportunities to work as a teaching assistant in the department during my doctoral program. I enjoyed my teaching work at the department and learned a lot from both students and the faculty members. I am lucky and proud of having all of you as my academic and professional advisors.

I would like to thank Simon Fraser University for offering me Graduate Fellowships which enabled me to carry out my research project without interruptions. As well, I thank all my study participants for volunteering in the study and for sharing their thoughts with me in the interviews. Thank you all for your time and ideas. I also learned a lot from friends, many of whom were former doctoral students in the same programme. Space does not allow me to mention many names, but I especially would like to thank Dr. Shiva Gol Tabaghi, Dr. Elena Halmaghi and Dr. Susan Osterle for their encouragement and generosity in sharing their research.

Lastly, special thanks go to my family for their enduring support and love. It has been a long and arduous journey, but I'm so happy to say THANK YOU to each one of you for supporting me and wishing me well.

Table of Contents

List of Tables

List of Figures

1. Introduction

This chapter has five sections: Section 1.1 provides a short history of my teaching experience, which was important in shaping my choice of research topic. In Section 1.2, I situate my study in statistical variability, followed with an example of variations in data distributions in Section 1.3. Section 1.4 provides my working definitions of the terms “static” and “dynamic” environments as used in this study. I provide an overview of the Chapters in the dissertation in Section 1.5.

1.1. The Beginnings

From 2002 to July 2008, I worked as a lecturer in the Department of Mathematics at Kyambogo University in Uganda. The Department of Mathematics offered courses for students majoring in mathematics, as well as those majoring in other subjects such as Economics, Geography, Business, Physics and Computer Science. I taught a Probability and Statistics course to students who were majoring in Computer Science. The teaching activities comprised of lecturing, weekly assignments, and tests; then the final examinations were written at the end of the semester. My students did not have the opportunity to use computers as part of their coursework. Although the Department of Computer Science had a computer laboratory, the computers did not have statistical data software installed on them; they were used for teaching Computer Science courses such as writing and testing computer programs. I had the SPSS data analysis software on my personal laptop computer, but was not able to assign work to my students given that they lacked the software in the laboratory to practice with. It was frustrating to teach students without involving them in some practical activities using concepts that they learned in class. Although the students did well in the course and the external examiners were satisfied with the standard of teaching, I felt that more work needed to be done to include more activities for students in the teaching and learning of statistics. However, I did not have a clearer suggestion than asking the University to purchase the license for data analysis software for teaching statistics. I passed my recommendation on in my course report to the head of the Department of Computer Science at the University.

In August 2008, I joined a PhD programme in mathematics education at Simon Fraser University. I also took up part-time position as a Research Assistant (RA) for a professor who was investigating the impact of dynamic technology on students' understanding of mathematical concepts. The project involved using *The Geometer's Sketchpad* (Jackiw, 1991, 1995), hereafter called *Sketchpad,* to investigate dynamic reasoning across the mathematics curriculum. We engaged university-level students with dynamic sketches and documented how they talked and gestured about the concepts targeted in the sketches. The dynamic mathematics sketches seemed to offer students new ways of thinking about abstract mathematical concepts. For example, it was motivating to observe students actively participate in the tasks, first by predicting what might happen, and then checking their predictions through interacting with the sketches. It was also motivating to see how the dynamic nature of the sketches, in addition to the particular tasks that we used, helped students develop meaningful ways of thinking about mathematical concepts they had previously only memorised, such as the sign of the product of two negative numbers. I also participated in interviewing mathematics instructors on what steps they took to solve some specific mathematical problems. Did they, for instance, use computers, or pen and paper? As would be expected, there were no straight answers as to how each individual solves a problem, given that problems vary and their solutions also vary. However, we drew some very insightful answers on using technology in general, and computer technologies in particular, to solve mathematical problems. It remained a question in my mind if students also used some of the techniques professional mathematicians applied while solving mathematical problems.

In 2010, while in the doctoral program, I was offered a teaching assistant (TA) job in the department of Statistics and Actuarial Science, and assigned to the Statistics Workshop (SW). The SW was a drop-in tutorial for students who needed help with specific concepts in statistics. My job included answering students' questions and assisting them to solve statistical problems on their own. The actual work varied from student to student, but in general, it involved discussing the concepts with the student by, for instance, reviewing the statistical principles that applied to the specific problems;

reviewing examples and some applications, as well as pointing the student to some additional information to help him or her solve the tasks as independently as possible. I also supported students in analyzing statistical data using computer software such as JMP and SPSS. It was quite motivating to work with the students and to observe them solve problems by themselves after the discussions.

However, in the course of my work, I noted that students more often asked for help with problems that required them to analyze data and provide the meaning of their results in the real context from where the data came, than with questions that asked for direct use of statistical formulas in calculations. I will provide an example of reasoning with graphical data in Section 1.3. Attending to students' questions enabled me to identify problems areas in terms of concepts that they frequently asked for help with. The problem areas included explaining the variations in distributions and making inferences from samples to the population. My interest in pursuing research study in the university level introductory statistics developed from attending to students' questions related to reasoning with data representations and interpreting the results. I assumed that the students' challenges were associated with the notions of statistical variability.

1.2. Situating the Study on Variability.

> Statistics is a general intellectual method that applies wherever data, variation, and chance appear. It is a fundamental method because data, variation, and chance are omnipresent in modern life.
>
> (Moore, 1998, p. 134)

I believe that Moore's statement cited above also says something about the capacity of "the modern life", or today's technologies, to measure variations in ways that were not possible in past decades. Quantitative information is available literally everywhere, including areas such as economics, psychology, health, industry and education. Thus, the need to evaluate different kinds of data is higher today than it was in the past. Moreover, numerical data are very often used with the goal of adding credibility to arguments (Watson, 2006), as well as in advertisements such as sales promotions, sports, and in politics. However, as Watson argues, critical evaluation of information, providing evidence-based arguments as well as critically evaluating data-based claims are skills that all students should learn as part of their statistics education. However, research studies reveal that many students view statistics concepts as challenging to learn (Garfield & Ben-Zvi, 2004). For example, it is not uncommon to hear a statement such as “I’m really not good at statistics” from students, even the ones who succeed in other mathematics courses.

From my review of related literature on the concept of variability, I found that standard deviation and the mean—concepts that are widely applied in university-level introductory statistics courses—were particularly challenging to students (delMas & Liu, 2005). Based on some of delMas and Liu’s recommendations from their study, as well as my own interest in dynamic computing technologies, I designed two sketches using *Sketchpad,* which I thought would make the functional connection among standard deviation, the mean and data distribution more obvious or visible. I choose these three concepts given that they are important to the understanding of statistical variability in a data set. I also designed related tasks in the static environment, which participants first solved before they embarked on solving similar tasks using *Sketchpad.* I wondered if the use of my sketches would help students develop a better understanding of statistical variability.

1.3. Static and Dynamic Environments

My working definition of an environment is a setting, physical and or non-physical, where a learning activity is enacted. By “static environment”, I mean a more traditional, paper-and-pencil-based classroom setting where written (including diagrams) and spoken words are the primary means of instruction. I use “dynamic environment” to refer to a computer–based setting where a student interacts directly with mathematical objects (including diagrams and models) that change continuously over time. I distinguish the static from the dynamic environment in that in the former tasks are performed without the computer technologies, whereas tasks in the latter are performed using dynamic mathematics sketches (as discussed in Hegedus & Moreno-Armella, 2009).

1.4. Explaining Variations in Distributions

Describing variations in graphs are concepts that are usually covered at the beginning of introductory statistics courses before introducing concepts such as the mean, median, mode and standard deviation to describe distributions. Following is an example (adapted from Moore, 2010), which involves basic reasoning with a graph: *A university conducted a study to investigate the mode of transport used by people entering the campus. From 7 A.M. to 10 P.M., monitors were posted at every entrance to the university on a chosen day. The monitors recorded the mode of transportation used by each person as he or she entered the campus. The following bar chart was constructed based on this information. If 1500 people entered campus on a particular day, (approximately) how many people arrived by car?*

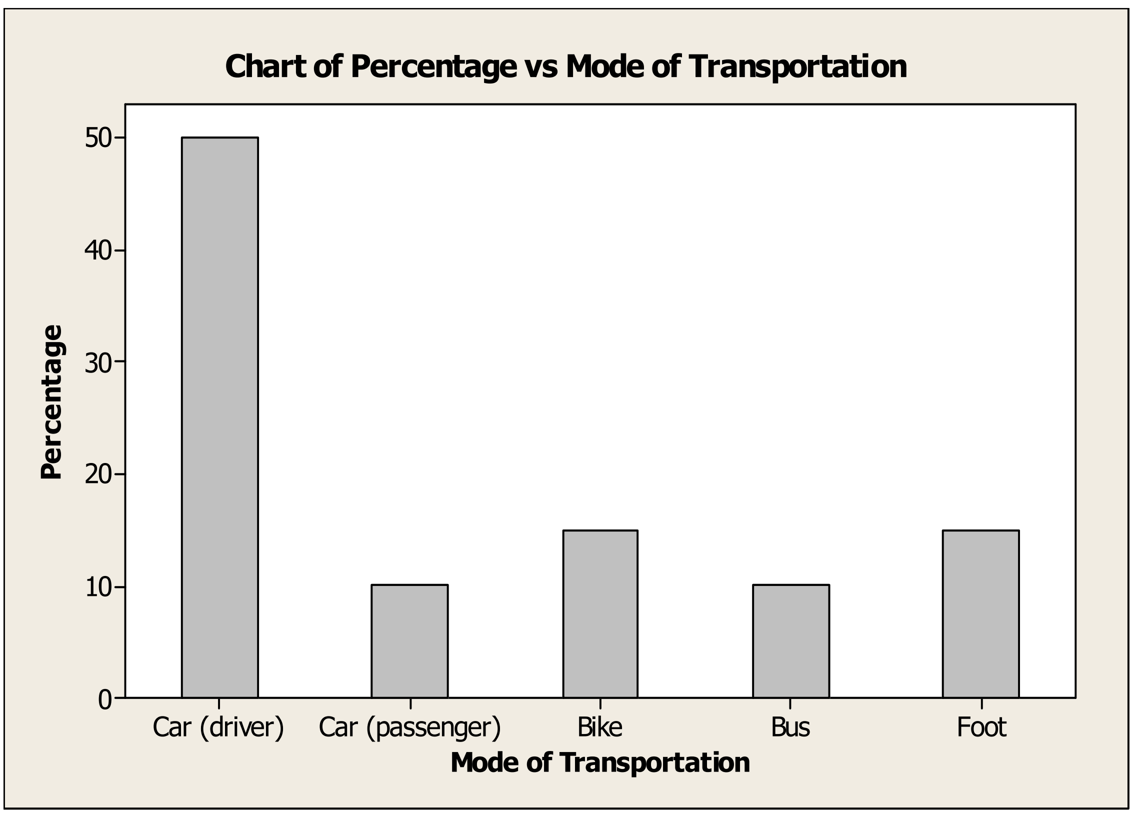


Figure 1: Bar graph showing mode of transport to a university campus (Adapted from Moore, 2010)

I choose this example to show a simple application of the notion of variability in data distribution in graphs and to point out the possible questions that students could ask for help with in solving the problem. Some students could ask why there are two separate variables for Car (driver) and Car (passenger). Other students could ask about the word "approximately", for instance, how is it possible to "approximate" a human being? Yet a few others may have problems reading off the graph and converting the percentage to the number of people who arrived by car. In general, most students in introductory statistics would probably solve this example by themselves. However, for the few who might ask for help, discussing with them the main ideas in the task, through questioning, estimating, and reasoning, would likely convince them that, approximately 60 percent of the 1500 people travelled to campus by car on that particular day. That puts the number to approximately 900 people. As for the Car (driver) being a separate variable from the Car (passenger), explanations would vary. The more likely explanation from the graph suggests that the majority of people who traveled to the campus on that particular day drove their own cars alone.

As I mentioned at the beginning of this chapter, in my Research Assistant (RA) work, I had been asked to design tasks that students could use to explore the meaning of certain mathematical concepts. I had noticed that the interactive and dynamic technology seemed to evoke students' interest in the concepts they engaged in, having them not focussing a lot on symbols or formulas as they interacted. I was seeing what Moreno-Armella, Hegedus, and Kaput (2008) have argued, which is that "the nature of mathematical symbols have evolved in recent years from static, inert inscriptions to dynamic objects or diagrams that are constructible, manipulable and interactive" (p. 103). According to them, students can use dynamic environments to reflect on their thoughts and construct new mathematical knowledge, for example, by exploring and defining properties of concepts that have been modeled graphically. The RA job also involved interviewing professional mathematicians on what steps they took to solve problems in mathematics. Working with dynamic technology motivated me to continue researching other mathematical topics using the same software. I will pursue the dynamic environments in the chapters that follow. Below is an overview of the chapters in this dissertation.

1.5. Overview of the Chapters

There are altogether seven chapters in this dissertation: Chapter 1 gives the background and my motivation to undertake a research study in statistical variability. In Chapter 2, I provide more working definitions of the terms that I frequently use in the dissertation, such as statistical reasoning, statistical thinking and statistical literacy (SRTL) as well as variation, variability, and signs. To provide some context to the study, I briefly review the historical developments in the current statistics curriculum, particularly from the 1990s to date. The evolving of the statistics curriculum document (GAISE, 2005), which is widely used in North American colleges and Universities, as well as in other parts of the world, is also reviewed. Moreover, I review research studies on the notions of distribution, the mean and standard deviation, which directly relate to my study.

In Chapter 3, I discuss two main theoretical perspectives that have influenced the current study. First, I discuss perspectives that relate to the foundations of statistical thinking, focusing on the considerations of variation (Wild & Pfannkuch, 1999; also Konold & Higgins, 2003; Reid & Reading, 2008; Carlson, Jacobs, Coe, Larsen & Hsu, 2002). Carlson et al.'s work does not come directly from a statistical research study, but it provides some insight into the notion of covariational reasoning that relates to my research. The foundations of statistical thinking provide me with a platform to discuss my findings in terms of participants' actions such as: noticing and describing variations in data distributions; measuring and explaining variation; predicting and testing the impact of variation; and looking for causes of variation in a data set. Second, I discuss perspectives on the applications of signs in mediating meanings of mathematical concepts as well signs as means of communicating with and learning from others in the same community (Vygotsky, 1978; as elaborated by Falcade, Laborde & Mariotti, 2007; and Bartolini Bussi & Mariotti, 2008). Vygotsky's socio-cultural historical perspectives relate to how tools aid one's thinking from interpersonal level, to a more personal level, through cognitive process Vygotsky calls internalization.

Having identified issues in the literature in Chapter 2 and adopted theoretical lenses to interpret my participants' expressions in Chapter 3, in Chapter 4, I discuss the tasks, procedures and materials that I use in collecting data for my study. Using

Sketchpad, and applying statistical principles, I design dynamic sketches for participants to explore the concept of variability. In particular, my study participants use the sketches to explore functional linkages among standard deviation, mean and the data distribution. I focus on standard deviation and its links with the mean and distribution of a data set.

In Chapter 5, I analyze the interview data from participants' expressions in the static and their interactions with the dynamic sketches in the computer-based environment introduced in Section 1.3. My analysis of the tasks with the dynamic sketches focuses on participants' interactions with the Dragging tool of Sketchpad and also on spoken words, during and after their interactions with the sketches. In Chapter 6 I discuss the data analysis from Chapter 5 and link the discussions to my research questions. Chapter 7 is my concluding chapter, where I respond directly to my research questions. In Chapter 7, I also discuss the limitations of the study as well as its contributions. As well, I suggest some implications for the teaching and learning of statistics at the university/college levels. In the same chapter, I also make recommendations for further study on the topic of variability.

2. The Teaching and Learning of University Level Statistics

> The cornerstone of teaching in any area is the development of a theoretical structure with which to make sense of experience, to learn from it and to transfer insights to others.
>
> (Wild & Pfannkuch, 1999, p. 224)

The review of literature in this chapter comprises of two main parts. In the first part, I review concepts that are more related to variability in statistics, whereas in the second part, I focus on the technology and related concepts. However, the dichotomy is rather artificial in that in Chapter 4, where I design interview tasks, the two parts mentioned above are not discussed separately. Section 2.1, discusses how I apply the terms statistical reasoning, statistical thinking, and statistical literacy in the dissertation. In Section 2.2, I review curriculum developments in the teaching and learning of statistics at the tertiary level from the 1990s and connect the developments with the current issues in the teaching and learning of statistics. Section 2.3 gives an overview of the current challenges in teaching and learning statistics, focusing on the challenges from students' side. In Section 2.4, I provide reviews of research studies on the concepts of distribution, mean and standard deviation. That ends the first part of the reviews. In the second part, which begins at Section 2.5, I review studies that involve computer technologies in general and dynamic computer-based technology in particular, for teaching and learning concepts in mathematics and statistics.

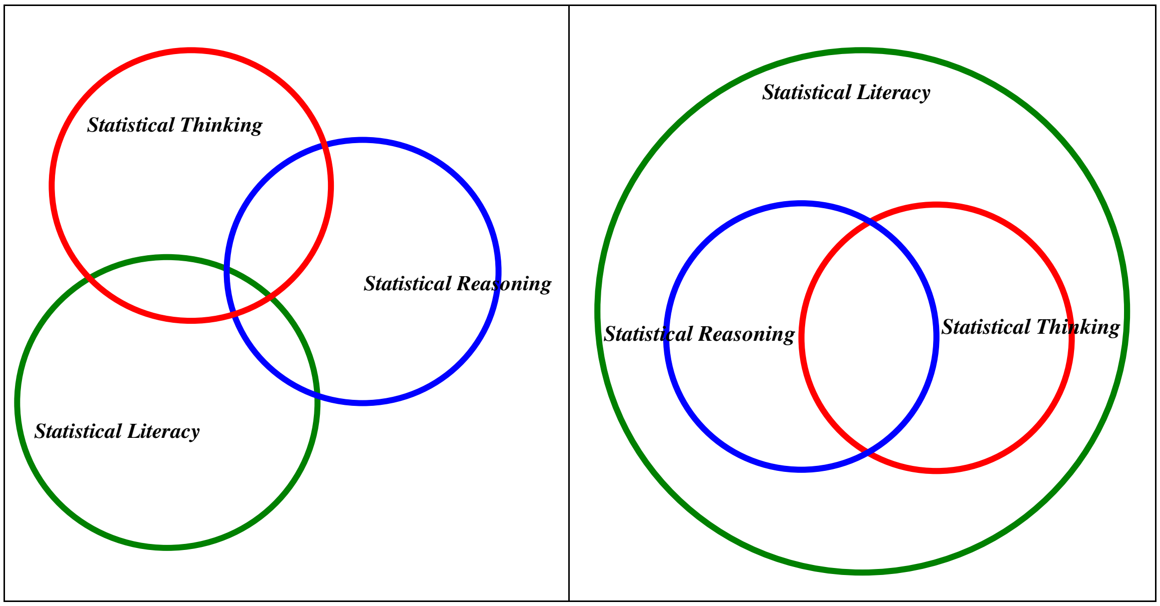
2.1. Statistical Reasoning, Thinking, Literacy (STRL)

Statistical reasoning, thinking and literacy (SRTL) are broad concepts that have attracted various theoretical perspectives from researchers. For instance, Gal (2002) considers statistical literacy as the ability to interpret, critically evaluate, and communicate statistical information. Other researchers (e.g., Garfield, 1999; Rumsey, 2002; Snell, 1999; and Garfield & Ben-Zvi, 2008) all agree that statistical literacy is about understanding and using the basic language of statistics as well as being able to interpret different representations of data. The common thread on statistical literacy among these researchers seems to be on interpreting and communicating statistical information using statistical language. Next, on statistical reasoning, Garfield & Ben-Zvi (2008) conceive it as "mental representations and connections that students have regarding statistical concepts" (p. 34). For example, reasoning about the spread of data from the centre and connecting that with magnitude of standard deviation as an index of spread. Garfield and Ben-Zvi's conception suggests a connection with Carlson et al.'s (2002) mental actions in covariational reasoning that I discuss later in this chapter. On statistical thinking, Wild and Pfannkuch (1999) propose being able to distinguish between different approaches to solving problems. For example, knowing why a particular method works better for a given task than another method. According to them, statistical thinking includes understanding the theories that inform statistical processes. Their view suggests a higher cognitive level for statistical thinking than statistical reasoning and literacy. Some models that describe the hierarchical structure of SRTL have been proposed.

delMas (2002) proposes two models of SRTL: the 'overlap model' and the 'subset model' (see Figure 2). In the overlap model, the three constructs share some common elements, but statistical thinking is considered at a higher cognitive level than reasoning and literacy. In Figure 2a, statistical literacy is modeled at the foundation, followed by statistical reasoning; and thinking is at the highest level. However, in the subset model shown in Figure 2b, statistical thinking and reasoning share some common elements, but the two are considered as elements of statistical literacy.

2.1.1. Critiquing SRTL

The models in Figure 2 show some overlap among the elements of the SRTL. However, given that reasoning, thinking, and literacy are cognitive functions, and thus cannot be directly measured, I believe that the overlap model is more suitable for describing SRTL. I take Gal's (2002) consideration of statistical literacy as the ability to interpret and critically evaluate statistical information, to relate to my study. Gal's consideration, though, is not radically different from Garfield and Ben-Zvi's (2008) position on statistical literacy as being able to use the basic language and tools of statistics to interpret different representations of data. However, Garfield and Ben-Zvi do not explicitly mention the critical evaluation of statistical information that Gal emphasizes. In this study, I would like my participants to critically evaluate statistical information and to communicate their ideas of variability in their own words.



(a) (b)

Figure 2. The overlap and hierarchical model of SRTL. (a) Literacy is modelled at the foundation and Thinking is modelled at the highest level; and (b). The subset model of SRTL where Thinking and Reasoning are subsets of Literacy (delMas, 2002).

I incorporate Garfield and Ben-Zvi's (2008) consideration of statistical reasoning as mental representations and connections that students make with statistical concepts. Moreover, I will use Wild and Pfannkuch's (1999) perspective on statistical thinking as the way mathematicians and statisticians solve tasks. According to them, statistical thinking also includes being able to understand the underlying principles behind statistical processes. Their views are consistent with the Vygotskian socio-cultural perspective of learning, which I adopt in my study.

2.1.2. Summary of the SRTL Working Definitions

Summarizing my working definitions of statistical reasoning, thinking, and literacy: I take statistical reasoning as mental representations and connections that students make with statistical concepts (Garfield and Ben-Zvi's, 2008). Statistical thinking is considered as having a mindset of statisticians/mathematicians, which includes knowing the principles behind different statistical processes. Statistical thinking also includes being able to freely choose alternative approaches to solving statistical/mathematical tasks as well as being able to communicate the solutions to others in the community (Wild & Pfannkuch). On statistical literacy, I take Gal's (2002) consideration of being able to interpret and critically evaluate statistical information.

2.2. The Reforms in the Introductory Statistics Curriculum

Before I discuss the current issues in the teaching and learning of statistics, I provide a brief historical context from the 1970s. The teaching and learning of statistics from the 1970s to date can be loosely characterized into two time periods or eras. The first era (1970-1989), that I describe as the "Kahneman era", referring to the work of Kahneman and his colleagues, comprised research studies that focused on: identifying students' misconceptions of statistics (e.g., Kahneman, Slovic & Tversky, 1982); comparing levels of students' statistical reasoning (e.g., Fong, Krantz & Nisbett, 1986); and testing models of statistical reasoning to explain why students reasoned incorrectly (e.g., Konold, 1989a,1989b; Pollatsek, Konold, Well, & Lima, 1984). Those researchers, to their credit, conducted many "diagnostic tests" (testing why students reasoned incorrectly), and set the stage for more research studies in the second era.

The second time period that I describe as the "post-Kahneman era" begins in early the 1990s. That period witnessed research studies and curriculum developments that aimed at improving the teaching and learning of statistics (e.g. Cobb, 1992, 1993; Moore, 1997; 1998; Garfield, 2000; GAISE, 2005). One of the important developments was the setting up of the curriculum working group by the Mathematical Association of America (MAA) that developed the new guidelines for teaching statistics Cobb (1992). The guidelines suggested three broad recommendations for the teaching and learning of statistics at the post-secondary level. First, it emphasized teaching students the basic elements of statistical thinking. Second, it recommended the use of data in teaching statistics, arguing that students learn better with examples that they can easily relate to in their daily lives (e.g. Figure 1). Third, the report advocated for active learning, arguing that learning should not be passive, but should involve appropriate activities, which can mediate abstract concepts for the students. Moreover, Cobb's (1992) publication also suggested specific elements that should be considered in each of the three broad recommendations mentioned above, such as paying attention to the methods of collecting reliable data, and being aware of the presence of variability in the data.

The recommendations from the working group chaired by Cobb were unanimously endorsed by the American Statistical Association (ASA) and expanded into a curriculum document called, *Guidelines for Assessment and Instruction in Statistics Education* (GAISE, 2005), which document is widely used in many post–secondary institutions in North America and other parts of the world. The GAISE (2005) curriculum recommends teaching approaches that support building conceptual understanding of students rather than teaching based on memorisation and calculations for their own sake. The curriculum also encourages students to use technology for learning concepts and for analyzing data, aiming to free students from calculations so that they can focus more on interpreting and reasoning with data output.

2.3. Challenges in the Teaching and Learning of Statistics

In spite of the achievements made in the curriculum developments in statistics, research studies reveal some challenges with respect to how students learn the concepts (see, for example, Ben-Zvi & Garfield, 2004). According to Ben-Zvi and Garfield, many students consider statistical concepts as too abstract to understand and sometimes counterintuitive. Thus, instructors may find it frustrating to motivate such students and have them actively participate in learning statistical concepts. As the example in Figure 1 attempts to show, statistical principles require some minimum level of mathematical knowledge (e.g., basic knowledge of fractions, proportions and decimals) to aid students' work with the analysis of data. However, as Ben-Zvi and Garfield suggest, students who have difficulty with basic mathematics are also more likely to have challenges learning statistical content. But university-level statistics courses require students to be able to explain their solutions in clear statistical words. However, Ben-Zvi and Garfield claim many students have challenges reporting their solutions in statistically acceptable words. Ben-Zvi and Garfield's claim suggest that students need more practice with communicating statistical information as they develop formal meanings of statistical concepts. In the next section, I review research studies on how student think about the notions of distribution, mean and variability, which are important in developing students' statistical reasoning and thinking.

2.4. The 'Big Ideas': Distribution, Centre, and Variability

The International Statistical Reasoning, Thinking, and Literacy (SRTL) research forum is a network of researchers who are interested in studying the development of students' statistical reasoning, thinking and literacy. The current studies have focused on some key ideas in foundation statistics, the so-called the "big ideas" of a discipline (Bransford, Brown & Cocking, 2000; Wiggins & McTighe, 1998). Three of the big ideas that relate to the current study are distribution, mean, and variability (or spread). I review research studies on each idea.

2.4.1. Research Studies on Distribution

Studies on how students solve statistical problems reveal that they tend to consider data sets more in terms of individual values rather thinking about data as a whole (e.g., Hancock, Kaput, & Goldsmith, 1992; Konold & Higgins, 2003). Konold and Higgins (2003) suggest that students should be engaged in learning activities that help them consider a distribution of data as a whole before they focus on unique cases in a distribution,

> If the data values students are considering vary […] why should they not […] think about those values as a whole? Furthermore, the answers to many of the questions that interested students for instance, "Who is tallest? Who has the most? […]—require locating individuals […] within the group (p. 203).

It stands to reason that if learners are not given tasks that ask them to think about distributions in aggregate, then they pay attention to the unique cases, such as the largest or the smallest values in a data set. In fact in 1997, Konold, Pollatsek, Well, and Gagnon conducted a research study with two pairs of high school students who had just completed a year-long course in probability and statistics. Konold et al. applied software to a large data set and asked participants in the study to explore and respond to different questions about the data set. Participants were also asked to support their answers with data summaries and graphs. The researchers reported that instead of considering data as a whole, their participants focused on comparing individual cases in each group.

In a related study, Konold and Higgins (2003, cited in Garfield & Ben-Zvi, 2008) propose four different ways that students think about data: i) data as *pointers*—seeing data as pointers to some events that do not relate to distribution; ii) *data as cases*—focusing on the identity of individual cases in the data set (e.g., small values or large values in the data set); iii) *data as a classifier*—paying attention to the frequency of particular data values (e.g., how many data points in the data set show large values compared to the total number of data values present); and iv) *data as aggregate*—focusing on the overall and the emergent characteristics of the data set as a whole (e.g., considering data distribution in terms of its shape, centre and spread of data values from the centre).

Other studies have focused on how students view shape, centre and spread of data as characteristics of a distribution. For example, Zawojewski and Shaughnessy (2000) analyzed data on students' responses on the National Assessment of Educational Progress (NAEP) materials spanning 15 years. Their finding suggested that students were challenged by the concept of the mean and its links to the data distribution. The authors speculated that students may have had difficulties connecting between the notions of the centre and spread of the distributions. Distribution seems so central to understanding other concepts that, for example, Mokros and Russell (1995) suggest that the concept of distribution should be taught before students are introduced to the mean, and other related notions.

In general, research studies suggest that students think about graphs of data distributions more as representations of facts rather than as tools for reasoning with, and learning something from the data set (Wild & Pfannkuch, 1999; Konold & Pollatsek, 2002). Thus, some researchers agree on teaching approaches that could help students make sense of data, through for example, detecting and discovering patterns in data, generating and testing hypotheses, noticing the unexpected (Pfannkuch, 2005; Watson, 2005); "unlocking the stories in the data" (Garfield & Ben-Zvi*,* 2008, p.171); and developing skills and abilities that can be used in interpreting statistical information (Bakker, 2004; Bakker & Gravemeijer, 2004; Garfield, and Ooms, 2005).

Summarizing my reviews on distribution, the main findings include Konold, and Higgins's categories of how students think about data distribution as *pointers*, *cases*, *classifiers*, and *aggregate*. According to them, students should be helped to consider distributions in aggregate. Overall, research studies show that students look at graphs as representations of factual information rather than as tools for reasoning about the data. It may be that the static graphs do not evoke the students' imaginations enough for them to consider graphs as reasoning tools.

2.4.2. Research Studies on the Mean

Research studies about the mean include, for example, Pollatsek, Lima and Well's (1981) study that asked college students to calculate the weighted mean of a data set. The researchers expected that the mathematically competent college students in their study could easily compute the mean of a group of two sets of numbers. Pollatsek et al. report that most students in their study answered the question without considering that the two means were taken from two groups of different sample sizes and scores. The authors conclude that for many students, dealing with the mean was more of calculation than conceptual act. Pollatsek et al. further contend that "computational rules not only do not imply any real understanding of the basic underlying concept, but may actually inhibit the acquisition of more adequate (relational) understanding" (p. 202). I add that designing learning tasks is as important as the results obtained from the tasks. Thus, if the learning goal is to move students away from focusing on calculations to learning the concepts, then the tasks should be designed so that calculations are not foregrounded.

In a different study, Hardiman, Well and Pollatsek (1984) tested whether working on tasks using a balance model would promote students' deeper understanding of the mean. A balance model depicts data values placed on a balance beam at given distances (deviations) from the mean such that the resultant 'weights' on either side of the fulcrum are at equilibrium (Hardiman, Well & Pollatsek, 1984; Strauss, 1987). Forty-eight college students participated in the study that included the pre-test, the training sessions, and the post-test of paper and pencil items. The researchers report that students who trained on the balance model performed significantly better on the post-test problems than those who did not. The findings suggest that the balance model may have contributed to the students' conceptual understanding of the mean as a point of equilibrium in a data distribution that changes as distribution of data points on the beam is changed.

In summary, Hardiman et al.'s (1984) balance model of the mean was a major contribution to understanding the concept of the mean. Their model that still applies today contrasts with Pollatsek et al.*'s* (1981) results based on calculations. Hardiman et al.'s findings suggest that using suitable models of concepts, be they physical and tactile models, could help students understand abstract ideas much better.

2.4.3. Research Studies on Variability

> If the concept of […] variation is puzzling even to statisticians and researchers, how much more puzzling must it be to those just embarking on their data handling careers?
>
> (Reading & Shaughnessy, 2004, p. 202)

As the Reading and Shaughnessy citation above suggests, the terms variation and variability have different meanings. In statistics education research, for instance, some researchers have applied variation and variability independently (e.g., Reading & Shaughnessy, 2004), while other researchers use the same terms interchangeably (e.g., Makar & Confrey, 2005; Garfield and Ben–Zvi, 2008). Variation is closely linked to the concepts of a variable (something that is liable to change) and to uncertainty (something unpredictable). Maker and Confrey (2005) define variation as the "quality of an entity to vary, including variation due to uncertainty" (p. 28). In this study, I use the terms variation and variability interchangeably, focusing on how students think about of standard deviation, mean, and distribution of a data set.

Educators and researchers agree that variability is an important concept in statistics (Rossman, 1996; Garfield & Ben-Zvi, 2008). We know from research studies that students struggle with the concept of variability although they can quite easily calculate indicators of variability such as the range, interquartile range and standard deviation in a data set. Garfield and Ben–Zvi (2008) explain:

> While students can learn how to compute formal measures of variability, they rarely understand what these summary statistics represent, either numerically or graphically, and do not understand their importance and connection to other statistical concepts. (p. 205)

Several other research studies have examined different aspects of variability (Meletiou & Lee, 2002; Lann & Falk, 2003; delMas & Liu, 2005; Reading & Reid, 2005, 2006; Slauson, 2008; Reid & Reading, 2008). Many of these studies were based in the classrooms where the researchers were also the teachers. Meletiou and Lee (2002) designed an experimental classroom study that followed a non-traditional approach. The researchers assumed that students' success in statistics depended on helping them gain sound awareness of variation. The authors engaged students on tasks with variation as the main principle. They report that the results of the assessment at the end of the course were encouraging in the sense that the majority of the students had a good grasp of the meaning and use of standard deviation. For example, the students were able to explain "[…] that one calculates standard deviation to obtain information about the distribution between the scores […] outside the centre" (p. 30). The authors propose that teachings that use realistic examples are more likely to connect to students' experiences and to help them improve their statistical reasoning than those which do not. Although Meletiou and Lee's study involved a relatively small number of students, which could restrict generalizing their study findings, their study nevertheless, was helpful in offering a methodology that other researchers can build on.

In another study on variability, Lann and Falk (2003) asked first-year university students to consider variability in a data set by incorporating measures such as the standard deviation, mean, and deviations from the mean in their considerations. Lann and Falk report that a large proportion of students chose the range instead of other measures of variability such as the standard deviation. The author's findings suggest that students were challenged differentiating among measures of variability. May be, as Mokros and Russell (1995) suggest, the concept of distribution was not well grounded to support the students' reasoning about the notion of variability. Moreover, in a different study, Slauson (2008) investigated college students' conceptual understanding of variability focusing on the standard deviation and standard error. She taught two sections of introductory statistics classes over one semester using two different methods. One section was taught using the formal lecture method covering topics on standard deviation, sampling distributions and standard error. The second section completed a hands-on, active learning laboratory covering the same topics as the first group. She collected data from the two sections and analyzed them. Her analyses revealed that:

> Students' conceptual understanding of ideas related to standard deviation improved in the active class, but not in the lecture class […]. The analysis of the qualitative data suggests that understanding the connection between data distributions and measures of variability is very important for students to successfully understand standard error. (p. iii)

Slauson's findings agree with Meletiou and Lee as well as Lann and Falk's findings, suggesting in general, that using hands-on and realistic examples were more likely to connect to students' daily experiences and to support them in their statistical reasoning (Cobb, 1992). Also important is students' awareness of the connections among data distributions and measures of variability that, in particular, Slauson's study mentions.

Furthermore, Read and Reading (2008) designed a sequence of teaching strategies with activities (e.g., short quizzes, assignment, and class tests), that aimed at assessing the understanding of variation by university-level students. Reid and Reading analyzed students' responses to test items, and from their findings, the authors developed four levels of students' consideration of variability (i.e., no consideration, weak, developing, and strong considerations of variation), each with a set of descriptors. For example, the hierarchy with no consideration of variation had two descriptors—i.e., students not being able to display any meaningful consideration of variation in the context of the tasks given; and they do not acknowledge variation in relation to other concepts. I will revisit Reid and Reading's consideration of variation hierarchies in Chapter 3. So far, I have discussed classroom-based studies on variability that have attempted to transform existing traditional methods of teaching statistics and make them more active in the sense of Cobb's (1992) recommendations and the new curriculum guidelines (GAISE, 2005).

I now review one more study on variability conducted by delMas and Liu (2005), which involves more active use of technology in designing learning tasks. delMas and Liu's study is particularly important to the current study in that I respond to some of their recommendations. The authors explored tertiary students' ability to coordinate how the mean varies with standard deviation and how such a connection can be used to measure the variability of distributions of data sets. delMas and Liu developed a dynamic conception of standard deviation that included a visual (graphical) understanding of distribution as data points on the number line at different distances from the mean. The authors assumed that the dynamic coordination of the mean with distance from the mean provided a more concrete, physical and interactive tool for mediating the abstract concept of standard deviation.

delMas and Liu further designed computer-generated graphs (e.g. Figure 3) to promote students' understanding of standard deviation. For example, in Figure 3a, the tallest bar has eight data points; each data point having a numerical value '1' marked on the horizontal number line. The location of the mean point is shown by an arrow along the number line whereas the arithmetical value of the mean is shown below the arrow. However, for standard deviation (not shown in Figure 3), both the horizontal bar and the arithmetic value appear below the mean.

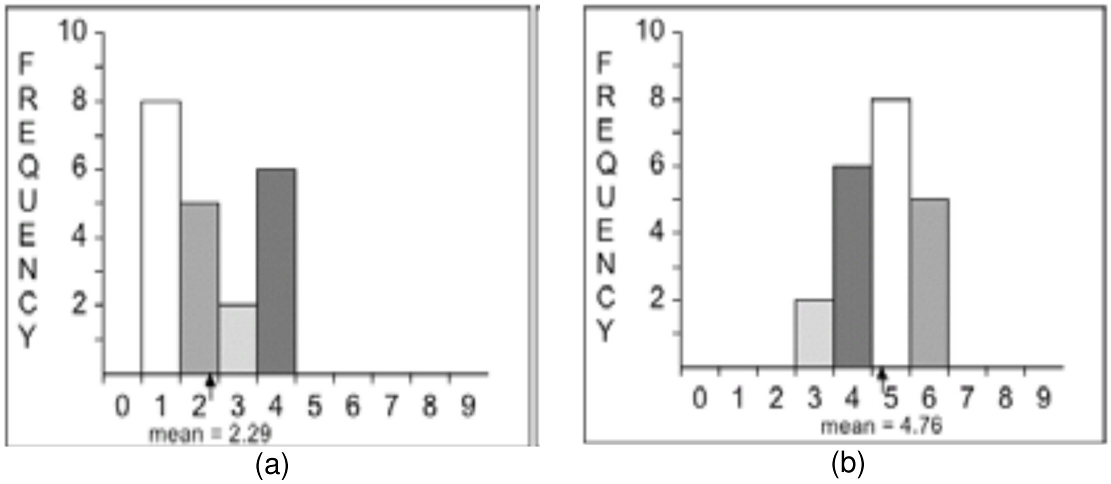


Figure 3. Comparing standard deviations in pair of graphs (delMas & Liu, 2005).

The graphs were designed so that for each pair of histograms, such as those shown in Figure 3, the numerical values for both the mean and the standard deviation were displayed for the graph in Figure 3(a), but, for the graph in Figure 3(b), only the mean value was shown (e.g., mean=4.76). Students were then asked to predict and justify their predictions, whether the standard deviation in Figure 3(b), would be larger, smaller or the same as that shown in Figure 3(a).

delMas and Liu report that although their tasks created some awareness by moving students away from initially focusing on the mean value, to considering the variability of data values about the mean, most students used a rule-based approach to compare variability across distributions instead of reasoning from conceptual representation of the standard deviation. By “a rule-based approach”, I suggest the researchers meant students focusing their comparison on individual characteristics between the two distributions, such as the tallest or shortest bars, rather than reasoning about the deviations of the bars from the centre. delMas and Liu’s findings appear consistent with Slauson’s (2008) findings, which is that understanding the connections between data distributions and measures of variability is very important for students’ learning of statistics. From their findings, delMas and Liu made some recommendations for future considerations. For example, they recommended more studies that would help focus on factors that affect the standard deviation in a given data distribution. They state:

> Is there a way to modify the task so there is less emphasis on a correct solution and more emphasis on exploring the relationships among the factors that affect the standard deviation? One possibility is to modify the software so that the mean and standard deviation are not automatically revealed. This might promote reflection. (p. 80)

My study builds on, and extends, delMas and Liu's recommendations in that it proposes tasks with emphasis on students "exploring the relationships among the factors that affect the standard deviation" (p. 80), the mean and distribution of data set. However, I adopt different theoretical and methodological approaches than those used by delMas and Liu (see Chapters 3 & 4).

Summarizing the articles in the first part of the chapter, although there is a general agreement that variability is a very important concept in statistics, the concept is not well understood by many students. Research studies reviewed in this chapter have shown that students tend to focus on comparing individual values in the data set rather than consider data in aggregate. Slauson's (2008) study, supported by findings from other studies, however, strongly suggests that understanding the connections between data distributions and measures of variability is very important in helping students develop a deeper meaning of variability. delMas and Liu (2005) proposed a conceptual model that provides the connections between the mean with distance from the mean, offering a more concrete, physical and interactive tool for mediating the abstract concept of standard deviation. Details of my study design, which uses dynamic sketches for connecting data distributions and measures of variability, are discussed in Chapter 4. In the second part of the current chapter, I focus on technology for teaching and learning statistics.

2.5. Technology for the Teaching and Learning of Statistics

This section provides an overview of the technologies that are used in statistical data analysis as well as those that are used for teaching and learning of concepts in mathematics and statistics. After a short introduction of the former, I will focus my discussion on the latter.

2.5.1. Data Analysis Software Packages

Statistical software packages are in general developed for analyzing statistical data. Examples of such software packages include SAS, Minitab, JMP, S-Plus, SPSS and R. Many of these packages, except for the R software, are menu-driven, in that they do not require detailed programming knowledge to be able to use them. However, most data analysis software are "black boxes" in the sense that the user does not see how the actual execution of the task given that it is performed inside the computer. Rather, the software provides the output (final product) after the user has entered in the data (the raw material) and instructed the computer to carry out the analysis. The process is however, slightly different for educational software packages.

2.5.2. Educational Software Packages

Educational software packages are those that have been developed to help students learn concepts in statistics or mathematics. Examples include *Fathom*, *Cabri*, *Logo* and *Sketchpad. Tinker Plots* and *Mini tools* have been used mostly in the elementary schools to help younger students handle statistical concepts using data (see Konold & Miller, 2005; Ben-Zvi, 2006; Bakker & Gravemeijer, 2004). *Fathom*, like *Tinker Plots* also offers a dynamic environment for teaching data analysis and statistics, using special features such as dragging, visualizing and simulating concepts (Ericson, 2002). I will, however, focus my discussion on the *Sketchpad* and will justify my choice in Chapter 4 in the design of the sketches for my study.

2.5.3. Dynamic Geometry for Teaching and Learning Mathematical Concepts

> Geometry is not merely an attractive side dish in a balanced mathematical diet, but an essential part of the entrée. (Goldenberg, Cuoco & Mark, 1998, p.3)

The main feature that sets dynamic geometry software (including *Sketchpad)* apart from other geometry software programs (such as *Turtle Geometry* or the *Geometric Supposer*) is its real-time transformation called “dragging.” The dragging feature makes it possible for users, after constructing a geometric figure, to freely move parts of the figure and observe how the parts respond dynamically and continuously to the movement. As the parts are moved smoothly over the domain where they exist (e.g., on the horizontal number line), some mathematical relations defined in the original figure are preserved as will be evident in the participants' activities with the sketches in Chapters 5.

Thus, the dynamic properties of mathematical objects can be used for developing mathematical concepts (Goldenberg, Cuoco & Mark, 1998; Moreno-Armella, Hegedus & Kaput, 2008). These researchers agree that students' experiences with dynamic geometry are based on their observing and interpreting the signs produced in the dynamic environment (Goldenberg et al., 1998). In fact, Moreno-Armella et al. further suggest that a dynamic environment can enable the user to explore mathematical properties through dragging using the mouse pointer and the Dragging tool of *Sketchpad.* According to them, dragging also promotes exploring concepts by for instance, questioning, "What does this surface look like when I drag and move the object?" (p. 103). I suggest that the dragging action proposed by the above authors generate signs which, according to Vygotsky (1978), can mediate meanings of mathematical concepts for students.

2.5.4. Pierce and Vygotsky's Perspectives on Signs

Pierce defines a sign "as anything which is so determined by something else, called its object, and so determines an effect upon a person" (Atkin, 2013, para. 1). Pierce's perspective suggests that the "object" of a sign could be associated with the meaning the user derives from the sign. I believe Peirce's definition is similar to Vygotsky's (1978) mediating perspective of sign in that both researchers suggest an indirect role of sign, for example, in the helping students construct meanings of mathematical concepts.

However, the Vygotskian perspective is specific on the mediation role of signs as well as the sociocultural and historical perspectives of learning such as the role language in cognitive developments. In terms of the former, a square, for instance, can function as a sign for the size of a data point's variable. Further, when such a data point is dragged along a line in the *Sketchpad* environment, the area of the square may increase or decrease in size. In that case, the sign is the changing size of the square, which is related to the mathematical meaning of variance. Vygotsky emphasises the mediating role of signs in helping learners construct meanings for concepts.

The second related aspect of sign from a Vygotskian perspective concerns the sociocultural applications of sign. Vygotsky assumes that speech, for example, is a sign that can be used to socialize. He elaborates that one may use interpersonal communication to ask for help from a member of the community who is more familiar with a given task. Vygotsky argues that, after mastering the task, one's social speech tends to turn inward so that instead of asking for help from someone else, one appeals to himself or herself for the solution. For Vygotsky, language (as a sign), which initially served an interpersonal function, now takes on an intrapersonal function that leads to internalization. He proposes that the process of internalization of social speech is linked to one's practical intellect (Vygotsky, 1978). In my study, I analyze both the sign generated due to participants' actions with physical cultural tools, and also their speech during actions with the tool.

Although the Piercian and the Vygotskian perspectives both extend the use of sign to cognitive functions, I believe Vygotsky stresses the socio-cultural and historical perspectives of cognition development more than Pierce does. Moreover, Vygotsky is quite explicit on the process of internalization, which he develops from the interpersonal (socialized) functions of signs to the intrapersonal (personalized) functions. However, by "anything which is so determined by something else, called its object", Pierce's perspective seems to suggest a more direct role of sign in the user's mind, while Vygotsky's emphasises the mediation role of sign, leading to personal awareness. In my discussions, I adopt the Vygotskian perspective of sign given that it suits the context of my study of using sociocultural tools for learning concepts.

2.5.5. The Application of the Dynamic Tools as Instruments for Learning Concepts

This last section of the chapter discusses a specific example (see Falcade, Laborde & Mariotti, 2007) that involved use of dynamic technology. The authors' assumptions are based on the Vygotskian perspectives, which I also adopt in the current study. Falcade et al. designed a teaching experiment and applied the Trace tool of the *Cabri* dynamic software for constructing the meaning of a mathematical function. The teaching experiment comprised of laboratory activities, individual student's reflections through notes-writing after the activities, and whole-class discussions facilitated by the teacher. Falcade et al. analyzed the teaching experiment based theoretical assumptions, summarized below:

> (i) The idea of variation (or co-variation) implied a relation between two changing entities, one depending on the change in other (similar to Carlson et al. 2002).
>
> (ii) Motion—defined as change in space according to change in time—is used as a metaphor for co-variation.
>
> (iii) The dynamic geometry environment (DGE) provided a semantic domain (space and time) within which variation was experienced as motion.
>
> (iv) The particular tools and objects that students interacted with in the DGE were considered as signs, referring to the notion of the function as co-variation. These (psychological) tools are considered as instruments of semiotic mediation (Vygotsky, 1978).

Assumption iv) suggests that the Trace tool of the Cabri software can be considered as a psychological tool and as such can be taken as an instrument of semiotic mediation (Vygotsky, 1978). The transformation of a physical tool into the psychological tool (instrument of semiotic mediation) is a process Vygotsky calls internalization. Vygotsky describes internalization as an "internal reconstruction of an external operation" (p. 56). From their study, Falcade et al., report some evidence of internalization when their students interacted with the Dragging and Trace tools to solve the tasks:

> The episode […] gives some evidence of the complex process of internalization through which the Dragging and Trace tool are transformed into psychological tools. […]. Furthermore, […] collective [whole class] discussion shows a particular way in which the Trace tool can function as a semiotic mediator. (p. 331).

The authors further report that their students reasoned about the meaning of variability in terms of motion, while the idea of co-variation was incorporated in the coordinated movement of points on the computer screen. The authors' notion of co-variation is consistent with Carlson et al.'s (2002) perspective of covariational reasoning that connects movements of one point on the computer screen with movements of another point on the screen. Similar research studies report, for example, how motion can provide both the physical and psychological grounding that could enable students to construct mathematical meanings (Healy & Sinclair, 2007); the differences between static and dynamic mathematical representations focusing on the teacher's expressions (Sinclair & Yurita, 2008); and the evolution of signs, showing more personal meanings of the learners shaping into mathematical meanings of concepts (Bartolini Bussi & Mariotti, 2008).

2.6. Summary

In this chapter, I provided fairly extensive reviews of research studies on the teaching and learning of introductory statistics from 1970s, but focused more on the periods from the 1990s to date. From the early 1990s to date, there have been noticeable shifts of focus from identifying students' learning difficulties, to designing strategies and implementing them with the goal of building students' statistical reasoning, thinking, and literacy (STRL). More recent research studies (e.g., Slauson, 2008), suggest that the connections between data distributions and measures of variability, such as the standard deviation and the mean, are important factors in enabling students to understand other concepts in statistics. I also reviewed studies on the application of signs, produced through actions with dynamic computing technologies and found some evidence that signs can mediate meanings of challenging mathematical concepts for students. The next chapter discusses the theoretical perspectives that informed the current study.

3. Theoretical Perspectives

In this chapter, I present two main theoretical perspectives that have influenced my study. The first perspective relates to the foundations of statistical thinking (Wild & Pfannkuch, 1999), also elaborated by other researchers (e.g., Konold, Higgins, Russell, & Khalil 2003; Reid and Reading, 2008). On the foundations of statistical thinking, my interest will be on my study participants' considerations of variability, for example, actions such as: noticing, estimating, predicting, testing and describing variation in a data set. The second theoretical perspective is Vygotsky's (1978) semiotic mediation and his socio-cultural historical theory of learning, also elaborated by other researchers (e.g., Falcade, Laborde & Mariotti, 2007; Bartolini Bussi & Mariotti, 2008). Vygotsky's theory incorporates the use of cultural tools for mediating meanings of concepts, which relates to my use of *Sketchpad* as a cultural tool to design learning tasks in a computer-based environment. I now discuss each theoretical perspective in the following sections.

3.1. Foundations or Statistical Thinking

According to Hacking (1975, cited by Pfannkuch and Wild, 2004), reasoning based on evidence from available data is relatively a new development and been slow in its growth and importance. Hacking suggests two changes that have shaped, in general, the current thinking, including what is considered to be the nature of knowledge and the nature of evidence. First, he submits that the concept of knowledge shifted from an absolute truth toward knowledge based on opinion. That shift, according to Hacking, resulted into thinking also shifting toward probabilistic perspectives. Second, the nature of evidence also shifted away from pronouncements, for instance, by those in positions of authority, toward inferences based on data (e.g., Stigler, 1999).

In his 1999 book, "Statistics on the Table", Stigler further adds to, and apparently agrees with, Hacking's propositions about the importance of data for supporting claims.

Stigler quotes a 1910 letter from Karl Pearson, a very prominent statistician, and regarded by some as the author of modern statistics, to The *Times* of London. Pearson was issuing a challenge to scientists, as well as to the lay-people, to do more than merely assert an answer—evidence must be provided, and that evidence need not be quantitative: “if the question is important and one position has been advanced with well-considered supporting evidence, then it is incumbent upon a critic to ‘put statistics on the table’” (p. 1). Stigler clarifies that “putting statistics on the table” is more than just putting forward numbers as evidence. It involves a careful analysis of the “forces that would affect any data, methods for measuring and expressing the uncertainty […] and the conventions for settling issues such as […] when an assertion should be rejected or not” (p. 1). These shifts in thinking and reasoning, according to Hacking, initiated a new way of thinking in the statistics community.

In the same year that Stigler wrote his book, Wild and Pfannkuch identified five elements (also called the foundations), which they considered to be fundamental to statistical enquiry. They are: i) considering variability; ii) reasoning with statistical graphs; iii) integrating statistical and contextual knowledge; iv) 'Transnumeration'—representing data in a way that enables a clear understanding of the concepts; and v) recognizing the need for data. Wild and Pfannkuch’s foundations of statistical thinking perspectives are indeed quite broad, having many applications outside statistics education (e.g. in industrial processes). However, I will focus my examples to statistics and mathematics education. In connection with the considerations of variability, I incorporate Carlson, Jacobs, Coe, Larsen and Hsu’s (2002) work on covariational reasoning, which compares change in variable ‘A’ with change in another variable ‘B’ in a functional relation. Although Carlson et al.’s covariational reasoning was not developed directly from a statistical research, but from mathematics, it has some relevance to consideration of variability.

3.1.1. Consideration of Variability

Wild and Pfannkuch (1999) make five assumptions in their considerations of variability: i) variability is an observable reality; ii) some variation can be explained whereas others cannot be explained based on our current knowledge; iii) random variation is the way statisticians model unexplained variation; iv) unexplained variation may be produced through random sampling; and v) randomness is a human construct for describing variation where patterns cannot be detected. Assumptions iii), iv) and v) which focus on random variation, are not discussed in the current dissertation. I discuss assumptions i) and ii) using two examples from Reid and Reading's (2008) work on the considerations of variations that I mentioned in Chapter 2. I will also briefly discuss Carlson, Jacobs, Coe, Larsen and Hsu's (2002) notions of covariational reasoning.

Consideration of Variation Hierarchies

Reid and Reading's (2008) study, introduced in Chapter 2, is based on Wild and Pfannkuch's considerations of variation principles. Reid and Reading analyzed students' responses to class tests and homework assignments, and also revisited similar work that they previously developed (e.g. Reading & Reid, 2005; Reid & Reading, 2004, 2006), to formulate a consideration of variation hierarchy (CVH) framework in Table 1. In Table 1 each of the four hierarchies (i.e., none, weak, developing, and strong) has a set of descriptors shown in the second column. The descriptors were developed from specific tasks, which could suggest that the hierarchies also depend on the learning tasks that are assigned to the students, and so they can be contextual to the tasks.

Table 1. Consideration of Variation Hierarchy (Reid & Reading, 2008).

Consideration of Variation Hierarchy	Descriptors
None	– do not display any meaningful consideration of variation in context – do not acknowledge variation in relation to other concepts (e.g., distribution)
Weak	– identify features of only one source of variation – acknowledge variation in relation to other concepts – incorrectly describe variation – do not base description of variation on the data – anticipate unreasonable amount of variation – poorly express description of variation – refer to irrelevant factors to explain variation – incorrectly refer to relevant factors to explain variation – do not use variation to support inference

Consideration of Variation Hierarchy	Descriptors
Developing	– recognize the effect of a change in variation in relation to other concepts – correctly describe variation – base description of variation on the data – anticipate reasonable amount of variation – clearly express description of variation – correctly refer to relevant factors to explain variation – use variation to support inference – do not link the within-group and between-group variation
Strong	– link within-group and between-group variation to support inference

Reid and Reading propose that to be categorized as “weak”, “developing” or “strong” consideration of variation, a student’s response on an assigned task should include the first item on the list of descriptors for that hierarchy. For example, in the ‘weak’ consideration of variation hierarchy (CVH), the first descriptor is that a student is able to ‘identify features of only one source of variation.’ However, given that responses to tasks may vary from one student to another, I suggest that using more than one descriptor in each hierarchy could capture a more accurate range of students’ consideration of variation. Given that the hierarchies were not developed with a focus on the use of particular computing tools, such as, in my case the dynamic technology, the CVH does seem to help me interpret data obtained from a similar environment, namely the static environment, not the dynamic environment.

Covariational Reasoning

Covariational reasoning is a framework developed by Carlson, Jacobs, Coe, Larsen and Hsu (2002) for interpreting university calculus students’ reasoning with co-varying mathematical entities in a dynamic setting. The authors define covariational reasoning as cognitive activities involved in coordinating two varying entities, while attending to the ways in which the entities change in relation to each other. For the authors, covariational reasoning includes mental actions, which can be described by observed behaviour of the student, but the behaviour is associated with cognitive functions of the individual student. For instance, mental action at level one is associated with coordinating the value of one variable with changes in another variable. The observed behaviour associated with this level is the verbal coordination of two entities such as a statement ‘entity A changes as entity B changes.’ Carlson et al. proposed five mental actions that were associated with students’ behaviours on a calculus task. However, only the first three (see Table 2) of the five Carlson et al.'s categories suit my study. The fourth category (i.e., coordinating the average rate of change); and the fifth category (i.e., coordinating instantaneous rate of change of the function), do not directly apply in my study, but they apply to calculus in mathematics.

Table 2. Covariational Reasoning Framework (Carlson et al., 2002).

Mental action	Descriptions of mental actions	Behaviours
Mental action 1 (MA1)	Coordinating the value of one variable with changes in the other variable	Verbal indications of coordinating two variables
Mental action 2 (MA2)	Coordinating the direction of change of one entity with change in the other entity.	Verbalizing an awareness of the direction of change in one entity while considering changes in the second entity.
Mental action 3 (MA3)	Coordinating the amount of change in one entity with changes in the other entity.	Verbalizing an awareness of the amount of change in one entity while considering changes in the second entity.

Carlson et al. assign a level of mental action to the learner according to the highest level attained. If, for example, a student’s mental actions reveal quantitative coordination—coordinating the amount of change in one entity with changes in the second entity—then the student is assigned at level three. From Table 2, and the discussions in this section, it is obvious that not all the levels of mental actions proposed by Carlson et al. apply to Wild and Pfannkuch’s considerations of variability in statistics. However, I modified the framework and used it in statistics for interpreting students’ reasoning with dynamic activities. Moreover, Wild and Pfannkuch emphasize statistical reasoning and thinking without putting any categories to student interactions. I will adopt Wild and Pfannkuch's considerations and categorize students' interactions from tasks, but not the students themselves.

3.1.2. Reasoning with Statistical Graphs

The second element in Wild and Pfannkuch's (1999) foundations is about reasoning with graphs. Although Wild and Pfannkuch's use the term 'model', I interpret a model to include a graph. For example, a statistical graph such as a histogram can be used to reason about the variability of a data set. Pfannkuch and Wild (2004) explain that:

> When we use statistical models [*graphs*] to reason with, the focus is more on the aggregate-based rather than individual-based reasoning […]; individual–based reasoning concentrates on the single data points and with little attempts to relate them to the wider data set, whereas aggregate-based reasoning is concerned with patterns and relationships in the data set as a whole. (p. 20)

Hancock, Kaput, and Goldsmith (1992) add that students' ability to reason from group tendencies rather than from individual cases is fundamental in developing their statistical thinking. Other researchers (e.g., delMas & Liu, 2005) have proposed using models, including dynamic graphs, to help the students to visualize patterns and to reason about the distributions of data sets.

3.1.3. Integrating Statistical Data with Context

Biehler and Steinbring (1991) contend that data cannot be detached from its context. For Biehler and Steinbring, exploring concepts with data should include asking and answering questions about the patterns that are observed in the data, and whether or not the outcomes relate to the context of the data. Other researchers (e.g., Shaughnessy, Garfield & Greer, 1996; Pfannkuch & Wild, 2004) agree that students need to have the mind-set of a 'detective', to look for information hidden in the data since data arise from specific contexts (Pfannkuch & Wild, 2004). I apply the notion of 'data detective' and 'looking for information hidden in the data' in my methodology in Chapter 4, where I ask participants to make conjectures about the patterns in the graphs, justify their conjectures and then later check them through interactions with the graphs.

3.1.4. Representing Data to Enable Clear Understanding of Concepts

Transnumeration is a term invented by Wild and Pfannkuch (1999) to describe a learning environment (and also outcome) where data are represented in clear ways through graphs or models, to help students understand generally challenging statistical concepts. The process of transnumeration can include representing data in a graphical form so as to make the patterns of variations in the data set easily visible to the students. For example, using a histogram to present data set so that the spread of data from the mean reveal the shape of the distribution, providing clearer information about the variation in the data. Ben–Zvi and Friedlander (1997) suggest from their study on data representations and analysis that students who handle multiple representations of data in meaningful and creative ways, for instance, by searching for patterns in the data to convey mathematical ideas, are more likely to reason and to think statistically than those who are not able to handle data, for example, in meaningful and creative ways.

3.1.5. Recognizing the Need for Data

According to Wild and Pfannkuch (1999), being aware of the inadequacies of anecdotal evidence and personal experiences to base decisions is very important. For Wild and Pfannkuch, basing decisions on deliberately collected data is “a statistical impulse” (p. 227). By that, I suggest, Wild and Pfannkuch probably mean that basing decisions on data is something statisticians take for granted. The authors' views are consistent with Hacking as well as Stigler’s notion of putting statistics on the table that I discussed earlier. I will use these perspectives in both the design and the analysis of my data in Chapters 4 and 5. I now present the second theoretical perspective that influenced in my study—semiotic mediation.

3.2. Using Socio-cultural Historical Tools to Mediate Meanings of Mathematical Concepts

The contribution of different artifacts and instruments in the activity of learning mathematics is documented in many research studies (e.g., Rabardel 1995; Verillon & Rabardel, 1995; Bartolini Bussi & Mariotti, 2008). Rabardel considers an artifact as a material object designed according to a particular goal and as such holds some specific knowledge. For example, the Compass Tool is traditionally used in mathematics for constructing circles. However, the functions of the Compass Tool can be extended beyond the original design if it is applied in combination with other tools such as the Line tool to construct, for instance, a square. However using a tool in combination with other tools requires specific knowledge and skills by the user. According to Rabardel and Samurçay (2001), an instrument is different from an artifact in that it possesses both an artifact-type (non-task specific) component as well as the utilization schemes (components of an artifact which are used to solve specific tasks). According to Rabardel (1995), the instrumental genesis—the process of transforming an artifact to an instrument—is not physical but a psychological process, which depends on the knowledge of the user and how he or she applies an artifact to a specific task.

Moreover, instrumental genesis includes two other processes—the instrumentalization (how a user applies an artifact on a task) and the instrumentation (how the artifact impacts the user's knowing)—such that the artifact is considered to have evolved into an instrument. The instrumentation process differs from the instrumentalization process in that in the former, the knowledge that a user may gain by using an artifact is constrained by the design of the artifact (Trouche, 2005). Going back to the earlier example of the Compass tool, whereas it can construct a circle, it cannot construct a triangle without combining its function with that of a Line tool.

However, critiques of instrumental genesis point out that it does not go far enough to explain how learners can gain an understanding of mathematical concepts (see Bartolini Bussi & Mariotti, 2008). Instead, instrumental genesis focuses more on the processes of transforming an artifact into an instrument than on learners' understanding of mathematical concepts. Bartolini Bussi and Mariotti caution that the contribution of an instrument to cognitive development is a delicate issue that should be carefully considered so as not to over-simplify a complex problem. In my study, I use semiotic mediation for two reasons. First, semiotic mediation fits in well with the use of social-cultural tools, which include modern computing tools as well as locally designed mathematical sketches. Second, semiotic mediation links the use of physical tools with cognitive tools in the activities of teaching and learning mathematical concepts. I take a

Vygotskian approach, which distinguishes psychological tools (internal tools) from corporeal tools and categorizes the latter as external tools.

3.2.1. Semiotic Mediation

According to the Vygotskian theoretical perspective, artifacts are products of human activity, which play a fundamental role in cognitive development (see Vygotsky, 1978; Falcade, Laborde & Mariotti, 2007; Bartolini Bussi & Mariotti, 2008). The Vygotskian perspective assumes a dialectical dependency between physical external tools and signs. By a dialectical dependency, Vygotsky (1978) highlights both the common elements and the differences between tools and signs:

> The tool's function is to serve as a conductor of human influence on the object of the activity; it is externally oriented; it must lead to changes in the objects […]. The sign on the other hand, changes nothing in the object of a psychological operation. It is a means of internal activity aimed at mastering oneself; the sign is internally oriented. (p. 55)

Vygotsky's perspective on tools and signs appears different from other psychological perspectives in that he highlights the different functions played by a tool and a sign in a cognitive process. He defines internalization as the process of transforming an external tool into an internal tool. The analysis of internalization involves analyzing a system of signs (e.g., words, drawings, gestures, and accompanying actions), which follow an activity with the external tool (Falcade, Laborde & Mariotti, 2007; Wertsch & Addison Stone, 1985). Falcade et al. (2007) elaborate that the use of particular signs, especially speech, contributes to mental processes such as concept formation. For Falcade et al, a specific tool may function as a semiotic mediator and by that they mean new meanings related to the actual use of a physical tool by a student, do evolve, often under the guidance of an expert, or a cultural mediator such as a teacher, to a correct understanding of mathematical ideas represented in the tasks.

Falcade et al. further elaborate that the Dragging tool may be considered as a sign (an instrument of semiotic mediation) referring to its role in bringing about changes in the different entities, for instance in a dynamic graph in a Sketchpad environment. .

The authors propose that personal meanings that emerge from students' dragging actions in the *Sketchpad environment* can evolve into mathematical meanings.

3.2.2. Analysing the Semiotic Potential of the Dragging tool

Bartolini Bussi and Mariotti (2008) propose two basic domains in a dynamic geometry environment (DGE): the construction domain and the motion domain. The construction functions of a DGE (e.g. Sketchpad) allow one to produce graphics on the computer screen by using construction tools (e.g., the Point tool, Line, and Bisector tools). However, the motion domain works through the dragging action of the Dragging tool so that dragging action preserves the properties that define a given sketch. For example, a properly constructed square in a DGE retains its mathematical features, even if the original size of the square increases or decreases through dragging.

Bartolini Bussi and Mariotti identify two kinds of motion in the DGE: the direct and indirect motion, both related to different utilization schemes. Direct motion occurs when a basic element such as a point is dragged on the computer screen by acting directly on it using the mouse. Indirect motion occurs if by dragging a basic point on the sketch, other points also move, but do not alter the geometrical properties of the sketch. Bartolini Bussi and Mariotti propose that using the Dragging tool to move the basic point within the sketch can enable the user to experience the functional dependency between direct and indirect motion so that the Dragging tool can be considered as a sign referring to the mathematical notion of variability. Moreover, the "the space/time movement of a point on the screen preserves the content of the sign as it evolves from the artifact sign (moving point) to the mathematical sign (variable)" (p.757). According to the authors, a sign can be used as "an index of the move from personal sense to mathematical meaning" (p. 757). In other words, the signs shape conceptual understanding of mathematics through moving from an explicit use of the artifact to the mathematic context. The authors propose three categories of signs that can evolve from a student's use of dynamic tools in a mathematical task:

Artifact signs refer to the context and activity that the artifact is applied to. Bartolini Bussi and Mariotti (2008) submit that the signs "sprout from the activity with the artifact, their meanings are personal and commonly implicit, strictly related to the experience of the subject" (p. 957). In other words, the artifact signs are similar, for example, to Noss and Hoyles' (1996) "situated signs"—how a learner develops mathematical meaning in a particular learning environment—and also to Radford's (2003) "contextual generalization", which describes the kind of expressions a learner applies in the context of using a tool on a particular task. Bartolini Bussi and Mariotti propose that artifact signs are connected to the mathematical signs through pivot signs (Fig. 4).

Pivot signs as stated above, link the artifact sign to the mathematical signs and they may refer to activities in both domains. Bartolini Bussi and Mariotti explain that pivot signs are characterized by their shared meanings in instrumented actions but also through natural language:

> The characteristic of these signs is their shared polysemy, meaning that […] they may refer both to the activity with the artifact; in particular […] to specific instrumented actions, but also to natural language, and to the mathematical domain. Their polysemy makes them usable as a pivot/hinge fostering the passage from the context of the artifact to the mathematics context. (p.757)

The diversity of meanings that characterises pivot signs could make them problematic to distinguish from artifact signs and from mathematical signs since they assume a position between the two signs. However, Bartolini Bussi and Mariotti add that pivot signs "express a first detachment from the artifact, but still maintaining the link to it in order not to lose the meaning" (p. 757), which provides more guidelines on pivot signs. Since each instrumented activity is assumed to be unique, the guidelines seem adequate to identify pivot signs from both the artifact and mathematical signs.

Mathematical signs differ from artifact signs in that they refer to the mathematical context and meanings shared within a cultural community of mathematicians (Sfard, 2008). The signs may be expressed as a "proposition, for example a definition, a proof, or a statement requiring proof according to the standards shared by the mathematical community" (Bartolini Bussi & Mariotti, 2008, p. 757). In the example that follows (Fig. 5), I discuss the emergence of artifact, pivot and mathematical signs from a teaching experiment (Falcade, Laborde & Mariotti, 2007; Bartolini Bussi &

Mariotti, 2008) conducted in 10th grade mathematics classes. The example was chosen to show how different signs can emerge from using an artifact in a mathematical activity.

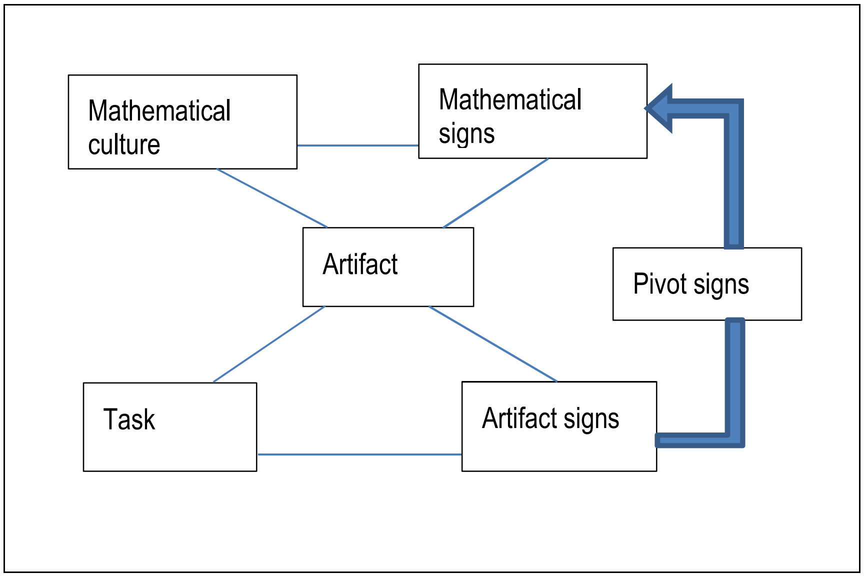


Figure 4. Relationship among the different signs in a mathematical activity. At the basic level are the artifact signs and at the higher level are the mathematical signs. Pivot signs link artifact signs with mathemtical sign (adapted from Bartolini Bussi & Mariotti, 2008).

Figure 4, provides an adapted version of how the artefact, pivot and mathematical signs emerge from an instrumented mathematical task. The artefact signs evolve in the activity of using a physical tool to solve a given task. Combining with other signs that are produced during solving a mathematical task, the artifact signs link to the pivot signs, which may evolve into mathematical signs. There may be situations where the pivot signs are indistinguishable from artefact signs or from the mathematical signs. In such a case, the pivot signs could appear to be by-passed by the artefact signs that evolve directly into the mathematical signs. The example below attempts to highlight some of the above scenarios.

3.2.3. Example

Three basic points A, B and P are shown in *Cabri* dynamic environment. The construction in Figure 7 provides point H as the orthogonal projection of point P onto line AB. Falcade, Laborde & Mariotti (2007) chose this construction because they wanted a strong visual representation of the notions of the domain and image of a mathematical function on the screen. Students worked in pairs and explored the effect of systematically moving one of the three points A, B and P, one at a time. They were to fill in a table explaining what moved and what did not move when they dragged each point. Finally, each small group was asked to write a common answer on the worksheet. Below is what student FE wrote:

FE: [1] Moving P, we realized that H is moving , whatever direction P is
[2] moving, except when P goes perpendicular to the line on which H is
[3] moving, (the line) passing through B and A. Dragging B, H forms a
[4] circle, passing through P and A.

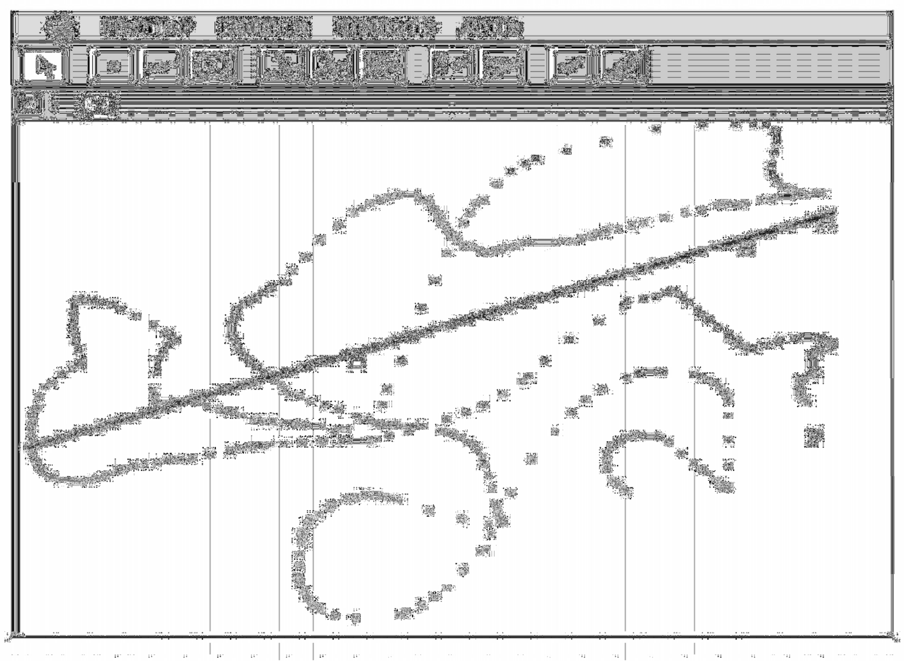


Figure 5. The graphic marks appearing on the screen after using the Dragging tool and Trace tool. The Trace tool can be considered as a sign referring to the mathematical notion of trajectory. (Bartolini Bussi & Mariotti, 2008, pp. 770-771).

In lines [1-2], FE's statements, "'moving P' … 'H is moving'…whatever direction 'P is moving'" are examples of artifact signs. However "'P goes perpendicular' [2] and 'H is moving'" can be considered as pivot sign that connects to the mathematical sign in lines [3-4], "dragging B, H forms a circle passing through P and A." In this example the Dragging tool is used as an instrument of semiotic mediation for the meaning of circle.

The second example, still based on Figure 5, is from a different student, JK:

JK: [5] The initial points are named independent variables, in fact they can be [6] moved individually and in our case all over the plane. H is called [7] dependent variable and we understand easily why, it cannot move by [8] itself, but always in function of some other moment (that is, it depends [9] on it).

JK's statement in line [5], "the initial points are named independent variables" suggests evidence of the mathematical sign, "independent variable", referring to the free movement of the initial points and stating the relation between motion and the variability of points "all over the plane." The statement in [lines 5-6], "the initial points are named independent variables, in fact they can be moved individually and […] all over the plane" combines both artifact signs ("moved individually and … all over the plane") and the mathematical sign (named "independent variables"). From Figure 5, it can be argued that the space/time movement of the basic points on the screen (artifact signs) evolves into a mathematical sign (variable). In this example, the pivot signs seem by-passed by artefact signs evolving directly into the mathematical signs. Dragging tool is an instrument of semiotic mediation for the functional connection among the moving entities A, B, P and H.

3.3. Summary

In this chapter, I have discussed two main theoretical perspectives that inform my research. First, I discussed Wild and Pfannkuch's foundations of statistical thinking, but focused on examples on considerations of variability, such as students reasoning with different types of statistical models, including graphs. Second, I used semiotic mediation theoretical perspective (Vygotsky, 1978). The Vygotskian perspective and distinguishes cultural tools (artifacts) from psychological tools in that the former serve external functions (interpersonal or social functions), whereas the latter serve intrapersonal functions (internal or psychological functions). Vygotsky uses the term internalization to denote a process whereby an external tool is transformed into a psychological tool. Similarly, the artefact signs can evolve into mathematical signs, for instance, from a student's actions with the tool as well as a realization that is mediated by the signs produced while he or she solves the task (Example in Section 3.2.3).

The main assumptions discussed in this chapter with regard to mathematical tasks implemented in a dynamic geometry environment (DGE) are: i) Dragging activities support learners' making and checking conjectures by, for example, directly interacting with the graphs in in a DGE (Arzarello et al., 2002); ii) The use of particular signs, for example speech*,* contributes to concept formation by leaners (Bartolini Bussi & Mariotti, 2008; Vygotsky, 1978); iii) Tool mediated actions can shape learners' evolving mathematical meanings (Falcade, Laborde & Mariotti, 2007); and, (iv) Learning takes place through internalization—a process whereby social (interpersonal) experiences are transformed into intrapersonal ones (Vygotsky, 1978).

Based on issues that were reviewed in Chapter 2, and on the theoretical assumptions summarized in the current chapter, my study explores participants' thinking about the notions of distribution, the mean and the standard deviation before, during, and after they have interacted with the dynamic sketches designed with the *Sketchpad*. The aim of my study is to gain an understanding of how students—in particular, students enrolled in a university-level introductory statistics courses—respond to the features of statistical variability, before, during and after interacting with the dynamic mathematics sketches. The specific research questions are:

1. What do students say about measures of statistical variability, such as *distribution*, *mean* and *standard deviation*, in a data set presented in a static environment?
2. How do students express the notions of variability while interacting with dynamic mathematics sketches?
3. How do students express notions of variability after interacting with dynamic mathematics sketches?
4. What might be the contribution of dynamic mathematics sketches to students' awareness of variability and statistical thinking?

Drawing on the assumptions discussed in this chapter, I hypothesize that my participants' descriptions of standard deviation, mean and distribution in the end will reveal more dynamic and physical expressions. I will consider such expressions as evidence of their consideration of statistical variability. I further hypothesize that my participants' interactions with the Dragging tool and dynamic sketches will enable them express meanings for standard deviation, mean and distribution more clearly than they were able to do in the static environment. Chapter 4 presents the methodology used in collecting data for the study. I will analyze the data in Chapter 5 and discuss the analysis of the data in Chapter 6.

4. Methodology of the Study

This chapter describes the procedures, materials and methods that I have used in the data collection. The chapter begins with an overview of the concept of variability, the standard deviation and the mean, aiming to set the background for the design of dynamic sketches, designed to explore the notions of statistical variability. Section 4.2 discusses further the design principles used in the dynamic sketches. In Section 4.3, I briefly outline the content that my study participants covered in their university-level statistics courses. I follow with a brief introduction of each participant in my study. Section 4.4 presents a report of a pilot study that I conducted, which enabled me to modify and improve on the initial study design. I present the interview tasks used in the data collection in Section 4.5. Given that I adopt a task-based clinical interviewing methodology (Ginsburg, 1981; diSessa, 2007), I discuss the merits of clinical interviewing in Section 4.6. In Section 4.7, I summarize the discussions in the current chapter.

4.1. Statistical Variability

In Chapter 2, I provided my working definition of the terms variability and variation, and in fact clarified that I use the terms interchangeably in this dissertation. Statistical variability, which is the focus in this study, is understood as a feature of the dispersion of data from its centre. Throughout this dissertation, the term variability is used to mean statistical variability, unless specific references are made to other types of variability. Although, the terms dispersion, spread and variability are often used interchangeably, I use variability and spread more frequently in my writing than I do dispersion. Variability (also called dispersion, scatter or spread) denotes how stretched or squeezed is a distribution (e.g., a theoretical distribution or the distribution of a statistical sample) from the centre. Common examples of statistical variability include the range, interquartile range, variance, and standard deviation. The current study focuses on the standard deviation, the mean, and distribution as features of statistical variability.

4.1.1. Standard Deviation and Variance as Features of Variability

The variance of a data set—symbolized by $s^2 = (1/(n-1))\sum(x_i - \bar{x})^2$ combines all the values in the data to provide a single measure of spread. Standard deviation, the square root of the variance—symbolized by $s = [(1/(n-1))\sum(x_i - \bar{x})^2]^{(1/2)}$ (where $\bar{x}$ is the mean, x_i is a given data point from the data set of finite size $n > 1$)—is the most commonly used measure of spread. If, for example, 2, 4, 6 and 4 are distances rounded to the nearest kilometre, from a central point, to four different locations and in four different directions, then the mean distance $\bar{x}$ = $(1/n)\sum_{i=1}^{i=4} x_i$ =4 km, the variance $s^2 = (1/(n-1))\sum(x_i - \bar{x})^2$ = 2.667km^2, and standard deviation, $s = \sqrt{2.667} = 1.633$km. Squaring the deviations from the mean and dividing the sum of the squared deviations by the degrees of freedom, (n-1), gives the variance. Variance can thus be described as the mean squared deviation of data values from the mean.

4.1.2. Standard Deviation and Mean as Features of Variability

Standard deviation can be conceptualized as a measure of the spread of data values from the mean. The magnitude of standard does not have a direct relationship to the magnitude of the mean in a data set. For example, the family of the Gaussian distributions in Figure 6, all the curves follow the general Gaussian density function given by $\Phi(x,\mu,\sigma^2) = [1/(\sigma(2\pi)^{1/2})]\exp[-(1/2)(x-\mu)^2/\sigma^2]$. The curves however, vary, one from another; those with relatively larger magnitudes of standard deviation (e.g. $\sigma = (5)^{1/2}$ and $\sigma = (1)^{1/2}$) appear more spread out than those with smaller magnitudes of standard deviation (i.e. $\sigma = (0.5)^{1/2}$ and $\sigma = (0.2)^{1/2}$). The blue curve with the smallest magnitude of standard deviation ($\sigma = 0.447$) shows the highest peak of the Gaussian curve among all the curves shown.

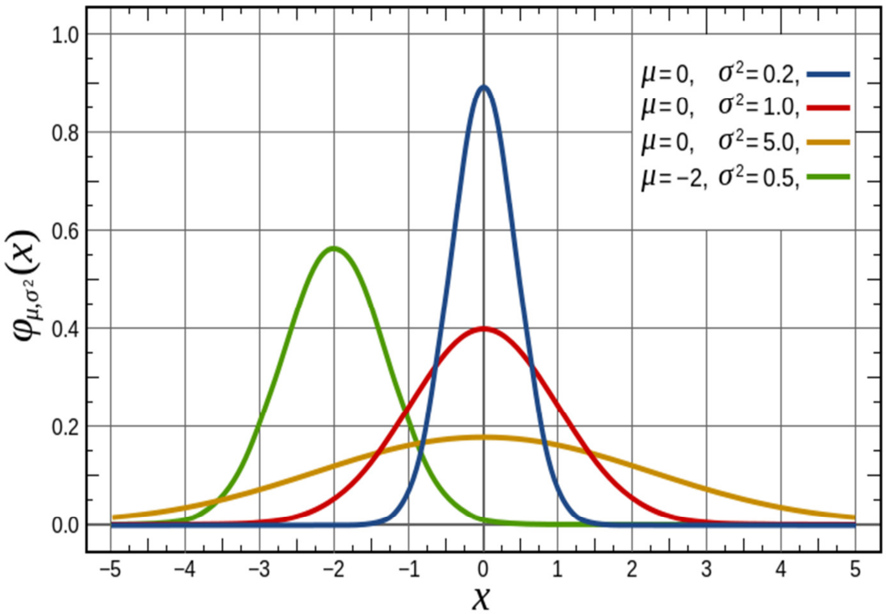


Figure 6. Gaussian density functions. The blue, red and yellow curves have the same mean. The shape of the distribution depends on the magnitude of standard deviation.

The yellow curve with the largest magnitude of standard deviation also shows the largest spread on the graph compared to the blue curve with the smallest magnitude of standard deviation. Figure 6 suggests the influence of standard deviation in controlling the spread of the Gaussian distributions.

4.2. Design of the Sketches

I designed two dynamic sketches using *Sketchpad. Sketchpad* was chosen among many similar software based on the background that I presented in Chapter 1, as well as the issues that I found in my review of literature in Chapter 2. Although some functions of *Sketchpad* are also available in other dynamic software, the dragging function *of Sketchpad*, which enables the user to interact with the sketches more directly and directly notice the changes, greatly influenced my choice of *Sketchpad.* In my Research Assistant work, discussed in Chapter 1, I had noticed that students pay more attention to the mathematical concepts built in the sketches if they have some direct involvement with the tasks.

I used a small sample (*n*=6) of data points in designing my sketches so that participants could more clearly notice the movements of individual points on the horizontal axis and focus their attention on exploring patterns in the data distribution. I relied on Konold's (2007) advice that I called 'parsimonious tool design'—in that it privileges using fewer variables over many variables in designing a tool, to help users find the essential elements they need from the tool. Konold explains that if "[...] you try to be helpful by including most of what everyone wants in a tool, it becomes so bloated that the users then complain that they cannot find what they want" (p. 9). My sketches incorporated as many features as were possible for participants to explore variability; but the sketches were also designed to ensure clarity of tasks to the participants.

4.2.1. The dyMS Sketch

The sketch in Figure 7, that hereafter I call the dyMS sketch (dynamic mean and standard deviation), is designed with six numerical data values positioned at points A, B, C, D, E, and F on the horizontal axis. Each data points on the number line take on a numerical value, which changes along the horizontal number line as the point is dragged horizontally from left to right and in the opposite direction. A data point can be selected using the mouse pointer, holding the left button of the mouse down and dragging the data point along the horizontal axis, using the Dragging tool of *Sketchpad.* As a data point is dragged to the left or to the right side of the mean line, the numerical scales for the mean and standard deviation change. In particular, the farther away a data point is from the mean the bigger is the size of standard deviation in the data set.

The mean-line in Figure 7 is a perpendicular line that passes through the mean (mean-point) of a data set distributed on the horizontal axis. The mean-line can be conceived as a geometrical representation of a moving mean. I define a mean-line by a perpendicular line that passes through the mean-point. The mean line also serves as reference line from where a square for each data point is constructed. I used a line instead of a point so as to provide a more visual representation of the centre of data as participants dragged data points on the horizontal axis. In fact, a major design consideration in using the mean-line over a point is that a line is more visual and physical for participants to easily see the "moving mean" as they dragged data points along the horizontal axis.

In Figure 7a, six data points have a mean-point, $\bar{x} = (1/n)\sum_{i=1}^{i=6} x_i$ =2.74 and standard deviation $s = [(1/(n-1))\sum_{i=1}^{i=6}(x_i - \bar{x})^2]^{(1/2)}$ =1.30. Each of six data points is positioned on the horizontal number line at some distance d_i ($i = 1,...,6$) away from the mean-point. In general, data point i at distance d_i from the mean line forms a geometrical square with area d_i^2. The sum of areas of all the squares is equivalent to the magnitude of sample variance $s^2 = (1/(n-1)) \sum_{i=1}^{i=6} d_i^2 = (1/(n-1))\sum_{i=1}^{i=6}(x_i - \bar{x})^2$. Hence, having data points at locations far away from the mean line implies that the sum of the squares will be large, resulting in large magnitudes of the variance and standard deviation. Standard deviation (the square root of the mean squared deviation of data values from the mean) is represented by the side length of the square obtained from the sum of the six squares constructed from the six data points to the mean-line. I used the square area so as to provide a visual representation of how variance (hence standard deviation) changes as data distribution changes through dragging points on the horizontal axis, using the Dragging tool of *Sketchpad.*

Figure 7a shows the six data points A, B, C, D, E, and F before any of them is dragged on the horizontal axis. In Figure 7b, data point F on the far right is dragged slightly toward the mean line, reducing the magnitude of standard deviation from 1.30 to 1.19. In Figure 7c, data point D on the right side of the mean line has been dragged across to the left side of the mean line and the magnitude of standard deviation has reduced further from 1.19 to 1.18.

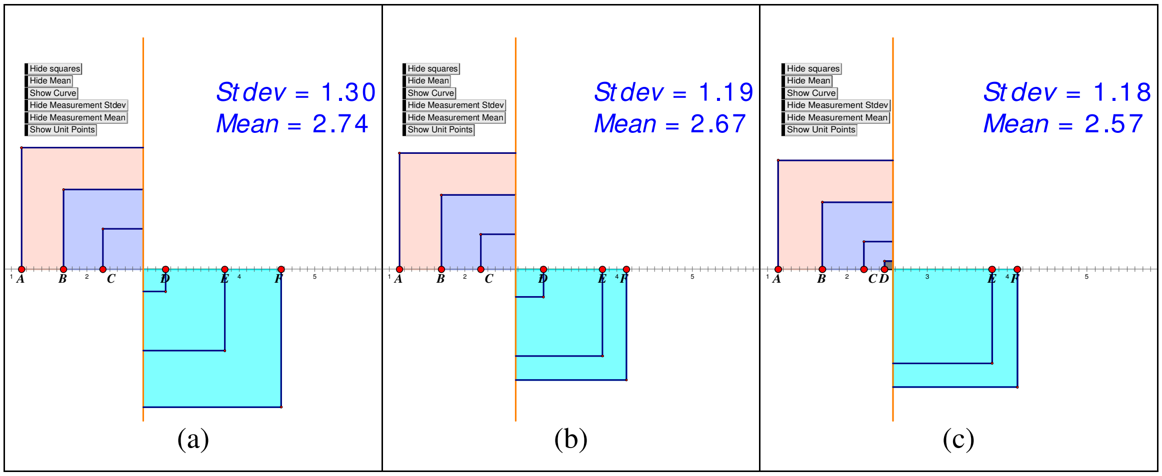


Figure 7. The dyMS sketch. (a) Before data points are dragged on the horizontal axis; (b) After data point F on the far right is dragged slightly to the left; (c) After data point D is dragged across the mean line to the left side.

The graphs in Figure 7 show that, standard deviation increases as the square areas increase. Conversely, standard deviation decreases as the square areas decrease. In other words, variability in the data distribution increases as data points are dragged farther away from the mean line.

4.2.2. The gC Sketch

Figure 8 provides the second sketch that I name the *g*C *sketch* (Gaussian Curve) with a density function given by $\Phi(x,\mu,\sigma^2)=[1/(\sigma(2\pi)^{1/2})]\exp[-(1/2)(x-\mu)^2/\sigma^2]$ as reviewed in Section 4.1.2. In designing the gC sketch, I assumed that the data sample used in the task is drawn from a normally distributed population. The normality assumption is important given that my data sample is small. Although participants did not have the density at each data point similar to what they would normally see, for example in a histogram, to be able to fit a Normal curve onto, I assumed, from a modeling viewpoint, that the density of each data point varied proportionately and symmetrically with its distance from the mean. The closer a data point is to the mean-line, the higher the density, so the peak of the gC sketch would be high at that point. Conversely, the farther away a data point was from the mean-line, the lower the density, resulting into a lower peak of the gC sketch. I expected that participants would be able to connect the distance of a data point from the mean-line with the change in the height of curve. Overall, my aim of fitting the Gaussian density function on numerical, but continuous data is to enable students analytically but also graphically explore changes in the height of the curve with the changes in the standard deviation as well as the distribution of data points on the horizontal axis.

Figure 8a shows the shape of the gC sketch before dragging the data points on the horizontal axis toward the mean line. The relatively large numerical value of standard deviation in Figure 8a confirms that there is more variability in the distribution as shown by the low height of the gC Sketch. In Figure 8b, data point P on the far right has been dragged closer to the mean line and the numerical value of standard deviation has reduced from 1.20 to 0.92. Figure 8c has each of the six data points slightly dragged closer to the mean line and the numerical value of standard deviation has reduced from 0.92 to 0.64, showing even less variability in the data distribution than in the other two sketches. Figure 8 shows that as the data points are dragged closer to the mean, variability reduces and the curve peak rises.

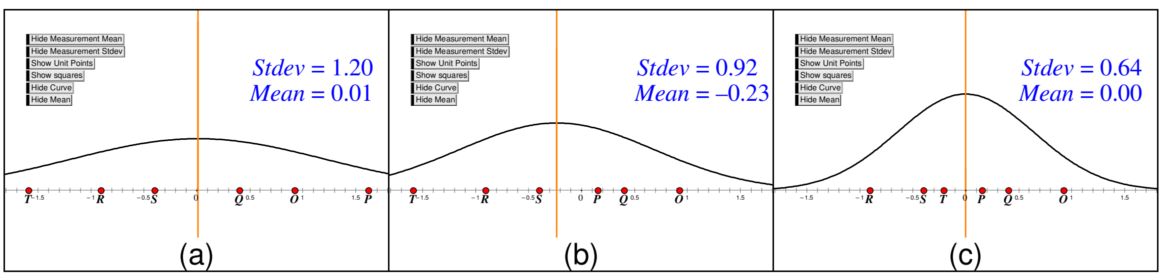


Figure 8. The gC sketch: (a) Original placement of data points; (b) After dragging far right data point F toward the mean line; (c) After dragging far left data point T toward the mean line.

The rise in the curve peak contrasts with the increase in the area of the square on the dyMS sketch as the magnitude of standard deviation increases.

4.2.3. Potential Limitation of the gC Sketch

The gC has some potential limitations which could include the following: Students may find it difficult to link a small number of discrete data points distributed on the horizontal axis with a normal curve that is fitted above the points (see Fig. 8). Students may be used to seeing, from their statistics courses and from textbooks, a normal curve that is constructed from a graph such a histogram. Unlike a histogram, the gC sketch may not provide students with a direct way of connecting the variability of each data point on the horizontal axis with the area under the approximated normal curve. To address that potential limitation, the gC sketch could be redesigned with each data point having unit width, but the heights vary along the horizontal axis to form a histogram that is approximately normally distributed. This may provide students with a clearer sense of the area under the curve, which is a very important component for determining statistical probabilities for continuous random variables located on the horizontal axis.

Given that the shape and position of a normal distribution curve depend on two key parameters—the mean and standard deviation—another possible option of designing the gC sketch would be to focus students' attention directly on how the mean and standard deviation vary with Gaussian curve. This could be done by designing sliders for the mean and standard deviation, and linking the two entities with the height of the Gaussian curve. However, this option may have a limitation as well in that it does not necessarily help students see how the curve relates to the actual changes in the distribution of data points on the horizontal axis. In others words, it does not seem to address Garfield and Ben-Zvi's concern that student generally have difficulties connecting the mean, standard deviation and the distribution of data.

4.2.4. Analysis of the Dynamic Sketches

Figure 9 shows how standard deviation and the mean vary as data points are dragged on the horizontal axis relative to the mean line. Dragging data points to the right side and away from the mean line increases the magnitudes of standard deviation and the mean, whereas dragging the points toward the mean line, but on the right side decreases both parameters.

Similarly, if data points are dragged on the left side of the mean line and away from it, the standard increases whereas the mean value decreases. Keeping on the left side of the mean line, if a point is dragged toward the mean line, the magnitude of standard deviation decreases, but the mean increases.

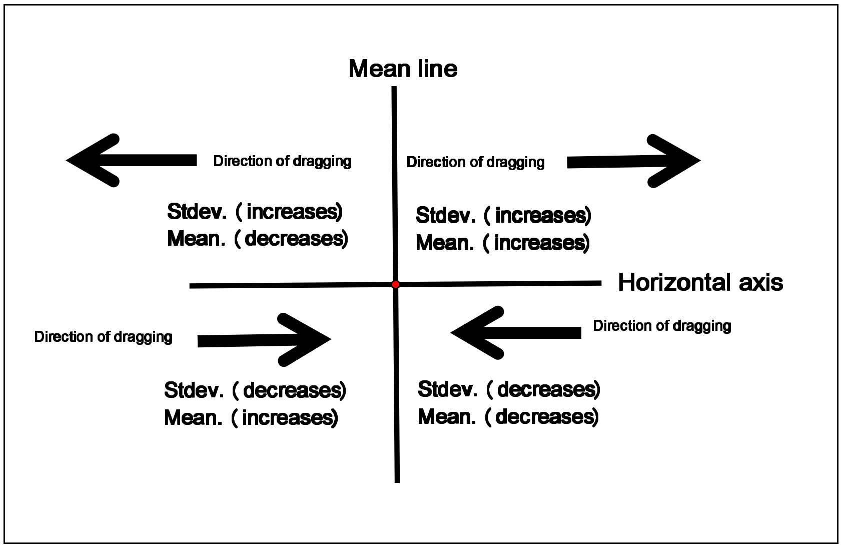


Figure 9. The behaviour of the standard deviation (Stdev.) and the mean as data points are dragged on the horizontal axis, to the right and to the left side of the mean line.

Moreover, if all the data points are selected and dragged on the horizontal line to the right or left side of the mean line, standard deviation remains the same, but the mean decreases as the points are dragged away to the left side and increases as points are dragged to the right side of the mean line. The conditions apply in both sketches in Figures 7 and 8. In this study, the dragging action is confined to the horizontal axis, either to the left or to the right side of the mean line.

4.3. Outline of the Concepts covered in the Introductory Statistics Courses

Having described the design of the sketches, I now present an overview of the topics that participants covered in introductory statistics courses before participating in the interviews. The participants all took a university-level introductory statistics course that included three to four hours of lectures per week for thirteen weeks and forty hours per week of drop–in tutorials, which were conducted in the Statistics Workshop. I believe that participants came into the study having covered all the concepts that they needed in in the interviews. Moreover, the interviews were not designed for students to explore the meaning of variability from a slightly different standpoint than what they may have learned in class.

The topics covered by participants in their statistics courses included: describing distributions with graphs (e.g., bar graphs, pie charts, box plots, histograms and stem plots); describing distributions with numbers (i.e., the mean, mode, median, standard deviation; and the Normal distribution (e.g. the Normal curve describing density, applying the 68-95-99.7 rule, and applying the Standard Normal curve to solve statistical problems). Additional topics covered by participants included scatterplots, correlation, simple linear regression, marginal and conditional distributions, the chi-square tests and introduction to tests of hypotheses. The topics are broad, but I focused on the foundation ideas of variability, which are important for participants to progress to the more advanced topics in their statistics courses.

4.4. The Study Participants

Participants in the study were university students currently enrolled in a one-semester first year statistics courses or who had completed the course in the previous semester. Instructors from the Department of Statistics taught the courses. I recruited volunteers for the study by contacting them directly from the Statistics Workshop, where I worked as a teaching assistant. At the beginning of the research study, I conducted a pilot study in which I interviewed seven volunteers. The analysis of the pilot video transcripts and notes enabled me to make three major changes in the final data collection framework.

First, the pilot study revealed that all my study participants had not used *Sketchpad*. In the final interview, I ensured that participants practiced with *Sketchpad* before they did the scheduled interviews. Second, the analysis of the pilot video also revealed that some participants did not justify their predictions as I expected them to do (see Appendix B and C), but began checking predictions on the computer immediately after stating them. In the final interview, I adjusted the protocol to ensure that participants stated their predictions and justified them before using the sketches to check predictions. One practical step that I took was to prompt participants to talk about their predictions *before* they could check on the computer. I also followed up their answers with further questions until they did not have more to say on a given thread of thought.

Third, I added more functions on the dynamic sketches to include “hide” and “show” buttons (Figs.7 & 8). For example, the “Hide measurement of standard deviation“ and “Hide measurement of mean” buttons made it possible for participants to make their predictions about changes in the mean and standard deviation without obtaining any hints shown on the sketches. Hiding the mean and standard deviation was my response to one of the recommendations that delMas and Liu (2005) made in their study (see Chapter 2). The information obtained from the pilot study contributed significantly to the changes that I made in the final study.

In the final phase, I contacted ten new students again through the Statistics Workshop. The selection of study participants was purposive in the sense that I picked students who took introductory statistics courses, and also who were active in the statistics workshop. I also considered diversity in my selection in terms of the subject major of participants, and their gender. Out of the ten students contacted, eight (four male and four female) participated in the interview. The two students who did not participate in the study had changes in their study schedules. After collecting data from the eight participants, I watched the individual video recordings several times. I also transcribed the video records into text and compared different responses to the interview questions. Moreover, I analyzed screen shots of participants’ actions during the interview as well. Based on my initial screening of data, I decided to exclude three sets of data based on the following grounds.

First, I left out data that had large portions of information that was not directly relevant to the topic of my study. For example, when asked about his/her thinking about the mean, one of the participants jokingly replied that the term ‘mean’ was to do a “mean person.” Another participant discussed the word “centre” in terms of the “epicentre” and gave an example of the “Tsunami” being the epicentre. Such were the portions of data that I felt would not relate to my research questions. Second, I also excluded data sets that I was unable to transcribe from audio to text because in some cases participants spoke so faintly that the words were not clear. Moreover, I did not have provision in my interview plan for recalling participants to clarify statements that they had made during the interviews. After very careful screening, I was left with data for five participants. Below, I present a profile of the five participants whose interview data are analyzed and reported in this dissertation. I have used pseudonyms Anita, Boris, Halen, Kars and Yuro to protect the identity of study participants as required by the ethics.

Anita
Anita was in her second year pursuing a bachelor's degree in health sciences. She had used SPSS software in her statistics course, but had not used *Sketchpad.* Anita practiced with *Sketchpad* before the final interview and was comfortable with the software during her interview. Anita did the interview in the last week of the semester before taking her final examination in statistics.

Boris
Boris was in his third year pursuing a bachelor of statistics degree, majoring in actuarial science. He had taken an introductory statistics course in the previous semester. Boris had not used *Sketchpad before* but had used JMP statistical software in his statistics course. He had a stronger background in statistics and mathematics than the other participants, partly as a requirement for gaining his current subject major. Boris did the interview toward the end of the semester.

Halen
Halen was in her third year pursuing a bachelor's degree in health sciences. Like Boris and Anita, Halen had not used *Sketchpad* before, but was familiar with the JMP and SPSS software from her statistics course. Halen practiced with Sketchpad before the interview and did not have problems using it in the interviews. She did the interview after writing all her semester examinations.

Maya
Maya was a third-year student pursuing a bachelor's degree in health sciences. Like the rest of the participants, Maya had not used *Sketchpad* before, but was familiar with JMP software. She practiced with *Sketchpad before* the interview and showed no difficulty using the software in the interview. Maya participated in the interview before her final examination in statistics.

Yuro
Yuro was in his third year pursuing a bachelor's degree in health sciences. He did the interview at the end of the semester after writing his final examination in statistics. Like all participants, Yuro had not used *Sketchpad before* but was familiar with the JMP software from his statistics course. Yuro practiced with Sketchpad before the interview and never showed any difficulty with Sketchpad in the interview.

As I stated earlier, the pilot study revealed that my study participants had not used *Sketchpad* before the interviews, but were familiar with statistical data analysis software. However, after some practice with *Sketchpad* before the interview*s*, participants showed no difficulty using Sketchpad in the interviews. Their familiarity with other statistical data analysis software could have contributed to their relatively short time to adjust to using *Sketchpad.* Also Sketchpad's dragging facility seemed to allow participants use it without prior knowledge of Sketchpad.

4.5. The Interview Tasks

I designed a set of four different tasks that altogether took approximately fifty minutes per interview session. The tasks include one short opening task (10-12 minutes); two longer tasks (30 minutes) that participants solved with the dynamic sketches; and one short task at the end of the session (10-12 minutes). In the opening task (Task 1), I asked participants to briefly describe their thinking about six terms in Appendix A. However, in the end I only analyzed their thinking about three terms shown in Figure 10, which I considered more related to my study on variability. I initially included the six items so as to gradually lead participants to the constructs that I was aiming for, rather than ask them directly about the three constructs. For example, I first asked them about “centre” before asking them to talk about the “mean.”

> **TASK 1:** In this activity, you are to briefly describe what comes to your mind when I mention the following terms: **(a) Distribution**; **(b) Mean;** and **(c) Standard deviation.**

Figure10. The interview task used to collect data at the beginning

The choice of the three terms in Figure10 was also informed by the issues that I found in my review of the literature (Chapter 2).

In Task 2 in Figure 11, I read out the instructions to the participants and clarified the tasks before they started to perform the task. I also ensured that they were comfortable with *Sketchpad* for checking their predictions.

TASK 2: You are to drag any of the six data points A, B, C, D, E and F (Fig. 7a) on the horizontal axis, to the left or to the right of the mean line and observe how the magnitudes of the standard deviation and the mean change. Before you drag any point, predict how standard deviation, the mean and areas of the squares change as you drag the point on the horizontal axis.

Figure11. The interview task based on the dyMS sketch

Task 2 relates to the properties of the standard deviation and the mean, which I reviewed in Sections 4.1.1-4.1.2.

In Task 3, shown in Figure 12, I read out the instructions to the participants and answered any questions before participants started on the task. I also ensured that participants explained their predictions before checking with the dynamic sketches. Task 3 connects changes in standard deviation and the mean, with changes in the height of the gC curve, as data points are dragged on the horizontal axis. Detailed description of Tasks 2 and 3 are in Appendices B and C respectively.

TASK 3: You are to drag any of the six data points T, S, R, P, Q and O (Fig. 8a), on the horizontal axis, to the left or right of the mean line and to describe how standard deviation and the mean, as well as how height of the gC sketch, change as you drag the points on the horizontal axis. Before you drag any of the points, first predict how standard deviation, the mean, and the height of the curve will change as you drag a data point.

Figure 12. The interview task based on the gC sketch

In Task 4, shown in Figure 13, participants reflected on a one-item task about standard deviation. I conjectured, based on participants' activities with the dynamic sketches that they would also include the notions of distribution and mean while reflecting on standard deviation. Hence, I did not ask participants about the mean and distribution directly.

TASK 4: In this activity, you are to say what comes to your mind when I mention the term Standard deviation.

Figure 13. The interview task at the end.

4.6. Data Collection

I used a one-on-one, task-based, clinical interviewing method (Ginsburg 1981; diSessa, 2007) to collect data from my participants. All the interview activities were videotaped. In Task 1 (Fig. 10), I asked participants to describe their thinking about the concepts of distribution, mean and standard deviation before they interacted with the dynamic sketches. These terms were chosen, among the six original terms, because they relate more to the features of variability in a data set. This interview segment informed the study on students' thinking about the notions of standard deviation, mean and the distribution before interacting with dynamic sketches. The data collected from this segment contributed to answering my first research question.

In Task 2, with the buttons for standard deviation and the mean in Figure 7a, turned off so that no hints were provided on the screen, I asked participants to predict how the mean, standard deviation and the squares would change, if they dragged the any data point on the horizontal axis (to left or right side of the mean line). My interview data were participants' verbal statements, their drawings on paper or in the air, as well as body movements, especially their hand movements. I recorded participants' actions during predictions and also during interactions with the sketches. The data collected from this segment contributed to answering my second, third and fourth research questions.

In Task 3 (Fig. 12), which was implemented with the gC sketch (Fig. 8), I ensured that the numerical values of standard deviation and mean were turned off. I then asked participants to predict how the magnitudes of standard deviation, the mean as well as the height of the gC curve would change if they dragged the data points on the horizontal axis. I did not specify the direction of dragging because I assumed that participants would include the direction of dragging from the mean line in their predictions. I followed participants' initial responses with probes until they exhausted their thoughts. After predicting the changes on the sketch, participants used the Dragging tool and moved data points along the horizontal axis as they talked about the changes that they noticed on the gC sketch. The aim this task was for participants to connect the changes in the standard deviation and the mean, with the changes on the curve peak and the data distribution. Task 3 contributed answers to my second, third and fourth research questions.

In the Task 4, (Fig. 13), participants reflected on the notion of the standard deviation, similar to one of the questions they answered at the beginning of the interview. This task informed the study on participants' thinking about variability after interacting with the dynamic sketches. Data collected from Task 4 contributed to answers to my third and fourth research questions.

The video records included participants' verbal statements, drawings, as well as gestures, especially, hand movements. I watched the videotapes several times and also took screen shots of participants' gestures and drawings as they expressed their thinking about the notions of variability. I transcribed all the video records of the interviews into text and analyzed them. In my analysis, I paid close attention to statements that related to aggregate reasoning with data. I also analyzed participants' statements on their use of the Dragging tool and how it enabled them to think about variability.

I adopted clinical interviewing method, drawing on the work of Piaget (1972), and the more recent work of Ginsburg (1981) and diSessa (2007). Although Piaget's work focuses on young children, his theory has been reformulated to accommodate other age groups as well. From my review of literature, I found that clinical interviewing technique can facilitate collecting very helpful insights from participants through their statements. Those statements and the different expressions from participants can provide information that may lead to, for example, redesigning the task to solve a given mathematical problem. Moreover, clinical interviewing also supports hypothesis testing (Ginsberg, 1981) for instance, by asking participants to make predictions on specific tasks and then check their predictions. Furthermore, according to diSessa (2007), clinical interviewing method is also suitable for identifying the cognitive processes behind intellectual tasks. In fact, diSessa work relates more to my study in the sense that his research involves clinical interviews with participants using computer technology. The next section provides some theoretical perspectives that informed my data collection.

4.6.1. Theoretical Perspectives used in the Data Collection

I designed my data collection procedures broadly based on the foundations of statistical thinking, but more toward the considerations of variability. I thus aimed at helping my participants engage and reason with data distributions in the dynamic sketches other than use the sketches as if they were representing static information (Wild & Pfannkuch, 1999; Konold & Pollatsek, 2002; Hancock, Kaput, & Goldsmith, 1992). I also aimed at moving participants away from focusing on calculations, to, for example, detecting and discovering patterns in the data distributions; generating hypotheses about data distributions, and physically checking their hypotheses with the dynamic sketches (Pfannkuch, 2005a; Watson, 2005; Ginsberg, 1981). Moreover, I incorporated in my study, developing participants' skills in making judgements in ways that would help them derive meaning from my sketches (Friel, Curcio, & Bright, 2001). Furthermore, I considered in my designs, delMas, Garfield, & Oom's (2005) proposal that incorporating the ideas of area and density in graphs were important in developing students' understanding of theoretical distributions.

4.7. Summary

In this Chapter, I have described the design of two dynamic sketches (the dyMS and gC sketches) whose main functions are to show: i) how standard deviation varies with data distribution from the mean; and ii) how standard deviation controls the height, and shape the Gaussian curve. I have also discussed the interview tasks that I use in my data collection and provided the rationale for choosing clinical interviews over any other interviews. One of the influences of clinical interviews on this study is that it can provide in-depth information on participants' mathematical thinking based on their interview data. Finally, I have provided the theoretical perspectives that informed my data collection procedures. For example, the tasks were designed with the aim moving students away from focusing mostly on the skills of calculations to helping them detect, discover and reason about patterns in data distributions. The tasks aimed at provoking participants to generate, justify and test their hypotheses about the changes in data distributions, the mean and standard deviation. In Chapter 5, I analyze data that I collected following the Methodology discussed in the current chapter.

5. Analysis of Data

In this chapter, I use the two main theoretical perspectives that I discussed in Chapter 3 to analyze my interview data. First, based on the foundations of statistical thinking (Wild & Pfannkuch, 1999), and focusing on the considerations of variability, I analyze episodes of my participants' considerations of features of variability in the static and in the dynamic computer-based environments. By the "considerations of variability", I mean participants' expressions that reveal evidence of their understanding of the features of variability, and the functional connections among standard deviation, the mean and the distribution of data set. By "understand", I mean participants' verbal or written expressions, which reveal mathematically/statistically accepted statements. The second theoretical perspective used in the analysis of data is semiotic mediation (Vygotsky, 1978; also elaborated by Falcade, Laborde & Mariotti, 2007; as well as Bartolini-Bussi & Mariotti, 2008). Using the semiotic mediation lens, I will focus my analysis on participants' use of signs produced during the tasks and how the signs mediate participants' understanding of statistical variability.

Given that I use two main theoretical perspectives to analyze data for each of the five participants, some episodes of participants' data will be analyzed from more than one theoretical perspective. I use the sign "[A#]" when citing an episode from participant A's data. For example, "[Anita, 5]" represents Anita's statement cited elsewhere in the dissertation. However, when citing in Anita's own data, I only use the "[5]" without her name. As well, the sign "[…]" is used to indicate lengthy statements that have been shortened for clarity, but without aiming at changing the original meaning. In my analysis I will use the terms "gC sketch" and "gC curve" to mean the same thing, especially when I refer to the "height" or "peak" of the gC curve. Lastly, the italicized statements in the square brackets are my own, aimed at providing additional information on the analysis. Data for the five participants are presented and analysed following the same order—Anita, Boris, Halen, Maya and Yuro—as presented in Chapter 4.

5.1. Anita

I first analyze Anita's interview data from Task 1, and then examine her interactions with the dynamic sketches in Tasks 2 and 3 (in Sections 5.1.2 and 5.1.3). In Section 5.1.4, I analyze Anita's final reflections in Task 4.

5.1.1. Activity in the Static Environment

In Task 1, I prompted Anita to "describe the term distribution."

> [1] Anita: Well, immediately what pops up is the normal distribution curve because that is part of what we have in statistics and also in my other courses; and we also learn every day in our lives as well [that] many things are normally distributed.

For Anita, her image of the distribution was the "normal distribution curve", because "that is part of what" she learned "in statistics and also in other courses." I asked Anita for some examples of the things that are "normally distributed" "in our lives."

> [2] Anita: For example, grades are normally distributed, so many of my course grades are normally distributed to get the normal standard for the class grades.

Anita's statements [in 1 & 2] linked the normal distribution to her "course grades", which she said were "normally distributed." Using Konold and Higgins' (2003) categories, Anita used the term distribution as a pointer to her course grades rather than distribution as a construct of how a data set is spread out from its centre.

I then asked Anita, "What about the term mean?" Anita's response focused on the "simple definition" of the mean and the procedure of obtaining the average of a data set.

> [3] Anita: I immediately think of averages because, um, well, simple definition of mean I guess would be adding up the numbers in the data set and then dividing it up by the figures you have [...].

By "adding up the numbers in the data set and then dividing it by the figures you have" [3], Anita's did not think about the mean as a balance point in a dataset that can (Hardiman, Well and Pollatsek, 1984). According to Reid & Reading (2008), descriptions that focus on single entities in a data set, with no attention to their connections to the distribution of data, show no consideration of variability. Anita's thinking of the mean focused on calculating a single value rather than considering the mean as the balance point in a data set.

When I asked Anita, "How about the standard deviation?" she paused for a moment and then said:

> [4] Anita: I think that [standard deviation] is kind of related to the mean. As I said before, if you […] can figure out the mean of data […], then you can derive the standard deviation, um and you can also figure out the normal distribution curve.

By "[standard deviation] is kind of related to the mean" in [4], Anita seemed to suggest that standard deviation is also obtained through a procedure similar to the way in which the mean is derived. She also correctly asserted that one needs the mean in order to compute the standard deviation. Further, she said that both the mean and the standard deviation were needed in order to "figure out" the normal distribution curve. I expected Anita to provide a more general sense of standard deviation as a measure of the spread of data from its centre. However, Anita's image of standard deviation focused on obtaining standard deviation by "deriving" it. Based on Wild and Pfannkuch's (1999) as well as Konold and Higgins' (2003) aggregate category, which considers a data set as a whole, including how it is spread out from the centre, Anita's consideration of standard deviation did not show aggregate reasoning. Rather she seemed incorrectly, to suggest that standard deviation meant the same thing as the Z-scores in the Standard Normal curve. I will now analyze Anita's interactions with the dyMS and the gC sketches.

5.1.2. Activity with the dyMS Sketch

In Task 2, based on Figures 7a and 11, I asked Anita to "predict how the squares would change if you move any of the data points on the horizontal axis?" Anita replied:

> [5] Anita: Well, if I take one [*data*] point for example B [below Fig. 14], if I move it away from the centre, um, square [*B*] might get ah, I think the [...] the square will move this way [*moved her left hand to left side of the sketch*] […] so it would get a bit bigger, that's what I think.

Although Anita had not yet used the Dragging tool to "move" the data points on the horizontal axis, her prediction "I think the […] square will move this way […] so it would get a bit bigger" seemed to consider the direction of movement and the magnitude of change in the square, which the interview question did not anticipate. The interview question anticipated that the considerations of variability would primarily be occasioned by dragging the data points on the horizontal axis using the Dragging tool. It seems that visualizing the sketch evoked an image of a physical motion in Anita's mind.

Moreover, Anita's prediction that "if I move it [point B] away from the centre […] the square will move this way […] so it will get bigger" acknowledged both the "centre" of the distribution and the deviation from the centre, suggesting an aggregate way of reasoning with data (Wild & Pfannkuch, 1999). Carlson et al. (2002) categorize statements such as [5], which consider change in the direction of movement of one entity in relation to change in magnitude of another entity, as evidence of quantitative coordination.

I asked Anita to talk a bit more about the change in the magnitude of the standard deviation, "What about the standard deviation?" I asked. Anita paused for a moment, and then replied:

> [6] Anita: If the mean gets smaller, then the standard deviation would um [pause] I think the standard deviation would um, get larger.

Using the analysis framework in Figure 9, Anita's prediction in [6] is correct, but only in the sense that dragging a point to the left side and away from the mean line would decrease the mean as standard deviation increases. Her prediction is consistent with Carlson et al.'s (2002) covariational reasoning, which accounts for a 'change in entity Y in relation to a change in entity Z'. In Anita's case, covariational reasoning is about the mean getting "smaller" as "standard deviation would […] get larger" through dragging a data point to the left side of the mean line. In general, the mean and standard deviation do not co-vary, but there are maybe cases in which both the standard deviation and the mean increase or decrease together (as in Fig. 9).

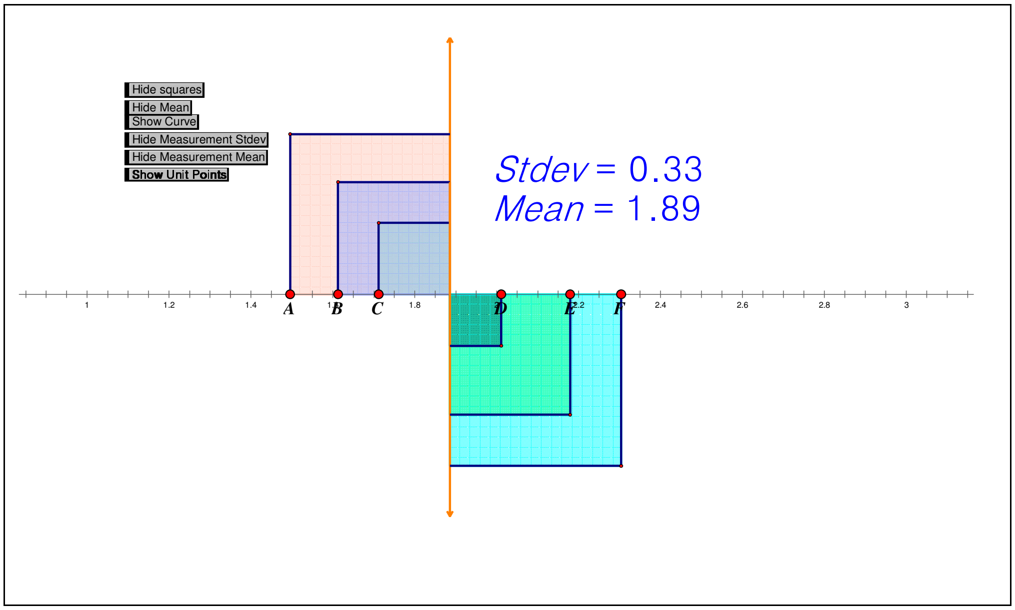


Figure 14. A snapshot of the dyMS Sketch before dragging the data points.

To prepare for checking her predictions, I asked Anita to click with mouse pointer, the buttons "Show measurement Stdev." and "Show measurement mean" on the dyMS sketch (Figure 14). Before Anita dragged a data point on the horizontal axis, the numerical values of standard deviation and the mean on the sketch were Stdev.=0.33; and Mean=1.89. I prompted Anita, "go ahead and […] check your predictions."

Anita selected data point **F** on the far right of the mean line with the mouse pointer and using the Dragging, she moved point F on the horizontal axis slightly to the right side. She stopped and said:

[8] Anita: The mean gets smaller [as] standard deviation gets larger.

It seems Anita expected the mean and standard deviation to change the same way as they did on the left side of the mean line [in statement 6]. I asked her to "move the data point [F] again and watch how the mean and standard deviation are changing". She dragged point F slowly to her right side and as the square grew bigger, she suddenly stopped dragging, suggesting some surprise. She said, "First, the mean got smaller, then it got larger, or maybe I'll drag it back again and see." She dragged the same data point briefly to her left and then back to her right side and continued dragging farther to the right of the mean line (Figure 15). Then she said:

[10]. Anita: Oh, so both of them *[the mean and standard deviation*] got larger, yeah, ok, so I thought the mean would get smaller and the standard deviation would get larger but actually both of them are increasing.

Anita's statement [10], "I thought the mean would get smaller and standard deviation would get bigger, but actually both of them were increasing" is consistent with the analysis in Figure 9. Her statement also suggests evidence that the Dragging tool enabled her to connect changes in standard deviation with changes in the mean, as the she dragged the data points on the horizontal axis. By "Oh, so both of them got larger", Anita showed some surprise on realizing that the standard deviation increased as the mean increased. She had predicted that the mean would decrease as standard deviation increased and her current result contradicted her prediction [8]. The contradiction seemed to cause a new realization by Anita, from her statement "Oh so both of them got larger, […]."

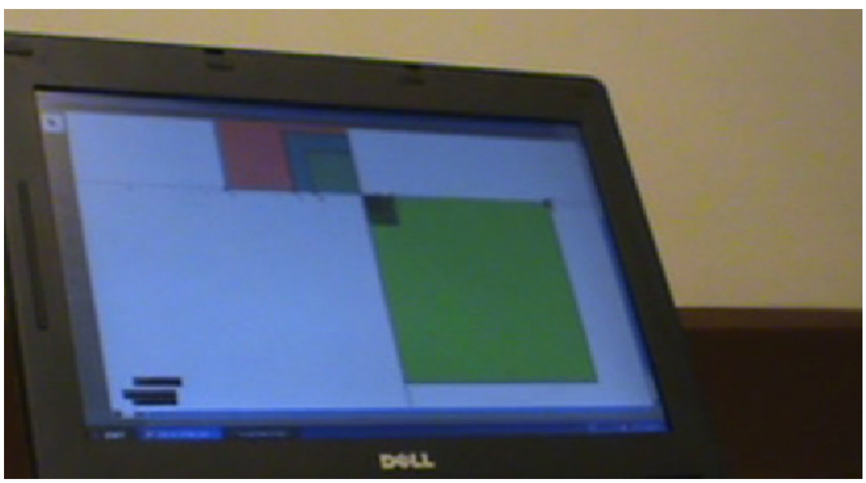


Figure 15. The square grew so big after Anita dragged data point F farther to the right.

I asked Anita to "explain why the mean increased as standard deviation increased." [10] Anita argued that if the distance from the mean of a data point was large, then the magnitudes of standard deviation and the mean in the data set would also increase:

[11] Anita: Yes, because um if you drag this data [*touched point F with mouse pointer*] farther to the right […] then the mean will increase and therefore the standard deviation will increase. If I move this [*touched point C,* Fig. 14*, with mouse pointer*] a little bit to the left, then it will decrease because I'm moving it to the negative side of the graph, so it slowly decreases as you could see.

The statement "if I move this [data point C] a little bit to the left, then it will decrease because I am moving it to the negative side of the graph […]" revealed Anita's consideration of variability in the sketch. Anita's may have used the change in the direction of dragging "to the negative side" to make a correct analogy with the "decrease" in the numerical values of the mean and standard deviation. Her statement supports my argument that the Dragging tool was used as semiotic mediator (Falcade et al., 2007) for Anita to link the changes in the mean and the standard deviation, with the direction of dragging the data point from the mean line.

Moreover, Anita's statement [11, lines 1-3], "if you '*drag'* this data", "the mean will '*increase***"** and "therefore standard deviation will '*increase',* suggests a linkage between the two entities, as well as the artifact signs describing actions with the Dragging tool (Bartolini Bussi and Mariotti, 2008). The statement "if you 'drag' […] farther to the right" [11, line 2] suggests a pivot sign linking the artifact sign to the mathematical sign, "then the mean will increase and therefore standard deviation will 'increase'." Based on Falcade et al. (2008), Anita's statement [11] supports my previous argument that she used the Dragging tool as a semiotic mediator to explore the relation between the mean and the standard deviation.

5.1.3. Activity with the gC Sketch

In Task 3, based on Figure 16, I asked Anita to "predict how the curve [*peak*] would change if you drag the data points on the horizontal axis." She predicted that the curve peak would "rise a bit more" if a point was dragged toward the mean line:

[13] Anita: I think if you move this data point at **O** (Fig. 16) toward the mean [*line*]*,* it will make the shape rise a bit more.

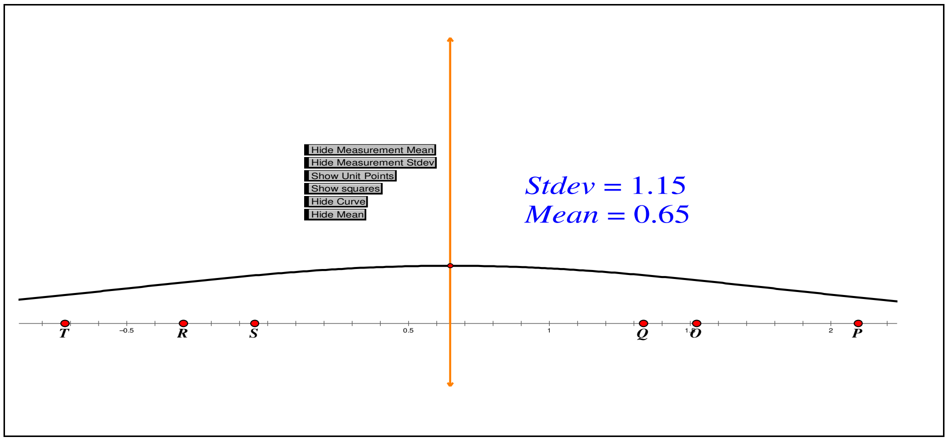


Figure 16. A snapshot of gC Sketch before the dragging action

Although Anita's statement [13] was correct, she did not justify it. I asked her why the peak of the gC sketch would "rise a bit more?" Anita argued that it was because at point P (Fig. 17), the curve "is really flat",

> [14] Anita: Because already at **P** point over here [Fig. 17] this part of the graph nearby the P point is really flat.

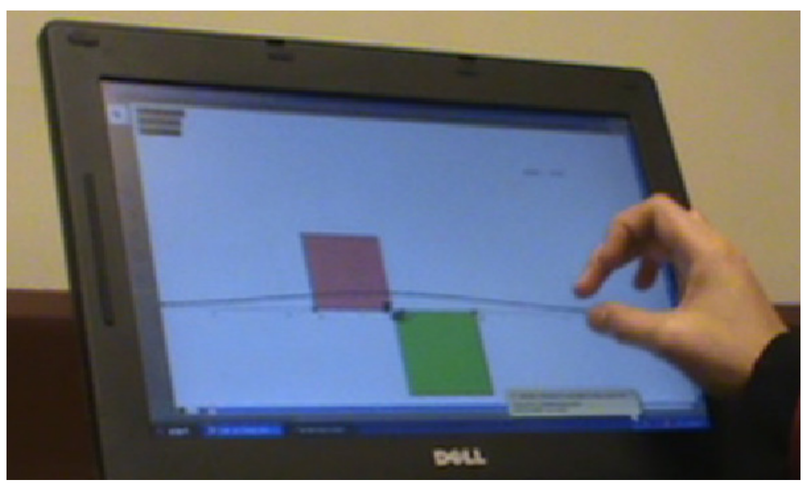


Figure 17. Anita gauges the flatness of the gC sketch at point P.

Anita positioned her thumb and index finger as if she gauging the thickness of the 'curve at point P (Fig. 16).

However, Anita's prediction was not explicit about the change in the magnitude of standard deviation in relation to the height of the GC sketch. In Chapter 4, I hypothesized that participants will be challenged predicting changes in the magnitudes of standard deviation and the mean, relative to the rise or fall in the height of the gC curve. I conjectured that participants would link the increase in the height of the gC curve with an increase in the magnitude of standard deviation and vice versa, contrary to the framework presented in Figure 9.

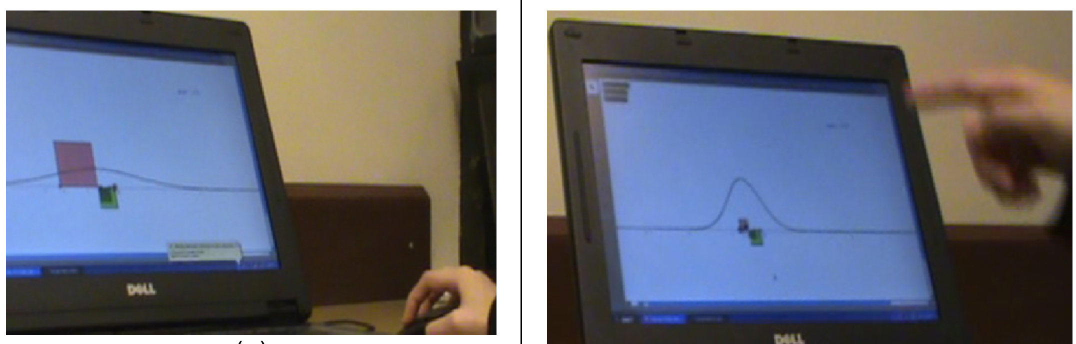
I probed Anita, "how do you know that the curve peak will rise as you move the data point toward the mean line?" She paused for a moment and then stated her "theory":

> [15] Anita: Ok, well um, I'm not exactly sure it [*the curve peak*] would actually rise or would be flat, but I have a theory that it will rise a little bit um, since it would be closer to these other data points [*touched on the data points nearer the mean line with the mouse pointer*]. I think with more clustered data points, the mean will also increase and therefore the graph would also rise.

Anita is correct in saying that clustering will make the peak rise, but she is incorrect in saying that clustering will cause the mean to increase. I asked Anita to "[…] go ahead and test" her theory with the gC sketch. She dragged data point O (Fig. 16) on currently on the right side toward the mean line. Then she dragged data point P, closer to point Q. As she dragged the points closer to the mean line, the curve peak kept on rising (Fig. 18a), and when she noticed the pattern, she said "oh!" as she momentarily stopped dragging. Anita dragged data points T, S and R, one after another toward the mean line and when the curve in Figure 18b showed, she said:

> [16] Anita: Actually, I didn't know that.

It's not clear if Anita "didn't know" because she seems to verify, through dragging, her prediction that clustering the data point will indeed increase the height of the curve. So perhaps she was not sure of her prediction. However, perhaps she noticed, as she dragged the data points towards each other, that the value of the mean was changing—when dragging points from the left toward the mean line, the mean increased and when dragging points from the right towards the mean line, the mean decreased. So her comment may indicate that she did not know how the mean would change when she clustered the data points.



(a) (b)

Figure 18. (a) Anita dragging data points closer to the mean; (b) Anita used her right index finger to draw a Normal curve in the air.

Based on Falcade et al. (2007) findings, I argue that the Dragging tool served as a semiotic mediator for Anita to connect the changes in the curve peak with changes in the spread of the data points on the horizontal axis.

5.1.4. Reflecting on the Standard deviation

In Task 4, I asked Anita "What do you say about standard deviation?" I expected her to connect the concept of standard deviation with the mean as well as with the distribution of data points on the horizontal axis.

> [17] Anita: Standard deviation, um, I did one of those examples based on the data points and the graphs, um, I realized that as the mean was increasing farther to the right, the standard deviation was also increasing, so that was a very direct relationship with the mean, […] whenever you moved a certain data point to the right or left, based on how much you moved it […].

By "I realized that as the mean was increasing farther to the right, the standard deviation was also increasing, so that was a very direct relationship with the mean", Anita referred to the functional linkage among the standard deviation, the mean, and the data distribution of data "whenever you moved a certain data point to the right […] the standard deviation was also increasing […]." Her statement, though incorrect, partially evokes Wild and Pfannkuch's (1999) aggregate reasoning perspective, since she seeks to establish a connection between different features of variability in a data set.

Anita's statement [17], "as the mean was 'increasing' [17, line 2] […] standard deviation was also 'increasing' [17, lines 2-3]" suggests the artifact signs (Bartolini-Bussi and Mariotti's, 2008). The statement "increasing farther to the right" suggests the pivot sign, which links the artifact sign with the mathematical sign, "the standard deviation was also increasing, so that was a very direct relationship with the mean" [17, lines 2-4]. Based on Bartolini-Bussi and Mariotti's classification of signs, Anita seem to move from the less mathematical and more action related statements (artifact signs), to the more formal mathematical statements (mathematical signs), such as, "I realized that as the mean was increasing farther to the right, the standard deviation was also increasing." In this example, the signs are the increasing numerical scales for the mean and standard deviation, as well as the statements that Anita makes about the changes she observes. I argue that the Dragging tool, used as semiotic mediator, enabled Anita to see the connections among the three features of variability—standard deviation the mean, and distribution.

5.1.5. Summary of Anita's Data Analysis

Analysis of Anita's data revealed that in the static environment, she focused mostly on the procedure for calculating the mean and deriving standard deviation from the mean rather than focusing on the meaning of those concepts. For example, to Anita, the mean was obtained by adding up data points and dividing the sum by the number of data points present. Anita also seemed to incorrectly suggest that standard deviation was the same as the Z-scores in the standard normal curve, instead of describing standard deviation more generally as how spread out or clustered a data set is from its centre. Thus, in terms of considerations of variability, Anita did not provide a clear link among the features of variability—standard deviation the mean, and distribution—being discussed in this study.

However, after interacting with the sketches, Anita showed some evidence of considerations of variability (e.g., in statements [13], [15], [16], & [17]). Anita's expressions after interacting with the sketches compared to her statements in the static environment suggest evidence that the Dragging tool mediated the mathematical meaning of standard deviation and enabled Anita to become more aware of the features of statistical variability.

5.2. Boris

Boris' followed the same interview format as Anita, starting with Task 1. In Sections 5.2.2 and 5.2.3, I analyze Boris' interactions with the dynamic sketches (Tasks 2 & 3). Section 5.2.4 provides Boris' reflections on the term standard deviation toward the end of the interview. In Section 5.2.5 I provide a summary of my analysis of his data.

5.2.1. Activity in the Static Environment

In Task 1, I asked Boris, "What comes to your mind when you hear the word distribution?" Boris replied:

[1] Boris: It's [the] observed frequency of some data [pause], yeah.

By the observed frequency of some data", Boris probably meant how a data set is distributed. Although Boris' thinking about distribution was in general a correct one, he did not refer to the centre of data and the spread of data from the centre. Boris' reply to my next question "what about the mean?" was equally a brief one. He answered in one word "average." Boris did not elaborate what he meant by "average" as he was generally brief in his answers. However, he probably meant the same process that Anita described about the mean in [Anita, 3].

I then asked Boris, "what about the standard deviation?" Boris paused for a moment, moved his right hand across toward his left side as he said:

[2] Boris: [*standard deviation*] kind of measures the variation of data from the mean [*pause*], standard deviation, deviation [pause], yeah finish with that.

Boris repeated the words "standard deviation" and "deviation" a number of times after stating that "[standard deviation] kind of measures the variation of data from the mean" [2]. Although Boris was generally brief, in general, his thinking about standard deviation is a more correct representation of standard deviation than Anita's. His thinking seems consistent with Wild and Pfannkuch's (1999) aggregate reasoning that emphasizes identifying the centre of data and deviation of data from the centre.

5.2.2. Activity with the dyMS Sketch

In Task 2, based on Figures 11 and 14, I asked Boris to "predict how the squares would vary if you drag any of the data points on the horizontal axis":

> [3] Boris: If you move [data points] away from the centre, the square is getting bigger and bigger because the square is the distance from the centre right? If you move away you get a bigger square […]

Boris' description, "the square is getting bigger and bigger […] if you move away from the centre", suggests dynamic thinking about the change in the square in relative to the distance of a data point from the mean line. His verbal statement not only suggests covariational reasoning (Carlson et al. 2002), but it also includes some aggregate reasoning with the graph as it relates changes in the spread of data with the "centre." It seems that the sketch in Figure 14 evoked the image of 'physical movement' in Boris' mind, as he imagined the square "getting bigger and bigger." Boris' statement also reveals a consideration of variability in that he recognizes "the centre" and correctly predicts that "if you move away" from the centre, you get a bigger square." The "move away", and "the square is getting bigger and bigger" are examples of artifacts signs (Bartolini Bussi & Mariotti, 2008), in that the signs relate directly to the activity of dragging the points with the Dragging tool.

Boris' predictions focused more on moving data points "away" from the mean line. I asked Boris, "what if you move data points on the opposite side of the mean line?"

Boris looked at the sketch (Fig. 14) for a moment. Then he said, "Well, if you move data points", [*paused, his eyes still fixed on the sketch*], then he continued,

> [4] Boris: Oh yeah, yeah, on the other side, it doesn't matter as long as you move away from the centre, you will get bigger square […].

By statement [4], Boris meant that dragging a data point away from the mean line did not matter "as long as you move away from the centre" the square area will increase.

His prediction in [4] is consistent with the analysis in Figure 9 in that standard deviation increases as data points are dragged away from the mean line on either side. I asked Boris "[…] go ahead and check your prediction" and he using the Dragging tool, he moved data point **F** to the right side of the mean line. As the square area **F** increased (Fig. 19), he said:

> [5] Boris: Yeah, yeah, they are getting bigger, the boxes are just scales; they are actually just scales to see how the relative differences are. That was interesting.

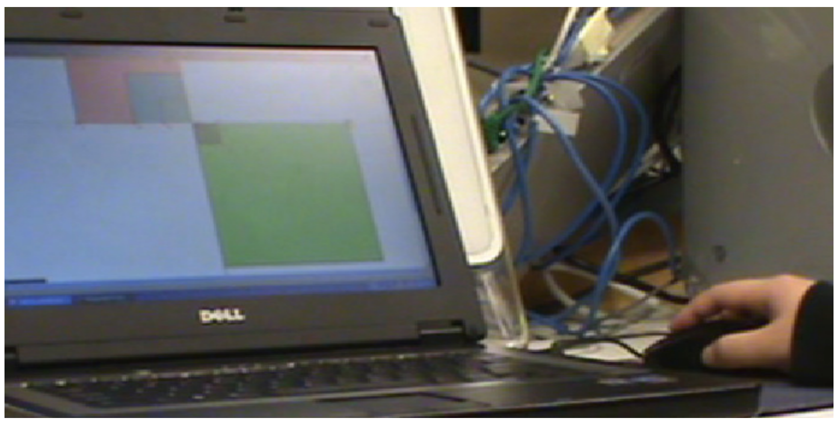


Figure 19. Boris dragging a point F farther away from the mean on the right

Given that Boris had correctly predicted the change in the squares in statement [4], his reaction in [5] only served to confirm his prediction. By "the boxes are scales", Boris suggested that the areas of the squares represented how far apart a given data point was from the mean line. A big square area meant that the data point was quite far away from the mean line, whereas a small square corresponded to a data point closer to the mean line. Boris' statement that "the boxes are scales" is interesting in that he evokes a new metaphor of weight (scales are used for measuring mass) to describe the relative differences in the spread of data points from the mean.

In terms of physical actions, Boris moved body, especially his hands a lot as he talked. He also produced artifact signs more frequently than he did mathematical signs. For example, the statement "if you move [data points] away from the centre, the square is getting bigger and bigger […]"in [3] is a set of artifact signs linked to using the Dragging tool to “move” or drag data points. It may be that artifact signs prompt more gestures than mathematical signs, which are often static and formal.

5.2.3. Activity with the gC Sketch

In Task 3, based on Figures 12 and 16, I asked Boris “to predict how the gC sketch” in Figure 16 “will change if you drag a data point on the horizontal axis.” He predicted that the gC sketch would be skewed on one side if he moved a point away from the mean line. However, if he moved the points closer to the centre, he went on, “the density will be more concentrated around the mean […]” [:6]:

> [6] Boris: Well if you drag a point, […] away from the mean, the curve will be skewed to that side [*moved his right hand his right side*] […]. As you move the points to the right, […] far away from the mean […] the curve will no longer be symmetric, instead it will be skewed on one side. If you move the points on that side [*moved his right hand to left side*], it will be skewed on that side. If the points are closer to the centre […] the density will be more concentrated around the mean […] you have like a sharp peak […]

It is not clear why Boris thought that the distribution would be skewed. Perhaps in the initial presentation of the gC sketch, the points are distributed more or less symmetrically along the number line, and the curve is also symmetric. He may thus have thought that moving one of the points would distort this distribution and then skew the curve. In fact, as I explained in Chapter 4, I assumed a normally distributed data set, but I had not told Boris this before asking for a prediction.

Boris’ prediction that “if the points are [moved] closer to the centre […], the density will be more concentrated around the mean […]” is one example of how standard deviation controls the shape of the curve, one main result that I was aiming for in the study. Although Boris did not mention standard deviation directly, his prediction suggested a connection with the magnitude of standard deviation. Boris was able to correctly link changes in the height of gC curve with the direction of movement of data points on the horizontal axis. I let him go ahead and check his prediction, “you can now test your predictions.” Boris used the Dragging tool and moved data point O (Fig. 16) on briefly to the right side and to his left, then back to his right. As he moved the point back and forth he said:

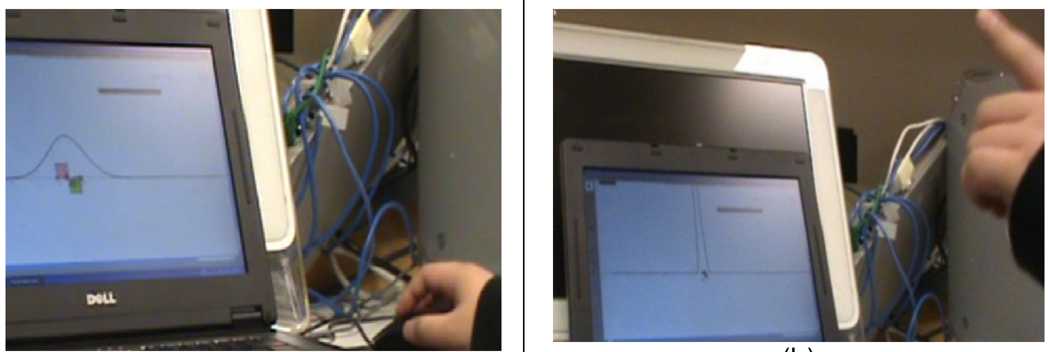
[7] Boris: Yeah, skewed to the left, move away […]

It was not clear if Boris actually observed a skewed distribution or he was recalling from his earlier prediction. But as he continued dragging the data point O toward the mean line the peak of the gC sketch continued rising (Fig. 20a), and Boris said, "if you move […] to the centre […] you get a sharp peak*."* Boris' last statement confirmed his prediction in [6] that, "if the points are closer to the centre […] the density will be more concentrated around the mean […] you have like a sharp peak." His statement also provides evidence that he used the Dragging tool as a semiotic mediator for connecting the rise in curve peak with the clustering of data points on the horizontal line.

I asked Boris, "What happens if you move all the data points closer to the mean line?" Boris did not answer immediately, but he dragged the data points one after another, toward the mean line. He said, "well, this (Fig. 20a) is not quite centre yet […]" as he continued dragging the points closer to the mean line. When he obtained the sketch shown in Figure 20b, he stopped dragging and said:

[9] Boris: […] if you move everything at the centre, it will be a degenerate distribution […]. Any points not at the centre will have a zero density. You have like a vertical line, all the mass are concentrated there

Boris did not answer my question but started dragging the data points perhaps to get answer to my question. It probable that he did not have the answer without first checking. After dragging the data points closer and obtaining Figure 20b, Boris talked about "a degenerate distribution", and explained that any data point not close enough to the centre "will have a zero density." His statement reveals the contribution of the Dragging tool that enabled Boris to easily and quickly produce different kinds of curves, including extreme examples as the one in Figure 20b.



(a) (b)

Figure 20. (a) Boris dragging data points on the gC sketch toward the mean line; (b) Boris pointing up at "the vertical line" with his left index finger.

The curve in Figure 20b also evoked a lot of mathematical signs from Boris compared with Anita, such as "degenerate distribution", "zero density" and "vertical line [with] all the mass concetrated there", which I had not anticipated in my design of tasks with the gC sketch. I attribute the generation of mathematical signs with the activity of dragging data points using the Dragging tool.

5.2.4. Reflecting on the Standard deviation

In Task 4, with the computer closed, I asked Boris, "What do you think about the term standard deviation?" Boris replied, as he moved his right hand to his right side, suggesting movement of points away on the right side of the mean line:

> [10] Boris: Standard deviation, as you move the points away from the mean, [*moved his right hand to his right side*], the standard deviation increases, that's what the graph shows. As you concentrate data at the center it gets less deviated, you get small values of standard deviation. If the data points are equal difference from each other, […] shift[*ting*] the data points to the left or right […] just shifts the mean, but it won't change the standard deviation.

Boris' statement [10] can be considered as a summary of the main constructs that he developed from his interactions with the sketches. First, the statement, "as you move the points away from the mean, the standard deviation increases" suggests dynamic thinking as well as a consideration of the functional connection among standard deviation, the mean and the changes in the data distribution with dragging. Second, the last three lines in statement [10], which I have labelled [10b], reads:

> [10b] Boris [...] If the data points are equal difference from each other, [...] shift[*ting*] the data points to the left or right [...] just shifts the mean, but it won't change the standard deviation.

Boris statement [10b] is still correct even if the data points are not placed at "equal" distances "from each other" on the horizontal axis. In fact, selecting all the six data points with a mouse pointer and dragging one of them on horizontal axis, will shift the mean, "but it won't change" the magnitude of "the standard deviation". In this example, the Dragging tool was used as a semiotic mediator for the functional connection between the mean and standard deviation as well as the data points. I argue that without dragging the data points and observing the patterns in the graph, it would have been more challenging for Boris to connect the patterns of change in the mean and the no change in standard deviation, with the distribution of data points on the horizontal axis. No other participant noticed this pattern on the sketches as Boris did.

5.2.5. Summary of Boris' Data Analysis

In Task 1, Boris was quite brief in his responses compared with Anita. However, Boris' statements in Task 1 provided generally correct statements about his consideration of variability compared with Anita. For instance, he considered standard deviation as a measure of "variation of data from the mean", a description which, in general, is consistent with Wild and Pfannkuch's (1999) as well as Konold and Higgins' (2003) aggregate reasoning. Boris' interactions with the dynamic sketches showed stronger and clearer considerations of variability than in the static environment in that he was able to connect the patterns of change in standard deviation with the dragging of data points [e.g. 10b]. After interacting with the dynamic sketches, Boris also showed more dynamic and physical thinking about standard deviation (e.g. in statement [10]) than he previously showed in the tasks. Boris' statements about the changes in different entities in the dynamic sketches (i.e., the standard deviation, the mean, the peak of the gC sketch, the dragging of the data points on the horizontal axis relative to the mean line) led me to think of extending the covariational reasoning perspectives proposed by

Carlson et al. (2002) as well as by Falcade et al. (2007). I will discuss the extension in Chapters 6 and 7.

5.3. Halen

Unlike Anita and Boris, Halen used a lot of drawings while solving Task 1. She also used gestures similar to those Anita and Boris made. In Section 5.3.1, I analyze Halen's transcripts in the static environment, followed by her interactions with the sketches in Sections 5.3.2 and 5.3.3. Section 5.3.4 provides Halen's reflections on the notion of standard deviation after the tasks. I provide a summary of analysis in Section 5.3.5.

5.3.1. Activity in the Static Environment

On Task 1, I asked Halen, "What comes to your mind when you hear the word distribution?" Like Anita, Halen said the "normal distribution" was "the first thing" that came to her mind:

[1] Halen: […] to me the first thing would be like the normal distribution […] and then it could be like other distributions, like […] skewed to one side, or to the other side.

For Halen, the first meaning of distribution "would be like the normal distribution" and "then other [...] distributions." Halen's image of distribution, like Anita, focused on one example, the normal distribution and did not include the spread of data values from the centre (Konold et al., 2003; Wild & Pfannkuch, 1999). To Halen and also Anita, the term distribution served as pointer (Konold & Higgins, 2003) to the normal distribution rather than how data is spread out from its centre. It may be that students pay more attention to the normal distribution than to distribution because it applies directly to their concerns such, as in grading their courses (e.g., [Anita, 2]). I asked Helen, "how about the mean?"

[2] Halen: Mean to me is like the average, like if you have a couple of numbers, you add them all up and then you divide by how many numbers there are, you get like the average or the mean.

Halen's thinking about the mean was associated with a formal mathematical process, "if you have a couple of numbers, you add them all and divide by how many you have." Her thinking of the mean was unlike the balance model proposed by Hardiman, Well, and Pollatsek's (1984). Hardiman et al.'s model considers the mean as a location of the centre of data, but the centre can change with the distribution of data. Lastly, I asked Halen about standard deviation, "what about standard deviation?" Halen's statement, like Anita, also suggested thinking about the Z-scores in the standard Normal curve.

[3] Halen: Standard deviation […] um, that is similar to the deviation […] If you have like this one right here [*she sketches the curve shown in* Fig. 21], like the standard deviation at the centre will be zero, and then […] you have in here like one standard deviation will be sixty eight percent […]. I forgot the numbers exactly but from here to here that will be the second standard deviation, and from here to here that will be the third standard deviation from the centre.

Halen sketched the normal distribution curve (Fig. 21) and explained that "[…] the standard deviation at the centre will be zero […] one standard deviation will be sixty eight percent […]". Halen's description of the so call "68-95-99.7" principle suggests that she recalled from her introductory statistics course. The "68-95-99.7" rule of thumb is a principle which relates the spread of data from the mean to the estimated percentage of the population that has been covered under a given investigation. For example two standard deviations away from the mean represent an estimated 95% coverage of the total population.

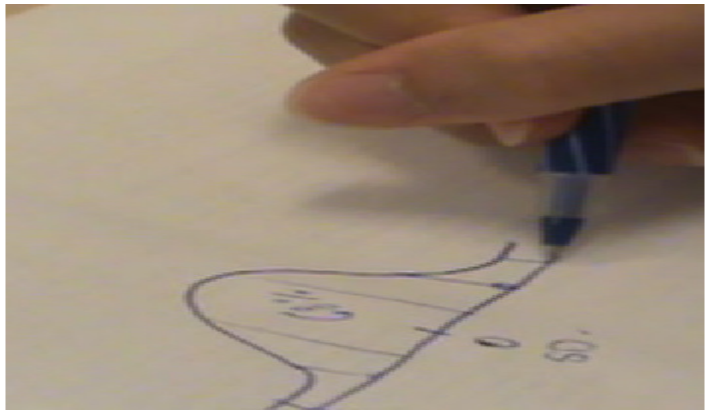


Figure 21. Halen sketch of the normal distribution curve

Although Halen's specified the centre of the distribution and the deviations from the centre in her sketch, she did not provide a clear image of the standard deviation. Her sketch instead focused on describing the Z-scores (also called the standard scores), which are measures of how many standard deviations are above or below the mean of a given data set. I expected Halen to relate standard deviation more generally to the measurement of variability or spread in a data set. Although Halen's drawing shows some considerations of how data vary from the mean, it is a special case of the Normal distribution curve. Moreover, her statement seemed to incorrectly suggest that distribution is the same thing as Normal distribution curve, and that the standard deviation is the same as the Z-score. According to Konold and Higgins (2003), Halen used the Normal distribution as a pointer to distribution and the Z-score as a pointer to standard deviation.

5.3.2. Activity with the dyMS Sketch

In Task 2, based on Figure 14, Halen was asked to "predict how the standard deviation, the mean and area of the squares change as you drag a point on the horizontal axis": Halen's prediction was not clear. It appears she imagined that the widths of the squares would change but not the heights if the data points were dragged on the horizontal line:

[4] Halen: Um, well if […] you were to stretch this way [*she move a data point away on the left side of the mean line*—Fig. 22], I think they

[*the squares*] will become […] a bit more narrow. […] as you go this way [*outward*] it becomes like skinnier and larger than this.

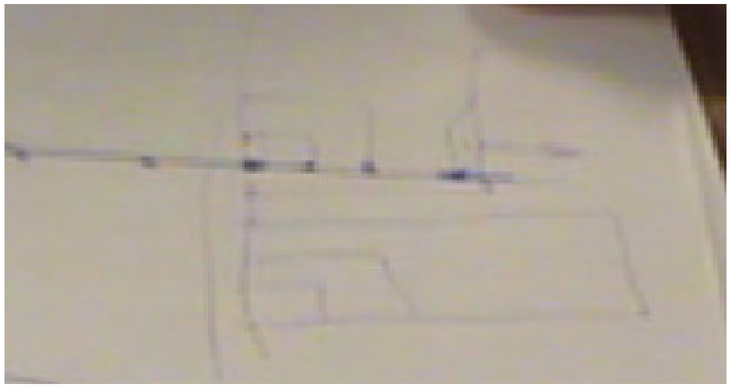


Figure 22. Halen's sketch showing the 'squares' looking "skinnier"

By the squares "become a bit more narrow as you go this way" Halen probably meant that the lengths of the squares would increase as the data points were dragged away from the mean line (Fig. 22), but not the heights of the squares. She seemed to be talking about the square stretching into a rectangle, which suggests that the squares were not functioning as signs for the size of the standard deviation. Indeed, the shape of the square does not change when the size of the standard deviation changes. I asked Halen, "What makes you think that the square will become more narrow […]?"

[5] Halen: Um, like when I look at it, it's just from imagination, like what I think will happen, just things going through my mind.

Halen imposed motion on the square even if the square looked static, predicting that the square "becomes […] skinnier" as a data point was dragged away from the mean line. Halen may have thought that the data point controlled the width of the square, instead of the area of the square.

I asked Halen to "go ahead and check" her prediction. She used the Dragging tool and selected data point C on the left side of the mean line (Fig. 14) and dragged it away from the mean line and back toward the mean line. As square area C increased and decreased (Fig. 23), Halen said "oh, that's so cool." "What do you mean?" I asked.

[6] Halen: Oh […] I thought like […] they would all go together […] if I did this [*she dragged data point C slightly to her left*], all of them would go together at the same time, but oh, ok, […] so, it affects both sides.

Halen's statement in [6] suggests some surprise at how the square areas changed. She noticed that dragging a data point on the horizontal axis "affects" the squares on "both sides" of the square. Her statement "yeah, so it affects both sides" revealed a change in her awareness considering her earlier prediction in [4] which incorrectly suggested the square " becomes like skinnier and larger" as a data point was dragged away from the mean line.

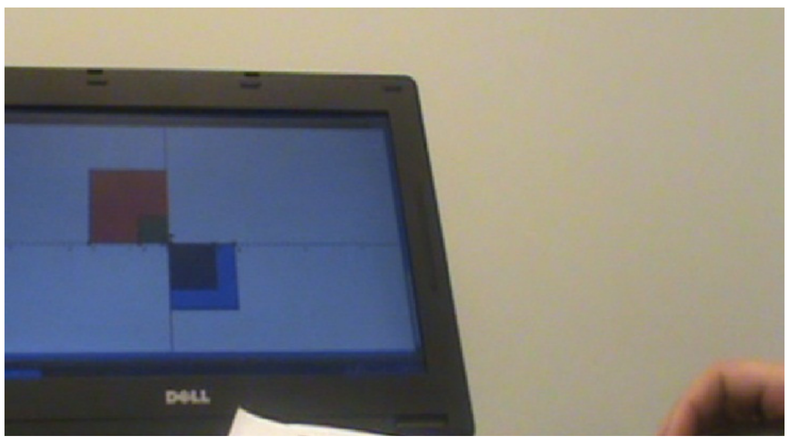


Figure 23. Halen drags data point C to the left side of mean line with the Dragging tool.

Statement in [6] suggests that Halen used the Dragging tool as a semiotic mediator for the movement of data points on the horizontal axis and connected the changes in the areas of the squares with dragging. Yet, the statement "oh, that's so cool" suggests affective dimension of Halen's interaction with the sketch. Overall, Halen's statements from her activity with the dyMS were different from those coming from the tasks she solved in the static environment. In fact in the static environment, Halen relied more on the formal procedures such as the "68-95-99.7" rule, as well as the algorithm for calculating the value of the mean in a set of numbers. However, with the dynamic sketch, Halen seem to do more exploring and discovering patterns by herself and her statements were more physical and dynamic in a sense. For example, statement [6], "I thought […] they would all go together […] if I did this […]."

5.3.3. Activity with the gC Sketch

In Task 3, using Figure 24, I asked Halen to “predict how the height of the gC sketch would change as you drag the data points along the horizontal axis?” Halen predicted that dragging the points away from the mean line “will flatten up the graph […]”

> [7] Halen: Ok, keeping the other one in mind [i.e., *Task 2 with the dyMS sketch*], well if I did [*drag*], let’s say, I did A [Fig. 25a], then […] um, I think if we pull data at A this way [*to the left side,* Fig. 25b], [*pause*], I think it will flatten up the graph. And if you did this one [*on right side of mean line*], the same thing, the line will move this way [*to the right side*] and the graph will become narrower.

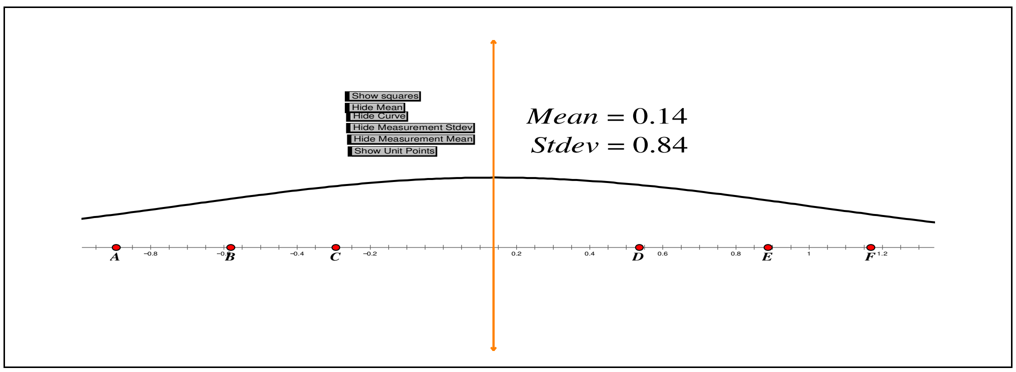
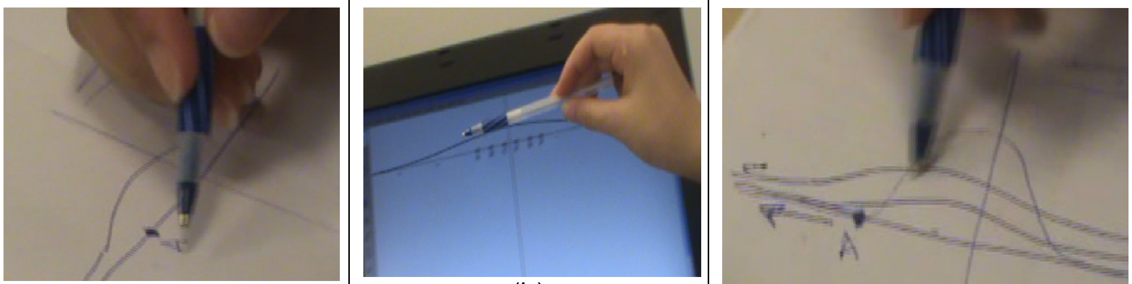


Figure 24. The gC Sketch.

Halen sketched Figure 25a and later used the pen to trace the curve (Fig. 25b) on the computer to confirm her prediction that dragging data point A to the left side “will flatten the graph.” By “flatten up the graph” Halen probably meant that ‘height of the gC sketch would go down' if the data points were dragged away from the mean line. Based on Halen’s opening remark, “ok, keeping the other one in mind […]”, there is reason to believe that her correct predictions on the gC sketch were mediated by her previous interactions with the dyMS sketch given that some entities, such as the data points looked similar on both sketches.



(a) (b) (c)

Figure 25. (a) and (b): Halen showing that dragging a data point to the left side of the mean line "will flatten up the graph"; (c) Halen traces the curve with the highest peak after gradually dragging the data points closer to the mean line.

I asked Halen, "what if you move all the data points […] closer to the mean line?" She said that "[…] maybe the peak will be higher […] if you crush them altogether" and sketched a curve (Fig. 25c) with the height of the curve gradually rising as data points were progressively dragged closer and closer to the mean line.

> [8] Halen: […] um, I think when you push them in the centre […], I think [the mean] line might deviate a little depending on what point you start at But then I guess [...] maybe the peak will be higher […] if you crush them altogether, maybe they will just be like this [Fig. 25c].

Helen traced with a pen, the final curve with the highest peak after dragging data points closer and closer to the mean line (see Fig. 25c). I probed her, "how do you know the curve peak will rise?" She said:

> [9] Halen: I don't know, I just think that when you push them [*data points*] together, the peak will become higher.

Although Halen stated in [9] that she did not "know" why the curve would rise, her sketch in Figure 25c seemed to convey her correct thinking—a change in the height of the curve as the data points were dragged closer to the mean line—(i.e. as standard deviation decreased, the height of gC sketch rose). I conjectured that Halen linked the physical action of 'crushing' [8] and 'pushing [*data points*] together' [9] to the rise in the peak of gC curve.

Halen's statement in [9] also showed some evidence of covariational reasoning—linking the rise in the peak of the curve ("the peak will be higher") with change in the direction of moving the data points inward ("if you crush them [points] altogether").

I asked Halen to "[…] go ahead and test your prediction." She used the Dragging tool and moved data point A on then left side of the mean line (Fig. 24) slowly to the far left and as the peak of the gC sketch decreased, Halen said, "ok, yeah, I was mostly right, the line moves that way." By "the line moves that way", Halen probably meant that the height of the gC curve dropped as she dragged the data point **A** farther away from the mean line. If that is what she meant, then Halen's last statement shows some evidence that the Dragging tool served her as a semiotic mediator to connect the change in the height the curve and the spread of data on the horizontal axis.

5.3.4. Reflecting on the Standard deviation

In the static environment, Halen's construct of standard deviation pointed to the Z-scores in the standard Normal curve. She also did not connect standard deviation, the mean, and data distribution in a functionally related way. In Task 4, and with the computer now closed, I asked Halen "do you […] see any links between the *standard deviation* and the *mean*?"

> [10] Halen: Well, if you change standard deviation, the mean is going to change. When I was looking at the graphs, I didn't realize that, like […] the red line [*mean line*], the red line like [...] if you move the points together [*she moved her hands closer together*—Fig. 26], the mean is going to change too.

Halen seemed to recall from her interactions with the sketches that "if you change standard deviation", the mean "will change." Although her statement showed evidence of covariational reasoning, the word "change" was not explicit. Moreover, although the mean and the standard deviation both change, there is no covariational relation between them (see Fig. 9). For example, as discussed before, moving a data point to the left side of the mean line causes an increase in standard deviation, but a decrease in the mean. However, from Halen's hand movement in Figure 26, it could be that by "change", she meant an increase in the height of the gC curve, but she was not explicit about what she meant. Her statement in [10], "when I was looking at the graph, I did not realize that if you move the points together, the mean is going to change too]" suggests some new evidence that she now became aware of the connection between the increase height of gC curve and the change in magnitude of standard deviation. I argue that Halen used the Dragging as a semiotic mediator for the meaning of the rise in the height of the gC and the distribution of data points on the horizontal axis.

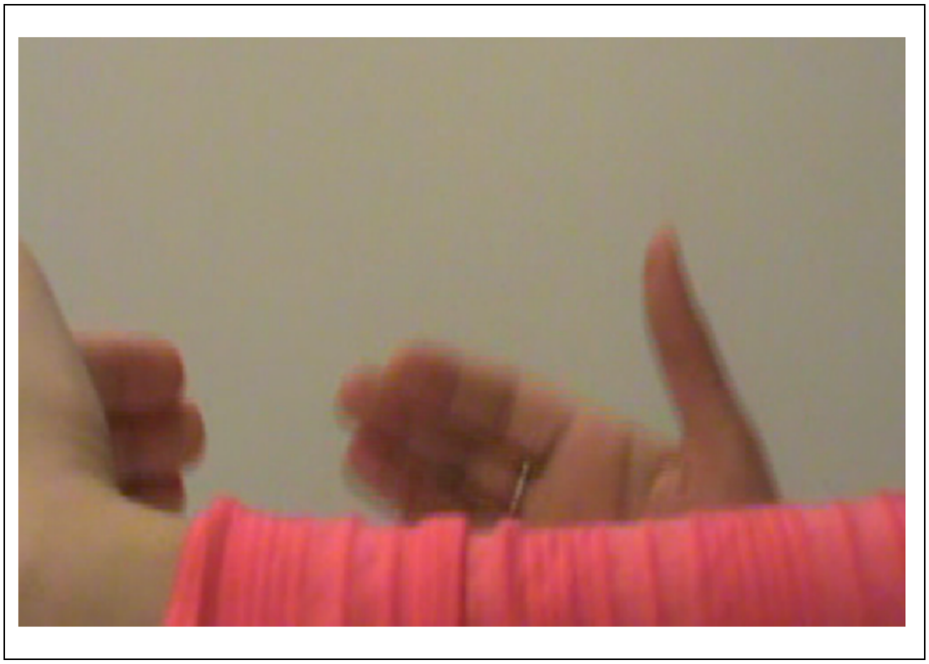


Figure 26. If you move the points together the mean is going change

5.3.5. Summary of Halen's Data Analysis

In the static environment, Halen did not clearly state the meaning of the terms distribution, mean and standard deviation in an aggregate way. Rather, she used the concept of distribution as pointer to the normal distribution. Halen's thinking about the mean also relied on the procedure of obtaining an average of numbers by adding them up and dividing by the total number. However, during and after her interactions with the dynamic sketches, Halen showed some considerations of variability. For example, her predictions about the changes in the gC sketch were generally correct (statements [8], [9], & [10]) although not as explicit and detailed as Boris' were. Nevertheless, she seemed to make generally correct but obvious mathematical statements such as "if you change standard deviation, the mean is going to change" and "if you move [data points] together the mean is going change too” which are examples of mathematical signs (Bartolini Bussi & Mariotti, 2008). Thus, it can be argued, more generally that the Dragging tool enabled Halen to construct the mathematical meaning of standard deviation as an important index of variability.

However, compared to Anita and Boris, Halen was not explicit about her use of the word “change.” In statement [10] she says, “If you ‘change’ standard deviation, the mean is going to ‘change’ too” but does not clarify the “change.” In fact as Boris' example [Boris, 10] shows, there are conditions in which the mean can "change", but standard deviation does not change through dragging the data points. Halen’s statement excludes such conditions, and focuses on the more general change when data points are dragged on the right side of the mean line.

5.4. Maya

Maya used gestures, especially hand movements quite often when responding to questions. In the static environment, like Anita and Halen, Maya also showed similar thinking, for example, the use of “formula” to “calculate out” the mean. I first analyze Maya’s transcripts in the static environment and then her interactions with the dynamic sketches.

5.4.1. Activity in the Static Environment

In Task 1, I asked Maya to “describe the term distribution”:

[1] Maya: Distribution is just how things are spread out, they can be evenly distributed or they can be randomly distributed, just the placing of values or data, or objects.

Although Maya’s image of distribution did not specifically mention elements such as shape, centre and deviation from the centre, which would conform to Konold and Higgins’s (2003) aggregate reasoning, in general, her reference to “how things are spread out” probably suggests some aggregate reasoning. In that case Maya’s image of distribution carries some element of consideration of variability (Wild and Pfannkuch, 1999).

I asked Maya, “how about the mean?” Like Anita and Halen, Maya’s thinking about the mean seemed to show more about "the answer to a formula […]". She said:

> [2] Maya: The mean is the answer to a formula where we add up all the values in a particular data set and divide by the number of values that are there, so mean is like a number, like a specific number, [the] mean is more specific, it’s like calculated out.

By the statement, “so the mean is like a number, like a specific number, [the] mean is more specific, it’s like calculated out”, Maya seemed to emphasize calculating the mean, an image which was not different than Anita and Halen’s thinking about the mean. Her statement confirms Pollatsek, Lima, & Well’s (1981) proposition that for many students, dealing with the mean is more of a calculation than a conceptual undertaking.

I asked Maya, “what about standard deviation?”. Like Anita and Halen, Maya suggested the the Normal curve although she did not explicitly say so:

> [3] Maya: Standard deviation? Um, standard deviation, I see standard deviation in *graphs*, there is like one, two, three, four; then there’s negative one, negative two, negative three, negative four. You can calculate standard deviation.

Drawing on a similar statement from Halen, Maya’s statement “[…] I see standard deviation in graphs [...] there is like one, two, three […], then […] negative one, two […]” strongly linked standard deviation with z-values found in standardized Normal curve. Her thinking also seemed to inform Halen’s drawing shown on the sketch in Figure 21. However, when I asked Maya “which graphs do you see standard deviation in […]?” she was not sure about the specific graph. Seemingly, the term “Normal curve” had disappeared from Maya’s immediate recall,

> [4] Maya: It’s like a graph and […] there is the middle and then there is standard deviation, right? It separates, there is a graph that separates the standard deviation into measures, is that standard deviation or is that central tendency?

Maya uses the terms “standard deviation” and “central tendency” with no clear connection between them. I assumed that Maya was looking for the term the Normal distribution curve given that her classmates, Anita and Halen had showed the same image. By “it’s like a graph […] there is the middle and then there is standard deviation […]”, I believed Maya was thinking about the Normal curve, but she had forgotten the name. I probed, “what sort of graph?”

> [5] Maya: I forgot […] a graph like a regular, um, I forgot what that graph is called, it’s a hill, right? Yeah, a graph that has a hill so um.

Maya’s statement “a graph that has a hill” [5] gave me more reason to believe that she was thinking about the Normal distribution curve. I then offered her a hint, “You mean the bell, bell curve?” Maya immediately responded, “bell shape curve, bell shape graph?” But she left a question mark that made me believe she meant another word for the “bell curve.” Then I suggested the “Normal curve?” which she responded to by repeating the “Normal curve” four times:

> [6] Maya: Normal curve, Normal curve, so anyways, there is, I see a graph, and there is a Normal curve and then the Normal curve has been separated into like eight sections. Is that standard deviation? I think that is what it is?

Like Anita and Halen, Maya seemed unsure of the distinction between standard deviation and the Z-scores. There was a general misrepresentation of standard deviation by my participants, as the Z-scores and distribution as the normal distribution curve. I expected participants to be aware that the Z-scores are the standardized values of data points that measure how many standard deviations a given point is above or below the mean value. The standard deviation in general, is a feature of how spread out or how clustered a distribution is from the centre. Thus, Maya’s image of “a graph that separates the standard deviation into measures” did not explicitly describe what the standard deviation is or what it does, rather, it described one of the applications of standard deviation, namely in the standard normal curve, as Halen had sketched in Figure 21.

I probed Maya more about standard deviation, “so, for you standard deviation is…?” Maya was quiet for a moment, and then she said:

> [7] Maya: Standard deviation is how far away from the mean a data point is, it is like a category, it categorizes how far away from the mean a data [point] is.

Maya’s statement [7] shows that she revised her thinking and provided a clearer understanding of standard deviation than her statement [6]. I believe it was through questioning and probing that evoked Maya's more correct thinking of standard deviation. I conjectured that questioning, probing as well as prompting can help students re-construct concepts that they have already learned, but which they may have forgotten. Bartolini Bussi and Mariotti (2008) have discussed the role of the teacher as a cultural mediator in assisting to learn and connect meanings between concepts. The role of a teacher as a cultural mediator in students' learning, however, is not the focus of the current study. I will now proceed to analyze Maya's interactions with the dynamic sketches in the computer environment.

5.4.2. Activity with the dyMS Sketch

In Task 2, based on Figures 11 and 14, I prompted Maya to "predict how the squares will change if you drag the data points on the horizontal axis."

> [8] Maya: I think when I move the points, the lines that the points are connected to will also move

Maya's statement [8] was not clear, but she probably meant that the 'vertical side of the square "will also move" if a data point was dragged. Her statement was obviously correct, but it did not offer much understanding of the features of variability in the curve. It may be that my original question was not focused, so I asked Maya more specifically about the changes in the square, "How do the squares change as you move the data points on the horizontal axis to the left or to the right of the mean line?"

> [9] Maya: Well, I guess when I move the points to the left, the square will increase.
>
> [10] Int: Why is that?
>
> [11] Maya: Because the farther away the point is from the centre, then the greater area it has.

Maya correctly defended her prediction in [9] by her statement in [11], "because the farther away the point is […] the greater area it has." It seems the distance from the centre and the change in the area of square were signs that mediated Maya's understanding of the patterns in sketches. However, Maya's use of the signs was at the artefact level, but not yet at the mathematical signs in that she did not link the change in the area of the square with the change in the magnitude of standard deviation. I asked Maya about the direction of moving the data points:

[12] Int: Ok, so your claim is that when you move the points to the left side of the mean line, the square area will increase. What if you move the points to the right side?

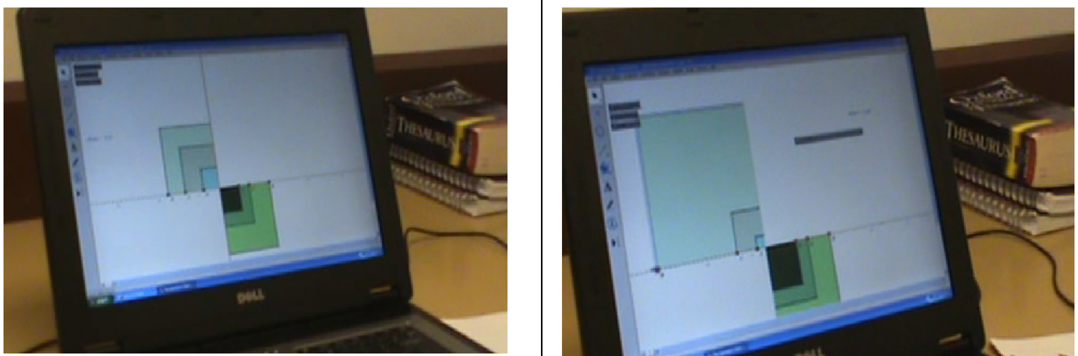
[13] Maya: To the left, the area will increase and to the right then the square area will decrease.

Maya's prediction "to the left the area will increase and to the right then the square area will decrease" was correct but only if the dragging activity was restricted to the left side of the mean line and away from it, as the framework in Figure 9 shows. If however, the dragging action was on the right side of the mean line, Maya's statement would not stand because both standard deviation and the mean increase with data points dragged away from the mean line. I asked Maya to "[…] go ahead and check" her predictions.

Maya used the Dragging tool and slowly moved data point B (Fig. 14) away to left side of the mean line. She continued dragging the point back and forth on the horizontal axis (Fig. 27b) and said:

[14] Maya: So, the mean increases as standard deviation […]. As standard deviation increases, the mean also increases. Oh no, the mean decreases right? […] Oh, this is nice.

Maya's initial statement in [14], "As standard deviation increases, the mean also increases" did not account for the direction of dragging the data points on the left side of the mean line.



(a) (b)

Figure 27. (a) The dyMS sketch before Maya dragged the points; (b) The dyMS sketch after Maya dragged a data point to the left of the mean line.

Maya continued dragging the data point B on the horizontal axis, but on the left side of the mean line. She was able to notice that "as standard deviation increases, […] the mean decreases" which statement provides evidence that the Dragging tool mediated the change in the standard deviation and the mean as Maya dragged data point B away on the left side. Her remark "oh, this is nice" at the end of [14] suggests the affective dimension of Maya's interactions with the sketch. It seems that Maya was pleased with what she noticed on the sketch.

5.4.3. Activity with the gC Sketch

In Task 3, based on Figures 12 and 16, I prompted Maya to "predict how the gC sketch will change if you drag the points on the horizontal axis."

> [15] Maya: […] I guess as I move the points, the purple line [*curve*] will also rise. So, when the standard deviation increases, the curve will also increase, sort of rise.

Maya predicted that "[…] when the standard deviation increases, the curve will also increase [...]" but she did not specify the direction of dragging the data points that would cause such a change. By the "purple line will also rise", Maya suggested that the height of the gC sketch would rise. Maya's prediction [in 15] that "when standard deviation increases, the peak of the gC sketch "will also increase" seemed to incorrectly suggest that the height of the gC curve co-vary with standard deviation. Maya's statement [15] supports my design hypothesis that some students may misinterpret the increase in the magnitude of standard deviation and link it to the increase in the height of gC curve, and vice versa. I challenged Maya,

[16]. Int: Since we now know that standard deviation increases as you move data points away from the mean line, do you mean that the curve peak will increase as you move data points away from the mean line?

Maya paused for a moment while looking at the sketch (Fig. 16). Then she predicted more correctly that moving data points away from the centre will cause the height of the gC curve to "plunge" or "go down":

[17] Maya: When I move the [data] points away from the centre, the curve will plunge, go down [*spread her hands out and briefly moved them down and up*—Fig. 28]; but when I move points closer to the centre, the curve will rise?

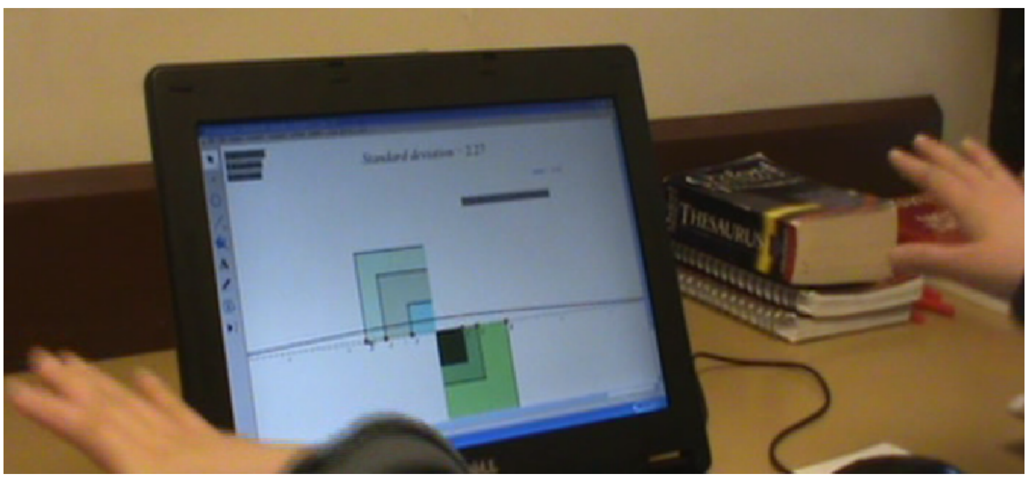
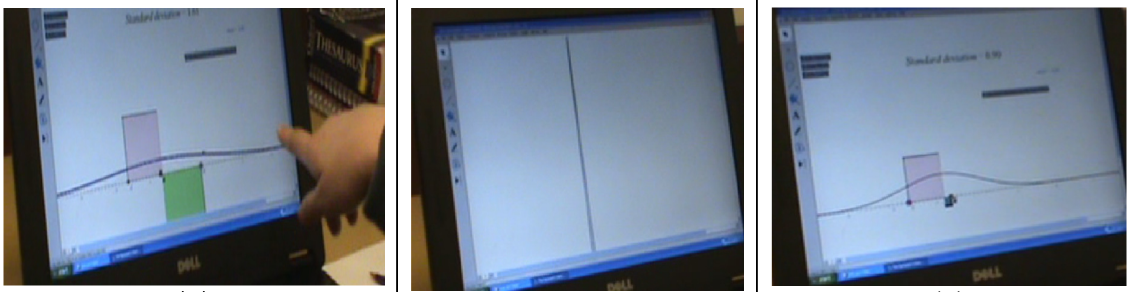


Figure 28. Maya spreading her arms across the screen and saying the curve peak will go down

As Maya said, "[…] the curve will plunge […]", she spread her arms across the computer screen (Fig. 28), and moved them down, maybe internalizing how the height of gC curve would "go down" as data points were dragged away "from the centre". It seems that statement [16] mediated Maya's more correct prediction in [17]. According to Vygotsky, interpersonal communication can lead to intrapersonal communication when an individual masters the task and is able to perform by her or by himself. Maya had also appealed to her own intellect (Vygotsky, 1978) and revised her predictions in Task 2, after similar interpersonal communication. It seems the sketch in Figure 28 also evoked a more physical and dynamic thinking by Maya, considering her use of the active verbs such as "*move* points away", "the curve will *plunge*", "but when I *move* points closer", "the curve will rise" in [17]. She also interiorized her statements by physical action, for example, spreading her arms across the screen in Figure 28.

I asked Maya to "[…] go ahead and check your predictions." Maya moved data point **O** (Fig. 16) on the right of the mean line, farther to right side and back toward the mean line. As she dragged the point back and forth, the height of the gC sketch decreased and increased alternately. Maya continued dragging the same point slowly back and forth and said "opposite, opposite, oh, opposite […] to what I got. It's like a hill […]." She drew a curve in the air with her right index finger (Fig. 29a) as she said "It's like a hill."

By "opposite […] to what I got […]" Maya seemed to notice something different in the gC curve that her prediction did not correctly state. In particular, she probably noticed, contrary to what she predicted that, by dragging the data points toward the mean line, she obtained something that looked "like a hill." I asked her to "drag all the points closer to the centre and describe what happens" My aim was to check her thinking about the link between standard deviation and the change in the height of the gC sketch. Maya slowly dragged all the six data points one after another, closer and closer to the mean line and as she dragged the points much closer together, the sketch in Figure 29b evolved and Maya she said "Wow, so big!"



(a) (b) (c)

Figure 29. **(a) Maya pointing on the rising peak of the gC sketch with her right index (b) Maya dragged the points much closer to the mean line and obtained the sketch looking like a 'stick' (c) Maya restored the sketch (in b) to a more normal size.**

By "wow, so big", Maya probably meant that the peak of the gC sketch was "very high". She also may have been surprised by the sudden change in the shape of the curve, which she probably had not seen before. Maya used her right index finger to draw a picture of the gC sketch with a lower peak in the air as she said "I thought the curve would just be like that." Her reaction to the sudden change in the height of the gC curve was different, for instance, than Boris, who correctly predicted how the height of the gC would change, and he seemed to showed less surprise at the result he obtained after checking. Maya asked if she could "see up" suggesting seeing the "tip" of the sketch in Figure 29b,

[19] Maya: Can I see up? Can I move the graph up?

I answered, "Yes, you can", but was not convinced that I gave Maya a satisfactory answer to her question. However, I anticipated that later she would discover the answer by herself. Maya scrolled up the page using the mouse pointer for about half a minute without getting near to the tip of the sketch. Then she asked "Is this forever, does it go on forever?" I did not reply. Instead, I posed a question to her, "What should we do to restore the sketch to its normal shape? Maya did not respond to my question but she kept looking at the gC sketch on the screen. After some moments of silence, I asked Maya to "scroll back to the horizontal axis and drag the data points apart". She scrolled to the horizontal axis, dragged some data points apart and was able to restore the peak to a more normal height (Fig. 29c). When she had finished dragging some data points apart, and restored the curve, she answered my earlier question, saying, “I just separate the points." I followed up Maya’s answer with a general comment: “Yes, and that tells you a lot about your data, how your data set is varying”, to which Maya responded:

> [20] Maya: Because standard deviation is like a measure of how far apart the points are from the mean […].

Maya correctly linked the changes she had observed in the height of the gC sketch to the standard deviation as “[…] a measure of how far apart the points are from the mean” [20]. Her statement showed that she used the Dragging tool (along with the sign of the gC curve) as semiotic mediator to construct the meaning of standard deviation, different than she originally did at the beginning of the interview through "deriving" standard deviation. Maya’s statement can also be considered as a mathematical sign (Bartolini Bussi & Mariotti, 2008). Her statement satisfies Wild and Pfannkuch’s (1999) aggregate reasoning with graphs in that she connects the centre of data with the spread of the distribution from its centre.

Furthermore, when I remarked, “[...] when the data set is spread out in a certain way, the gC sketch also behaves in a certain way.” Maya responded:

> [21] Maya: When the data points are farther away from the centre, then the Normal curve will fall, and when the points are towards the centre, the Normal curve will rise.

At the beginning of Task 3, Maya had a hard time predicting how the height of the gC sketch would vary with changes in the magnitude of the standard deviation. Based on Maya statements [20, 21] after interacting with the sketches, I propose that the dragging action enabled her to connect the changes in the magnitude of the standard deviation with changes in the height of the gC sketch. I argue that Maya used the Dragging tool as an instrument of semiotic mediation for the mathematical meaning of standard deviation and its connection to the height of the gC sketch.

Maya did not get to the final result in statements [20 & 21] all by herself, but initially she obtained some partial assistance from me, acting as a cultural mediator. However, as Maya became more aware of her task, partly through our interpersonal communication (e.g. probing, questioning, predicting, and recalling some known ideas), and also by her own action with the Dragging tool, Maya seemed to stand on her own and make correct mathematical statements, for example, “because standard deviation is like a measure of how far apart the points are from the mean […].” Maya’s episode validates the Vygotskian hypothesis that meaningful use of a cultural tool includes interpersonal interactions, as Vygotsky claims, before intrapersonal interaction (or internalization). I argue that Maya became aware of the mathematics involved in the task (i.e. the link between standard deviation and the change in the height of the gC sketch) by utilizing the Dragging tool as an instrument of semiotic mediation as well as the interpersonal communication (or language) which got transformed into intrapersonal communication for Maya (Vygotsky,1978).

5.4.4. Reflecting on the Standard deviation

In Task 4, (Fig. 14) and with the computer closed, I prompted Maya, “what do you think about the term standard deviation?” Maya’s said:

[22] Maya: Standard deviation is certain point away from the centre of a population and there is negative and positive standard deviation. I have a picture of a normal distribution divided into sections, which are called standard deviation and those sections are not equal unless they have the same positive and negative value. There is a formula, I forgot but it’s like standard deviation equals the square root of the variance.

Maya’s thinking of standard deviation as “a certain point away from the centre of the population” was probably linked to the spread of data points from mean, which would generally go for correct statement in that her statement recognizes the “centre of a population”, which in general, agrees with Wild and Pfannkuch’s (1999) as well as Konold et al.’s (2003) perspectives of aggregate reasoning. However, Maya’s “picture of a normal distribution divided into sections” confuses the construct of the standard deviation with the Z-scores. The Z-score is used in the standard normal curve to provide information on how many standard deviations above or below the population mean, a given data point is. The farther away a data value is from the mean of the population, the less likely that the overall impact of that data point on the whole population can be ignored, or will simply happen by chance. Based on her statements after interacting with the sketches, I described Maya’s thinking about variability as mixed (i.e. showing both static and aggregate) consideration of variability. It is not clear why Maya suddenly fell back to the static mode of thinking, having shown very clear mathematical meaning of standard deviation in statements [20-21]. Maya's mixed consideration of variability suggests that the dynamic sketches were important in evoking her thinking of variability; but without the sketches, Maya seemed to fall back to the static mode of thinking. It may be that Maya, as was the case with Halen, needed some more time to interact with the dynamic sketches, to enable her attain a more stable understanding of the features of variability.

5.4.5. Summary of Maya’s Data Analysis

In the static environment, Maya provided a relatively clear image of the term distribution compared to Anita and Halen. However, Maya's thinking about the mean and standard deviation was similar to Anita and Halen's in the sense that, for the mean, Maya also relied on the processs of adding numbers up and dividing by the total number present in the data set, an image that supports Pollatsek et al.’s (1981) finding that many students consider the mean more as a number rather than an important concept for reasoning about the centre of data and the distribution of data points about the mean.

Moreover, during her interactions with the dynamic sketches, Maya initially had challenges predicting changes in the sketches, but after using the sketches, she showed clearer awareness of the connections among standard deviation, the mean and distribution. I suggest that Maya’s interaction with the Dragging tool occasioned her consideration of the features of variability. However, after using the sketches and having the sketches put away from her, Maya seemed to fall back to her original thinking of distribution in the static environment. She also used standard deviation similarly as a pointer (Konold & Higgins, 2003) to the Z-scores in the standard Normal curve. I conjectured, based on previous studies (e.g. delMas & Liu, 2005; Konold & Higgins, 2003), that students who consider distribution as a pointer to the Normal distribution curve are more likely to use standard deviation as a pointer to with the Z-scores. Moreover, such students are more likely to fall back to the static thinking about variability when they are not actively engaging with the dynamic tool.

5.5. Yuro

Yuro was also quite brief in answering some questions in the interview, similar to Boris. Like all other participants, Yuro also moved his hands as he responded to the interview tasks. In the first section, I analyze Yuro's transcripts from the tasks he performed in the static environment. In Sections 5.5.2 and 5.5.3, I analyze episodes of Yuro's interactions with the dynamic sketches. Sections 5.5.4 and 5.5.5 are respectively Yuro's reflections on the term standard deviation after the tasks, and a summary of my analysis.

5.5.1. Activity in the Static Environment

In Task 1, I asked Yuro, "What comes to your mind when you hear the word distribution?" Based on similar responses that I received from his classmates (e.g. [Halen, 3]), Yuro also seemed to consider distribution in terms of standard deviation and the normal distribution curve:

> [1] Yuro: Standard deviations, how far are um, how far […] points are […] around the dots […] one point on the left, one point on the right and then one point in the middle and then all these random points around it.

Although there is link between the "distribution" of a data set and the "standard deviation", Yuro's image of distribution in this particular question did not present a clear connection between the two constructs. When Yuro describes "one point on the left, and one point on the right and then one point in the middle", he may be referring to the units of standard deviation from the mean that he would have seen in class, which are usually represented on the normal curve (as was seen also in Halen's sketch).

When I asked, "What comes to your mind when you hear the term mean?" Yuro responded in one word: "Average", and did not elaborate his answer. However, by "average", I assumed that Yuro meant adding up numbers and dividing the sum by the total number in the data set, given that his course mates, Anita and Halen had already given similar statements. In fact, all participants provided a similar response for the mean. On the last question in Task 1, I asked Yuro, "What comes to your mind when you hear the term standard deviation?"

[2] Yuro: Spread […] same things, there's standard deviation and deviation […] same things to me.

Yuro's thinking of standard deviation as "spread" was correct, but not completely developed in that he did not include the construct of "centre" of the distribution of a data set. Yuro also seemed unsure about the meanings of "deviation" and "standard deviation". To him "standard deviation and deviation" seemed to mean the "same things." However, I expected Yuro to consider deviation as the distance of a single data point from the centre of data, whereas standard deviation accounts for the entire data set as discussed in Chapter 4. It may be that students find the distinction between deviation and standard deviation unclear because deviation is not as much talked about in statistics courses as standard deviation.

5.5.2. Activity with the dyMS Sketch

In Task 2, and using Figure 14, I prompted Yuro to predict "how the squares will change as you drag the data points on the horizontal axis." Yuro predicted that dragging a data point toward the mean line would cause the corresponding square "to be smaller." I asked Yuro "Why?" and he replied:

[3] Yuro: Um, why? Oh yeah, I just think it's going to move this way […]. I just think this is going to move this way […] ah, what should I say ok, I'm going to move this way […], the whole thing is going to go this way.

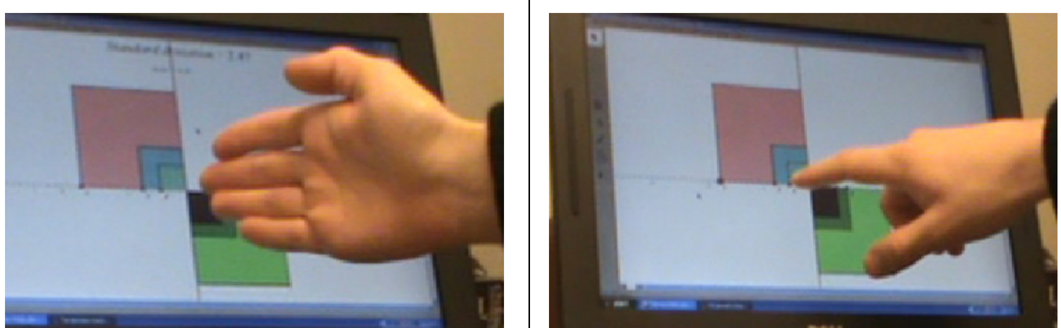
Yuro did not give reason "why" the square was going "to be smaller". Rather, he focused on the dynamic movement of a data point, "it's going to move this way", "this is going to move this way", "I'm going to move this way", and "the whole thing is going to go this way". In fact, Yuro's was not specific about what exactly was "going to move this way". I restated the question more specifically, "[…] If you moved a data point to the left or to the right side of the mean line, how would the squares change?" Yuro replied:

[4] Yuro: Oh, if you move it this way [he *move*s *his left hand to* the *left side*] it's going to increase, and if you move it this way [*move*s *his right hand to the right*], it's going to decrease.

Yuro accompanied his predictions with hand movements, to the left and to the right as he said "this way it's going to increase [moves his *left hand to the left side*] and this way it’s going to decrease” [*moves his right hand to the right-side*]. When I probed, “Why is that?” Yuro more or less restated his prediction in [4]:

> [5] Yuro: Because over here [*stretches his hand on the screen near the mean line*—Fig.30a] this is kind of like a central point, and then these are distributions around the central point, and if you are going to move this one, your distribution is going to increase.

By “[…] your distribution is going to increase” Yuro was not clear but he probably meant that the square area was going to increase if a data point was dragged away to the left side of the mean line.



(a) (b)

Figure 30 (a) Yuro stretches his right hand to the screen to show the distribution around the mean line. (b)Yuro points the horizontal length of a square, which he believed would ‘increase’ if data points were dragged away from the centre.

I asked Yuro to click on the buttons for standard deviation and the mean to enable him observe how the scales changed as he dragged the data points on the horizontal axis. After Yuro clicked on the buttons with the mouse pointer, I asked him to “[…] go ahead and check your predictions.” Yuro dragged data point **B** on the left side (Fig. 14), away from the mean line and then back toward the mean line as the area of square B increased and decreased. As he continued dragging point B back and forth and noticing the changes on the sketch, Yuro said “Oh, I was wrong.” I followed up his remark, “why” he thought he “was wrong.” Yuro explained as he pointed to the left side of the mean line (Fig. 30b), that "he felt" only the horizontal length of the squares increase, but not the height.

[6] Yuro: Because I only felt that this [*put his index finger on the horizontal line*—Fig. 30b] was going to increase but not this one [*vertical length of the square*]. But now that I think about it, it makes sense because […]. I thought […] the heights would be constant […].

Yuro's statement, "but now that I think about it, makes sense […]" can be considered as evidence of his considerations of variability, which I suggest, was occasioned by the dragging action and by his coordinating the signs produced in dragging the data points. The Dragging tool made it possible for Yuro to "make sense" of (or internalize) the changes in the squares as he dragged the data points on the horizontal axis. However, Yuro's initial "wrong" thinking about the square is not an isolated case in the sense that Halen also had similar difficulty (see Fig. 22), with the horizontal length of the square increasing, but not the height. It is not clear why Halen and Yuro thought about the square that way, but I believe it is challenging for some students who, in general, learn concepts using static tools, to imagine a square dynamically increasing or decreasing in area without its losing its mathematical "squareness"—i.e. maintaining the same width and height on all four sides, at the four right angles.

5.5.3. Activity with the gC Sketch

In Task 3, I asked Yuro to "predict how the peak of the gC sketch would change if you move the data points on the horizontal axis?"

[7] Yuro: […] I don't […] remember exactly, but I think that if you move [*the data points*] the curve is not going to be affected by standard deviation that much, but you have to move it may be extremely.

By "the curve is not going to be affected by standard deviation that much" [19] Yuro was not clear in his prediction but he probably meant that moving the data points on the horizontal axis was not going to cause a big change in the peak of the height of gC curve, except if "you have to move" the data points "extremely." This of course would be correct if each of the six points were one of the infinite points that make up the normal distribution. But Yuro did not specify the direction of moving the point, so I asked, "In which direction is the data point being moved?"

[8] Yuro: Just any […] like if you move this point [*pointed at* ***O***—Fig. 16] all the way to the right […], the curve should go like, should change its height […].

Although Yuro provided a clear direction of dragging the point, “all the way to the right”, he did not clarify if the “change” in the “height” of the gC sketch was going to be an increase or a decrease. That led me to ask:

> [9] Int: Should the curve peak go down or go up as you drag the data point O “all the way to the right?”
>
> [10] Yuro: The height? It should go down. If you move it that way [*left side of mean line*], it’s going to go down as well.

Yuro’s prediction in [10] was correct, suggesting that he could connect the change in the height of the gC sketch with the change in the distribution of data points on the horizontal axis. However, when I changed the question “what if you move data point T and then point O (both on opposite sides from the mean line, Fig. 16), toward the mean line?” Yuro predicted that the curve peak was “still going to go down”,

> [11] Yuro: It’s [*the peak*] still going to go down […] not flat, it’s going to decrease slightly. Doesn’t matter where you move it this way, it will decrease […].

Predicting changes on the gC sketch seemed problematic to Yuro, and more generally to the other participants as well except Boris. In fact, Anita, Halen and Maya had hard times predicting the patterns of change in the gC sketch. It is not clear if participants' incorrect predictions were partly contributed to by the design of the gC sketch, which only provides individual data points on the horizontal axis and a normal curve fitted on them. However, Boris' correct prediction with the same sketch weakens the design claim. Interestingly, all the participants were familiar with the Normal curve, as they all covered the normal curve in the statistics courses. May be the difficulty was not with the normal curve per se, but with dynamic nature of the normal curve that participants had not encountered before. I expected Yuro to apply his response in [10], but in the opposite sense, and predict that moving two data points toward the mean line would show an increase in the peak of gC sketch.

With the numerical scales for the mean and standard deviation turned on by Yuro, I asked him to go ahead and “check your predictions [...]”. Yuro dragged data point O (Fig. 16) away from the mean line and back toward the mean line and said:

[12] Yuro: Ok, so it decreases [*as he dragged data point **O** away on the right*]; it decreases [*as he dragged point T away on the left side away*]. Oh, it increases [*as he dragged point T toward the mean line*].

As Yuro dragged data point **O** and then after he dragged **T**, away from the mean line, he noticed that the height of the gC sketch decreased. He then changed direction and dragged data point T back toward the mean line and the peak of the gC sketch slightly increased as he said “Oh, it increases” [12]. Based on the semiotic mediation perspective, Yuro’s statements, “Ok, so it decreases”; and “Oh, it increases” revealed instances where the Dragging tool, along with the visual feedback provided, enabled him to relate the changes in the gC sketch with changes in the magnitude of the standard deviation.

I asked Yuro to check “what happens if you drag all the points closer to the mean line?” Yuro responded by slowly dragging the data points, one after another, closer to the mean line and as the sharp peak (Fig. 31) emerged he said:

[13] Int: Oh, ok the centre point changes and the curve becomes narrower […] ah ok, yes it becomes narrower and there is less error in that curve.

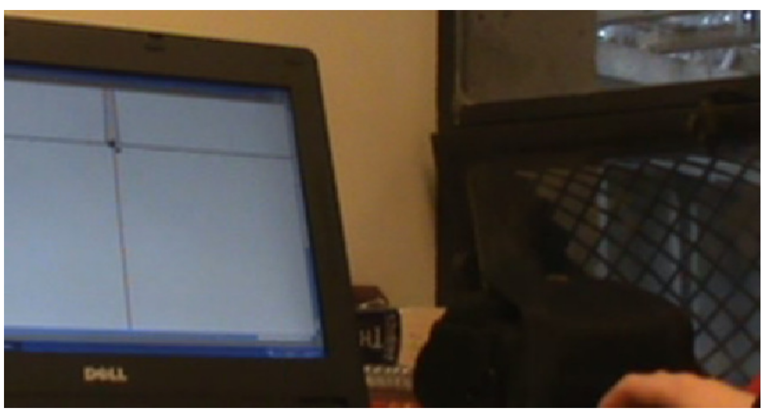


Figure 31. Yuro dragged data points closer to the centre and noticed that “the curve becomes narrower and there is less error.”

By “the centre point changes and the curve becomes narrower […]”, it is not clear what Yuro meant, but he may have observed from his dragging action that, as the magnitude of standard deviation decreased, the peak of gC sketch increased and the curve became "narrower" having "less error." By "the centre point" changes, Yuro probably meant the mean line changing its location as data points were dragged on the horizontal axis. Still, statement "the centre point changes" suggests an artifact sign whereas the statement " the curve becomes narrower" suggest a pivot sign connecting the more mathematical sign "[…] and there is less error in that curve" (Bartolini Bussi and Mariotti, 2008). His last sentence "[…] and there is less error in that curve" suggests that the Dragging tool served as a semiotic mediator for the magnitude of error in the curve, the "narrower" the curve, the "less error in that curve." Moreover, it is not clear if by "less error in that curve" Yuro had in mind the small magnitude of standard deviation associated with such a sketch in Figure 31, but Yuro did not mention standard deviation explicitly.

5.5.4. Reflecting on the Standard deviation

In Task 4, with the computer closed, I asked Yuro to reflect on the notion of standard deviation after the tasks. "What do you say about the term standard deviation?" I asked. Yuro paused for a moment and said:

> [14] Yuro: Standard deviation is the spread-like distances from the mean. I think the same thing […] yeah distances from the mean. Spread, standard deviation [*pause*]. You know what, actually I think now spread is actually different […]

Yuro appear to correctly associate "standard deviation" with the "spread" describing it as "spread-like distances from the mean." However, from his last statement in [14], Yuro seemed to change his mind, thinking that "spread is actually different" from standard deviation. I asked him, "Have you changed your mind?"

> [15] Yuro: Yeah, that's what, because now that I think about it [*standard deviation*], I use in a different way. Yeah, standard deviation I just think of it as the distance from the mean, to the points that are around the mean.

Yuro's consideration of standard deviation as "distance from the mean to the points that are around the mean" may have been evoked by his interaction with the dynamic sketches, for instance, the changes in the gC sketch in Figure 31. Although Yuro's statement [15] suggests a distinction between spread and standard deviation, his main challenge appeared related to the use of statistical terms rather than his considerations of variability. In fact, Yuro's statement in [14], "standard deviation is the spread-like distances from the mean" and statement [15] "standard deviation, I just think of it as the distance from the mean", are not very different. Thus, I believe that Yuro used the terms standard deviation and spread as if they were different and yet he likely meant the same thing.

5.5.5. Summary of Yuro's Data Analysis

In the static tasks, Yuro was rather brief in his answers, but he provided adequate information to enable me understand his thinking about variability. His thinking about of distribution and the mean and standard deviation in the static environment seem to rhyme with those of Anita, Halen and Maya. Yuro's interactions with the dynamic sketches in the computer-based environment revealed much clearer consideration of the connections among standard deviation, the mean and data distribution than in the static environment. At the end of the task, when the computer was already closed, Yuro struggled with the meaning of spread and standard deviation. However, it was not clear if Yuro's struggle was about reconciling the meaning of standard deviation from the dynamic sketches and the one he learned in class, which was more formal and static. It seems that Yuro's more correct thinking about standard deviation dependent more on him interacting with the dynamic and physical sketches than without the sketches.

5.6. Summary of Chapter

Table 3 summarizes participants' notions of standard deviation before, during, and after interacting with the dynamic sketches. Before interacting with the dynamic sketches, four of the five participants were more likely to link standard deviation with the normal distribution curve. Moreover, all the five participants considered the mean as "average", a number obtained by adding up a set of numbers and dividing the sum by the number of data values in the set.

Table 3. **Summary of Participants' Notions of Standard Deviation before, during, and after Interacting with the Sketches**

	Before	During	After
Anita	As I said before, if you […] can figure out the mean of data […] then you can derive the standard deviation, um and you can also figure out the normal distribution [...] curve. **[Static/Procedural= S/P.]**	If you drag [a data point farther to the right] then the mean will increase and therefore the standard deviation will increase. If I move this [point to the left] then it will decrease because I'm moving […]. **[Dynamic/Physical=D/Ph.]**	Standard deviation, […] I did one of those examples based on […] the graphs, um, I realized that as the mean was increasing farther to the right, the standard deviation was also increasing, so that was a very direct relationship with the mean, […] whenever you moved a certain data point to the right or left, based on how much you moved it […]. **[D/Ph.]**
Boris	**[**Standard deviation] kind of measures the variation of data from the mean**. [Aggregate reasoning= AGR].**	If you move [points] away from the mean line, the square is getting bigger and bigger because the square is the distance from the mean line right […]. **[D/Ph.]**	Standard deviation, as you move the points away from the mean, the standard deviation increases, that's what the graph shows […]. If the data points are equal difference from each other, […] shift the data points to the left or right […] it just shifts the mean but it won't change the standard deviation. **[D/Ph.]**
Halen	Standard deviation […] that is similar to the deviation […] if you have [the Normal curve], the standard deviation at the centre will be zero, and […] one standard deviation will be sixty eight percent […]. **[S/P.]**	If you […] stretch this way *away from the mean line*, I think they [*the squares]* will become [...] a bit more narrow […] becomes like skinnier and larger.[…]. **[D/Ph.**]. I thought like […] [the squares] would all go together […] ok they just don't […]. Like if you move this side it […], yeah so it affects both sides. **[AGR.]**	Well, if you change standard deviation the mean is going to change. When I was looking at the graphs, I didn't realize that, […] if you move the points together […], the mean is going to change too. **[D/Ph.]**

	Before	During	After
Maya	[…] I see standard deviation in *graphs*, there [are] like one, two, three, [and] four; then there's negative one, negative two, negative three, and negative four. You can calculate standard deviation. **[S/P.]**	Standard deviation is how far away from the mean a data point is, it is like a category, it categorizes how far away from the mean a data [point] is. [**AGR.]**	Standard deviation is a certain point away from the centre of a population […] **[AGR.].** I have a picture of a normal distribution divided into sections, which are called standard deviation […].There is a formula, I forgot but it's like standard deviation equals the square root of the variance. **[S/P.]** **[AGR.], [S/P.] = MIXED considerations of standard deviation.**
Yuro	Standard deviations, […] how far the points are […] around the dots […] one point on the left, one point on the right and then one point in the middle and then all these random points around it. **[S/P.]**	Oh, if you move [a *point to the left side, the squar*e] is going to increase, and if you move [*the point to the right side, the square*] is going to decrease. [**D/Ph.** Because I only felt that [the horizontal line] was going to increase but not [vertical sides of the square]. But now that I think about it, it makes sense because […]. I thought […] the heights would be constant […]. **[AGR.]**	Standard deviation is the spread-like distances from the mean […]. Yeah, standard deviation I just think of it as the distance from the mean to the points that are around the mean. **[AGR]**

During interactions with the dynamic sketches in *Sketchpad,* all the participants were able to link changes in standard deviation with the dragging of data points on the horizontal axis, and with the mean. After the tasks, four of the five participants were more likely to coordinate changes in the data points with changes in the magnitudes of the standard deviation and the mean. One participant (Maya) showed mixed considerations of the meaning of the standard deviation after the tasks. By mixed considerations, I mean that, Maya showed aggregate reasoning about standard deviation, but she also reasoned about standard deviation in a static way. For example, Maya's statement that "standard deviation is a certain point away from the centre of a population" [Maya, 22] generally satisfies aggregate consideration (Wild & Pfannkuch, 1999), but the statement "[…] standard deviation equals the square root of the variance shows a static and more procedural consideration of standard deviation. Hence, I described Maya's constructs of the standard deviation as "mixed consideration of variability."

From Table 3, four categories of participants' constructs of standard after using the dynamic sketches stand out. They are, not in any particular order: i) static/ procedural; ii) aggregate iii) dynamic/physical; and iv) mixed consideration. The static considerations reveal thinking about standard deviation, and in general reasoning about data, based on more formal procedures (e.g. text book formulas) rather than provide qualitative meanings of concepts. The aggregate considerations of standard deviation identify the centre of data and the deviations of the data values from the centre, while dynamic and physical considerations involve a lot of body (e.g. hand) movement while participants explain their thinking about a concept. In general, dynamic/ physical interactions include aggregate reasoning, except it involves more physical and dynamic expressions. Finally, mixed construct includes some elements of aggregate reasoning as well some elements of static reasoning. Table 3 shows that in general, participants progressed from static/ procedural thinking about standard deviation to more physical and dynamic constructions of the meaning of standard deviation. I argue that the Dragging tool was used as semiotic mediator for the meaning of standard deviation and its functional linkages with the mean and data distribution on the horizontal axis. In Chapter 6, I discuss the analysis of data and link the discussions to my research questions. I will respond more directly to the research questions in Chapter 7, as I conclude the study.

6. Discussion

In Chapter 5, based on two major theoretical perspectives, I analyzed my participants' considerations of variability on the tasks that they solved in the static, and in the dynamic environments. Based on the Vygotskian socio-cultural and historical perspectives, I argued in Chapter 5 that the Dragging tool was used by participants as an instrument of semiotic mediation for the meaning of standard deviation and its links with the other features of variability such as the mean and data distribution. Participants' thinking about standard deviation showed four categories: i) static/procedural ii) aggregate; iii) dynamic/physical; and iv) the mixed considerations. Moreover, based on the foundations of statistical thinking, focusing on the considerations of variability, my analyses revealed that the dynamic sketches supported participants' reasoning about data distributions in aggregate, by stating the connections between the centre of data, and the spread of data values from the centre.

The current chapter, which comprises of three sections, discusses results of the analyses in Chapter 5 with reference to the issues that were raised in the previous chapters, particularly in Chapters 2 and 3. Section 6.1 discusses participants' considerations of variability as they solved tasks in the static environment. I base my discussions on Wild and Pfannkuch's (1999) foundations of statistical thinking, focusing on the consideration of variability. In connection with the considerations of variability perspectives, I also include the use of the dynamic sketches for constructing mathematical meanings (Konold & Higgins, 2003). Konold and Higgins (2003) propose four different ways that students think about a given data set. They are data as: i) *pointers*; ii) *cases*; iii) *classifiers*; and iv) *aggregate.* Aggregate consideration focuses on the overall characteristics of a data set and not only on the characteristics of single data points. I used aggregate reasoning perspective to analyze participants' statements both in the static and in the dynamic environments, whereas Reid and Reading's (2008) consideration of variability hierarchies (CVH) suited my analysis of tasks in the static environment.

In connection with the considerations of variability, I also discuss Carlson et al.'s (2002) covariational reasoning perspective. Covariational reasoning did initially enable me to attend to the participants' thinking about the changes in the magnitude of standard deviation in relation to changes in the magnitude of the mean as the data points were dragged along the horizontal axis. However, it is important to underline that the mean and the standard deviation do not in fact co-vary. Moreover, the participants were moving points, and not changing the mean or the standard deviation directly. Further, out of the five levels of Carlson et al.'s covariational reasoning framework (i.e. coordinating value, direction, quantity, average rate, and instantaneous rate), only three levels (i.e. value, direction, and quantity) applied to my interview data. Nevertheless, I used this covariational reasoning perspective as a starting point for developing a multi-variation reasoning perspective, which I discuss in Chapter 7.

In Section 6.2, I discuss evidence of participants' considerations of variability while they solved tasks with the dynamic sketches. Using the Dragging tool of *Sketchpad,* participants explored the functional connections between the notions of standard deviation and the mean as they dragged data points along the horizontal axis. To consider the complex relationships between physical tools, and signs produced by the tools while participants solve the tasks, I based my interpretations on Vygotsky's socio-cultural historical perspective of learning and on semiotic mediation. According to Vygotsky (1978), the social environment influences learning through the use of its artifacts, that is, through its cultural objects and language. Moreover, social interactions can transform students' learning experiences from interpersonal to more personal awareness through a process of internalization.

Section 6.3 discusses participants' considerations of variability as they reflected on the notion of standard deviation at the end of the interview tasks. I chose standard deviation given that it is an important feature, among others, for describing variability, for instance in in a Gaussian distribution. Standard deviation also links with other concepts such as the mean and distribution to provide more information about a given data set. My expectation was that after the tasks, participants would be able to state the functional connections among the standard deviation, the mean and the data distributions in their own words, as well as to reason with graphs of distributions in aggregate. I now discuss in detail participants' interactions in each of three sections mentioned above, and link the discussions to my research questions.

6.1. Participants' Considerations of Variability in the Static Environment

The notion of variability encompasses many constructs in statistics, but I chose three constructs in my study: standard deviation, mean and distribution. In Chapter 2, I reviewed research studies about students' understanding and difficulties with the notion of distribution (e.g., Konold, Pollatsek, Well, & Gagnon, 1997). I also reviewed students' challenges with the concept of sampling distributions, which includes the ideas of variability (Chance, delMas, & Garfield, 2004; delMas, Garfield, & Chance, 2004). In connection with the notion of sampling distribution, delMas and Liu (2005) propose that part of students' challenges with this concept may be related to their difficulties with concepts such as the mean, distribution and standard deviation as well as the linkages among these concepts. For example, a study by Garfield, delMas and Chance (1999) reveals that some students struggle comparing distributions in histograms, focusing on the top part of the bars (i.e. smooth or irregular) instead of comparing the relative density of the data points around the mean of the distributions.

Analysis of my data revealed that in the static environment, participants' descriptions of distribution, mean and standard deviation tended to focus on the normal distribution. Although the normal distribution is one example of a family of distributions, it seemed to be the prototypical one for the participants. For instance, when I asked what the term distribution meant to them, Anita and Halen responded thus:

> Anita: Well, immediately what pops up is the normal distribution curve because that is part of what we have in statistics […]
>
> Halen: […] to me the first thing would be like the normal distribution […] and then it could be like other distributions […]

Reactions from my participants show that the normal distribution is "the first thing" that "pops up" when they hear the term distribution. The normal distribution curve is an important theoretical model in introductory statistics and has many applications in modeling real situations. For example, to Anita, the normal distribution is “part of what we have in statistics”, suggesting that it was one of the main theoretical models that she learned in her statistics course. She also stated, rightly or wrongly, that the normal distribution was used in grading her courses. From my participants’ statements, the normal distribution seem to distract their attention from a more general understanding of distribution, which includes the shape, centre and spread of data from that centre. It seems that students pay more attention to the concepts that directly apply to them, and in the process, may ignore the other meanings related to such concepts. Anita and Halen paid no attention to the concept of distribution beyond the normal distribution. Based on Konold and Higgins’ (2003) work, my study found that in the static environment, participants were more likely to reason with data as pointers (e.g., Anita’s “immediately what pops up is the normal distribution [...]”)*;* data as cases (e.g. Halen’s “to me, first thing would be like the normal distribution and then […] other distributions”); and using data as a classifier (e.g., Anita’s “I immediately think of averages because […] simple definition of the mean […] would be adding up the numbers […]”). These examples exclude aggregate reasoning, which Konold and Higgins propose as evidence of students’ consideration of variability and statistical thinking.

Moreover, four out the five participants in my study were more likely to consider the mean as a “number” obtained by a process of “adding up the numbers […].” For example, Anita expressed her thinking about the mean as follows:

> Anita: I immediately think of averages because, um, well, simple definition of the mean, I guess would be adding up the numbers in the data set and then dividing it up by the figures you have.

The students’ challenges with the construct of the mean, or the arithmetical average, are not new. Nineteenth-century scholars faced the same dilemma. As Stigler (1999) describes, Quetelet and Jevons wrestled with questions such as: How should we interpret a social average when the identity of each of the elements in the aggregate is well known? For example if “the heights of Adelphe Quetelet and William Stanley Jevons are 5´1˝ and 5´9˝, what then is 5´5˝? Surely not the height of Adelphe Quetelet or William Stanley Jevons” (p. 3). A related question was, should we model the behaviour at the individual level (micro-level), or should the model be at the level of the group distribution (macro-level), through a bell–shaped Normal curve as Galton and Edgeworth would have it? Stigler submits that although the individual-level (micro-level) models had the appeal of capturing much of the individual information in the data, the data usually failed the checks because the models did not "capture between individual dynamics or correlation" (p. 4). The group (or aggregate) models had the benefit of "much wider applicability as a compensation for less sensitivity to individual dynamics, and less ability to incorporate individual characteristics" (p. 4). Modern statistics has preserved both approaches, but there have been more appeals in favour of the group model over the individual-level model. However, my study revealed a strong tendency of my participants to reason about the mean at the micro-level, as a single value. Thus it is not surprising that my participants did not use Hardiman, Well & Pollatsek's (1984) balance model of the mean, by considering the mean in aggregate.

On the concept of standard deviation, participants who linked the image of distribution with the normal distribution curve were more likely to associate the Z-scores with the standard deviation. Halen, for example, sketched a Normal curve [Fig. 21] and used it to describe her thinking about standard deviation:

> Halen: Standard deviation […] if you have like this one right here [*she points at her sketch in* Fig. 21], like the standard deviation at the centre will be zero, and then […] you have in here, like one standard deviation will be sixty eight percent […], from here to here that will be the second standard deviation […].

Another example was Maya, who reported that she "sees standard deviation in graphs", which "graphs" she later clarified meant the "Normal curve". According Maya, you can also "calculate standard deviation",

> Maya: Standard deviation? Um, standard deviation, I see standard deviation in *graphs*, there is like one, two, and three four; then there's negative one, negative two, negative three […]. You can calculate standard deviation.

My findings on students' thinking about standard deviation agree with delMas and Liu's (2005) submission that:

> Most instruction on the standard deviation tends to emphasize teaching a formula, practice with performing calculations, and tying the standard deviation to the empirical rule of the normal distribution (p. 56).

Although delMas and Liu's quotation above is about instruction for which I have no direct evidence, it suggests some anecdotal evidence on the down-side of teaching with a focus on calculations and algorithms. delMas and Liu contend that instructions that emphasize calculations and procedures do not necessarily promote conceptual understanding of standard deviation. The author's contention is supported by similar findings by Pollatsek, Lima & Well's (1981) on their students' understanding of the mean of a data set. Pollatsek et al. contend that the computational rules do not imply any real understanding of the basic underlying concept, but in addition may actually inhibit students' understanding of other related concepts. Both delMas and Liu as well as Pollatsek et al. draw researchers' and educators' attention to designing learning tasks that do not focus students on calculations as the primary method of teaching and learning concepts.

However, given that students forget some concepts that they have learned in class, it is not possible that classroom instruction is entirely responsible for all the gaps in students' understanding of concepts. Moreover, there are a lot of supplementary materials outside the classroom, (e.g. text books, applets etc.) that students can use to support their classroom learning. Many of these resources in undergraduate basic statistics are designed following the GAISE's (2005) learning goals of teaching and learning statistics (in Chapter 2). For instance, Moore's (2013) textbook, "*The Basic Practice of Statistics",* which is widely used in introductory statistics courses*,* fully supports GAISE's learning objectives, such as using 'actual' or 'real' data to help students construct meanings of concepts; emphasizing conceptual understanding rather than focusing on applying formulas and algorithms; encouraging collaborative learning such as group projects, and class discussions; and having students use well-selected computing tools for teaching and learning concepts. In general, many of GAISE's teaching and learning goals are implemented with the static tools, such as paper and pencil, graphs and diagrams in the text books. What is not well known is how students engage with physical and dynamic models of statistical variability. I discuss this question in the next section. The discussions in the current section contribute to answering my first research question.

6.2. Participants' Considerations of Variability while Solving Tasks with Dynamic Sketches

In Section 6.1, I discussed the analysis of my participants' constructs of variability in the static environment, which related to my first research question. In this section, I discuss findings from the analysis of participants' interactions with the dynamic sketches in three overlapping categories, namely: i) dynamic and physical expressions, including gestures; ii) constructions of mathematical meanings; and iii) affective and aesthetic expressions (e.g. fun, surprise, and amazement) while engaging with the tasks. Lastly, I discuss participants' understanding of the constructs of standard deviation and statistical variability in Section 6.3. I will use some episodes presented in the previous chapters as examples.

6.2.1. Physical and Dynamic Expressions

At the prediction and the testing of predictions, my participants' showed more physical body movements and gestures than they did in the static environment. For instance, they used active verbs such as "increasing", "decreasing", "going down", "getting smaller" and "clustering", which Bartolini Bussi & Mariotti (2008) associate with artifact signs, that is, signs associated with using a physical tool on a given task. For example, while interacting with the dyMS sketch, Anita said, "I saw the mean 'getting smaller' and standard deviation 'gets bigger' […]". Boris predicted that "if you 'move' [data points] away from the centre, the square is 'getting bigger and bigger' […]." For Boris, the size of the square was a metaphor for "scale", which represented how far a data point was from the mean line. Boris' example revealed dynamic and physical thinking, which seemed to be evoked by the perceived action of dragging the points on the horizontal axis. Maya's predicted that, "when I move the [data] points away from the centre, the curve will plunge, go down" and accompanied her prediction by spreading out her arms across the table as she said "the curve will plunge" (Fig. 28).

Moreover, when asked to predict how the squares on the dyMS sketch would change if he dragged the data points on the horizontal axis, Yuro attention was focused on the activity of "moving" rather than on the object that was moving or being moved: "it's going to move this way"; "this is going to move this way"; "I'm going to move this way"; and "the whole thing is going to go this way" [Yuro, 3]. It is interesting that Yuro's own "moving" did not come first but after "it" and "this", suggesting that he was engaged more with the motion than the object being moved. His expressions support Arzarello et al.'s (2009) proposition that dynamic thinking is multimodal (i.e., involves many different aspects). It seems from Yuro's statements that multimodal thinking is a complex manner of thought, which involves some kind of physical movement.

The examples given so far (and there are more to come), suggest that my sketches evoked dynamic and physical thinking from the participants. My participants' interactions with the sketches also linked changes in two or more varying entities in the dynamic sketches, such as Maya's statement in Chapter 5, "When I move the [data] points away from the centre, the curve will go down, but when I move points closer to the centre, the curve will rise." Although there is no covariational relationship in the context of Carlson et al. (2002), in Maya's statement, her statement specifies the direction of moving the "[data] points away from the centre" in connection with the change in the height of the gC sketch as she states, "the curve will rise." Carlson et al. consider the statement "the curve will rise" as quantitative coordination, in that it accounts for the amount and direction of change in a given entity. According to them quantitative coordination is one of the higher cognitive levels of covariational reasoning.

Overall, the tasks with the gC sketch showed to be more challenging for the participants to predict than those on the dyMS sketch. The main challenge seemed to arise from participants assuming that the height of the gC sketch would increase as data points were dragged away from the mean line. I had anticipated in my design that some participants might associate an increase in standard deviation with increase in the height of the curve and a decrease in the standard deviation with a decrease in the height of the curve. Hence, participants' incorrect prediction about changes on the gC sketch was not particularly surprising.

After interacting with the sketches, participants thinking about the concepts seemed different from their thinking at the beginning of the interview. For example, Yuro associated the sharp peak of the gC curve with "less error" in the curve. From his statement it is not clear if by less "error" Yuro meant reduced magnitude of the standard deviation. Participants' considerations of the variability after interacting with the dynamic sketches support de Freitas and Sinclair's (2011) proposal that diagrams can give rise to new ways of thinking, as well as new kinds of awareness. I argue that the sketches evoked dynamic and physical, time-dependent thinking in my participants, which seemed to support their considerations of variability. Section 6.2.1 contributes to a response to my second, third and fourth research questions.

6.2.2. Using the Signs to Construct Mathematical Meanings

Studies suggest that students are more likely to consider functional relations among varying entities in terms of discrete values than in a qualitative way (e.g., Trigueros & Ursini, 1999; Thompson, 1994). In designing my sketches, I incorporated both the qualitative and numerical considerations. For instance, I included the numerical scales for standard deviation and the mean in order to enable participants confirm their predictions (Chapter 4 & 5). During interactions with the dynamic sketches, participants were able to link changes in standard deviation with changes in the mean, as well as the time-dependent movement of data points on the horizontal axis. After her interactions with the sketches, for example, Halen indicated more awareness of the links between standard deviation and the mean,

> Halen: Well, if you change standard deviation, the mean is going to change. When I was looking at the graphs, I didn't realize that […] if you move the points together […] the mean is going to change too

Although Halen was not as explicit in her use of the word "change", she acknowledged the functional connection between standard deviation and the mean, in her statement "[…] if you change standard deviation, the mean is going to change.", Halen "realized" that change in the standard deviation was linked to change in the mean, led me to argue that she relied on the signs produced in the dragging action to internalize the changes on the sketch. Wertsch and Addison Stone (1995) submit that internalization is an evolving connection between the physical changes produced from using an artifact, and the internally-oriented signs. For Wertsch and Addison Stone, internalization represents the process of constructing individual knowledge as generated by a shared experience. Halen's statement suggests that she became aware of the connection between the standard deviation and the mean as well as the movements of the data point on the horizontal axis, though at the level of using the artefact.

A similar episode occurred when Anita solved the task on the dyMS sketch. She dragged a data point away from the centre on the right side of the mean line, and after realizing that both the standard deviation and the mean increased in magnitude, contrary to her predictions, said:

> Anita: Oh, so both of them got larger, yeah ok, so I thought the mean would get smaller and the standard deviation would get larger but, actually both of them are increasing.

Anita's statement, "[…] actually both of them are increasing" suggests that she became aware of the changes in the magnitude of standard deviation and the mean and related them with the direction of dragging the data points on the horizontal axis, away from the mean line. Drawing on Bartolini Bussi and Mariotti (2008), I argue that the Dragging tool was used by Anita as an instrument of semiotic mediation for the links between standard deviation, the mean and distribution of a data set. The discussions in this section contribute to a response to my second, third and fourth research questions.

6.2.3. Engaging in the Tasks: Surprise, Amazement, Fun

> When I look at the tool like this one [*the dynamic sketch*], my first question is, what would I use it for? I am always keen on anything that gets people to play, anything that brings a sense of discovery, and wonder, the fun thing (Interview from a mathematician, cited in Ekol, 2011)

This section briefly discusses the affective and aesthetic dimensions of learning (i.e., episodes in which participants expressed surprise, amazement, and fun during their interactions with the dynamic sketches). The affective components of teaching and learning are rarely discussed in mathematical activities. The citation at the beginning of this section is from a professional mathematician who was asked to evaluate a dynamic sketch for solving mathematical problems. His response, as well as evidence from other studies (e.g. Sinclair & Gol Tabaghi, 2010), show that professional mathematicians' activities do include affective and aesthetic elements such as elegance, fun, clarity, simplicity, brevity, structure and power (e.g. see Jacobsen, 2010). However, it is unclear if mathematicians agree that these elements can be developed and promoted by the teacher through mathematics teaching and learning activities.

During their interactions with the sketches, my participants used statements that suggested surprise and amazement at some of the results they obtained from the activities. For example, Maya's "oh, this is nice"; Boris' "that was interesting"; Halen's "Oh, that's so cool" and Maya's "Wow, so big!" all suggested affective dimensions. In the case of Maya, she seemed so amazed at seeing a very high and thin gC sketch that she said "Wow, so big!" It seems Maya had not anticipated such a sketch to evolve from a normal distribution curve. Researchers (e.g., Sinclair, 2001; Sinclair, Zazkis, Liljedhal, 2002; Sinclair, Pimm, & Higginson, 2006) agree that expressions such the above have some aesthetic as well as affective dimensions. The authors further agree that the affective dimensions can include sensations of pleasure as a result of apprehending and discerning information, including patterns in a learning activity. In fact, Sinclair (2001) conceives an aesthetic response as an act of acknowledging structure or order perceived as being intuitive and pleasing. I believe the responses from my participants support Sinclair's and the above researchers' findings.

Furthermore, although the affective aspects of learning are not usually given much attention in the teaching and learning of mathematics nor in research, some researchers (e.g., Sivan, 1986; Schunk, 1995) disagree with the classical views that affective dimensions are entirely an internal state, or wholly dependent on the environment as predicted by reinforcement theories (e.g., Skinner, 1953). Rather, Sivan (1986) contends that motivation depends on the cognitive activity in interaction with sociocultural and instructional factors, which include language and other forms of assistance to the learners by the teachers. Moreover, Schunk (1995) believes that good instruction can raise motivation for learning and motivated learners tend to look for affective learning environment.

More recent studies show that although students recognize that affective components are important, they do not integrate these components in their individual studies. For example, Petocz and Reid's (2003) research studies on students' experience of learning statistics showed that students rated enthusiasm as very important to good teaching. However, interestingly, students' scores on the affective scales were consistently independent of aspects of teaching and learning such as how organized the learning materials and activities were. Thus, the affective components were completely de-linked from students' discussion of their own learning. Petocz and Reid conclude that students "value enthusiasm and motivation, but they believe that it is an aspect of their study that comes from outside rather than from within." (p.50) I have argued that students' responses to the affective components reflect how the teaching and learning institution considers them. In general, the affective dimensions are not considered as contributory to learning.

In fact, Roth and Lee (2007) point out tensions between the epistemological and ontological aspects of human development (Packer & Goicoechea, 2000); the differences between de-contextualized and embodied knowledge (Lave & Chaiklin, 1993); the difficulty of planning for specific forms of learning (Holzkamp, 1992); and the apparent disjunction between individual learners and their social environments (Barab &; Shultz, 1986). Roth and Lee use these examples to argue strongly that, by excluding affective dimensions, contemporary applications of the Vygotskian historical socio-cultural theories have not taken a holistic approach as Vygotsky may have intended. The discussions in this section have attempted to show that the affective components are indeed important and need to be considered more in the teaching and learning designs. Section 6.2.3 contributes to a response to my second, third and fourth research questions.

6.3. Students' Considerations of Variability after the Tasks

> […] Our investigation, […] showed that as the basic forms of activity change […] and a new stage of social and historical practice is reached, major shift occur in human mental activity. These […] involve the creation of new motives for action and radically affect the structure of cognitive processes. Luria (1976, p. 161; cited by Bartolini Bussi & Mariotti, 2008, p.747):

This section provides an overview of participants' awareness of the meaning of variability after the tasks. I cite Luria (1976) to provide some context to the discussions in this section. The main argument in this section is that the signs produced by my participants' interactions with the sketches supported them to consider variability in a way that was different than at the beginning before they interacted with the dynamic sketches. I use an episode from Boris' data to exemplify my argument.

Toward the end of the interview, I asked Boris to reflect on the term standard deviation:

> Boris: […] as you move the points away from the mean, the standard deviation increases […]. As you concentrate data at the center […] you get small values of standard deviation. If the data points are [at] equal difference from each other, without changing the difference between them, shift[*ting*] the data points to the left or right […] just shifts the mean, but it won't change the standard deviation.

It is worth noting that Boris stated the above results after interacting with the dyMS and gC sketches. His statement reveals a deeper awareness of how standard deviation may not change as the mean changes with the movement of data points on the horizontal axis. Boris' results cannot be easily stated without using the sketches to observe the changes as the points are dragged physically. Other participants in my study also showed more awareness of the meaning of standard deviation and its applications after using the sketches (for example, Anita's, statement "[…] I thought the mean would get smaller and the standard deviation would get larger but actually both of them are increasing", suggests that after using the sketches, she became more able to link changes in standard deviation to changes in the mean. I argue that the signs produced by participants' interactions with the dynamic sketches supported them to link the features of statistical of variability. This Section contributes to a response to my second, third and fourth research questions. I now move to Chapter 7, and respond to my research questions, as well as discuss some contributions of my study to statistics/mathematics education.

7. Conclusion

Chapter 6 discussed findings from the data analysis chapter and linked the discussions to my research questions. In this chapter, I summarize the analyses and respond directly to my research questions. As well, I discuss some contributions of my study to research in mathematics/statistical education in general, and to the teaching and learning of statistics at the university (and perhaps secondary) level in particular. There are altogether six sections in this chapter: Section 7.1 provides specific answers to my research questions. In section 7.2, I discuss limitations of the study and issues of validity of my results. Section 7.3 proposes some contributions of my study to research in mathematics and statistical education, in the area of variability. I discuss implications of the study for post-secondary introductory statistics curriculum in section 7.4. In section 7.5, I suggest areas for future research studies. I conclude the chapter with brief personal reflection on the study, in section 7.6.

7.1. Responding to my Research Questions

As articulated in Chapter 3, the following are the specific research questions that the current study set out to answer:

1. What do students say about measures of statistical variability, such as *distribution*, the *mean* and *standard deviation*, in a data set presented in a static environment?
2. How do students express notions of variability while interacting with dynamic mathematics sketches?
3. How do students express notions of variability after interacting with dynamic mathematics sketches?
4. What might be the contribution of dynamic mathematics sketches to students' considerations of variability and statistical thinking?

I will respond to the questions one by one.

7.1.1. What do Students say about Measures of Statistical Variability such as Distribution, the Mean, and Standard Deviation, in a Data set presented in a Static Environment?

Based on my analyses and discussions in the previous chapters, the current study found that the majority of my participants were more likely to consider the normal distribution curve as the prototypical (and perhaps even only) example of a distribution. I expected participants to think about distribution more broadly in terms of how a data set is spread out from its centre rather than focusing on one example, which is the normal distribution. In fact, four out of the five participants in my study provided the normal distribution curve when asked about distribution. Hence, I argue that the image of the normal distribution seems to distract students from the general image of distribution.

Moreover, three participants (Anita, Halen, and Maya) considered the mean as a number obtained by 'adding up the numbers and dividing by the number of data values in a data set.' Two of the participants (Boris and Yuro), were quite brief in their answers and each described the mean as the "average", which I interpreted to mean the same thing as adding up the numbers and dividing the sum by the number of data values in the data set. Thus, all the five participants shared the same static image of the mean based on using the formula for the mean rather than the conceptual meaning of the mean. Moreover, three of the five participants incorrectly equated standard deviation with the Z-scores in the standard Normal curve rather than considering it as a measure of how spread out a data set is from its centre. My findings on the notions of distribution, mean and standard deviation were not surprising given findings from previous studies (e.g. Pollatsek, Lima & Well, 1981; Hardiman, Well, & Pollatsek, 1984; Konold & Pollatsek, 2002). Pollatsek, Lima & Well (1981) had shown that many students think about the mean in terms of calculations rather than as a conceptual model. Pollatsek et al. (1981) contend that "computational rules [...] may actually inhibit the acquisition of more adequate (relational) understanding" (p. 202). In my study, participants who considered the mean and standard deviation in terms of calculations at the micro-level (Stigler, 1999) were more likely to be challenged by relating the mean and standard deviation in a dynamic and functional way. For example Anita's, "if you can figure out the mean of a data set […] then you can derive the standard deviation" (Anita [4]) focused on calculations (deriving) rather than on the qualitative links among the mean, standard deviation, and the distribution of data.

Like Anita, Halen's thinking of the mean was, "if you have a couple of numbers, you add them all up and then you divide by how many numbers there are, you get like the average or the mean." Halen thought of standard deviation as a pointer (Konold & Higgins, 2003) to standard normal curve (Fig. 21) rather than a feature of variability in a data set. Anita and Halen's examples represent the thinking that most of my participants had about the notions of distribution, the mean and standard deviation in the static environment. To summarize my answer to the first research question: my study participants were more likely to consider the term distribution as the 'normal distribution'. The term mean was more likely to be considered as a number obtained through the process of adding up numbers and dividing by the total number of data points present; whereas standard deviation was more likely to be linked to the Z-scores in the standard Normal curve. Hence, participants' constructions of distribution, the mean, and the standard deviation showed static images, suggesting the links with images found in text books. That leads to my second research question.

7.1.2. How do Students Express Notions of Variability while Interacting with the Dynamic Mathematics Sketches?

My participants used semiotic resources (Arzarello et al., 2009) such as gestures, drawings, and words, to construct the meaning of features of variability, such as the standard deviation, the mean and distribution. Through moving data points along the horizontal axis using the Dragging tool of *Sketchpad,* participants used the signs produced in the activity to explore the patterns and to interpret the relationships among the changing entities. Thus, based on a semiotic mediation perspective, the signs produced in the interactions supported participants' considerations of statistical variability. I argue that the dynamic sketches contributed to my participants' considerations of the relations between data distributions, the mean and the standard deviation. Moreover, the physical and dynamic expressions by the participants, for example, Boris' statement that "the squares are getting bigger and bigger" support my hypothesis in Chapter 3, that students' thinking about standard deviation after interacting with the sketches would show more dynamic and physical expressions. I propose that the sketches evoked dynamic and physical thinking in my participants, which seemed to support their considerations of variability in aggregate.

Second, the use of signs enabled participants to interpret complex dynamic connections on the sketches and to communicate these patterns in a more mathematical way. For example, Halen's statement, "When I was looking at the graph, I didn't realize that […] if you move the points together […] the mean is going to change too" suggest that after the interactions she did consider the changes in the graph and became more aware of the dynamic links between standard deviation and the mean. Vygotsky calls such 'realization' by Halen "*internalization"* (an internal reconstruction of the meaning of variability through an activity with a physical tool). In Chapter 3, I assumed that my participants' interactions with the dynamic sketches would enable them express meanings of standard deviation, the mean and distribution more clearly than they did when solving similar tasks in the static environment. My findings so far support this assumption. I propose that the signs that evolved from participants' expressions while dragging points using the Dragging tool supported them in constructing mathematical meaning of variability.

Third, during their interactions with the dynamic sketches, participants used verbal expressions such as "oh, that's so cool", "wow, so big!", "oh, this is nice" which, suggested their affective engagements with the tasks. Research studies indicate that the affective and aesthetic domains are rarely considered in mathematics teaching and learning activities, partly due to lack of a robust instrument to assess it. However, Roth and Lew (2007), with some examples, strongly contend that affective dimensions should be an integral part of teaching and learning. Several other studies (e.g. Sinclair, Pimm & Higginson, 2006; Sinclair, Zazkis & Liljedhal, 2003) also show that the affective and aesthetic dimensions are present in the ways mathematicians solve problems. To sum up my answer to the second research question: analyses of my data reveal that the activities with dynamic sketches evoked physical and dynamic expressions from my participants, which helped them to develop the meaning of statistical variability. In particular, the dragging actions seem to help participants reason about the features of variability and to state the patterns they noticed without using formulas or algorithms. Moreover, the sketches also evoked affective dimensions from my participants, such as surprise, pleasure, amazement and some fun, which I found to be consistent with the problem-solving practices of professional mathematicians (e.g., Sinclair & Gol Tabaghi, 2010; Ekol, 2011).

7.1.3. How do Students Express Notions of Variability after interacting with the Dynamic Mathematics Sketches?

Analyses of my data revealed improved considerations of the meanings of standard deviation, mean and distribution among participants during, and after the tasks in the dynamic environment compared with the static environment. I propose that dragging actions, and the signs produced in the interactions enabled participants to internalize the mathematical features of statistical variability. The dynamic sketches also evoked physical and dynamic expressions from my participants, for instance, Maya's statement, "When I move the [data] points away from the centre, the curve will plunge, go down" was accompanied by a physical action of opening her arms and lowering them (Fig. 28), suggesting the deictic gesture for "plunge" or "go down." Moreover, the dynamic sketches also evoked considerations of motion in participants' statements of variability (e.g., [Yuro, 3]; and [Boris, 5] in Chapter 5). These findings support Angel and Gibb's (2013) more recent proposal that digital environments, such as *Sketchpad* have strong connection to the affective modes of communication.

7.1.4. What might be the Contribution of Dynamic Mathematics Sketches to Students' Considerations of Variability and Statistical Thinking?

The major contribution of the dynamic mathematics sketches, as seen in participants' interactions, was that the sketches provided participants with physical and concrete mathematical tools with which to reason about abstract mathematical constructs, such as standard deviation, mean and distribution, and to relate these constructs to statistical variability. Moreover, the tasks with the dynamic mathematics sketches moved participants away from considering features of statistical variability merely as numbers obtained from calculations using formulas, but to begin to attend to the patterns and meanings of these features in aggregate. From a Vygotskian perspective, the signs produced in participants' interactions with the dynamic mathematics sketches mediated their constructions of the meaning of statistical variability and enabled them to reason statistically. Moreover, the sketches also evoked aesthetic and affective aspects of the mathematics which seemed to motivate participants. For example, the statement below shows how Maya reasoned with a statistical graph (the dyMS sketch); it also has some affective elements after Maya solved the task:

> So, the mean increases as standard deviation […]. As standard deviation increases, the mean also increases. Oh no, the mean decreases right? […] Oh, this is nice [Maya, 14].

Maya seemed pleased with her finding after successfully solving the task. However, the affective and aesthetic components were not noticed in my participants' statements when they performed tasks in the static environment, which suggests that dynamic and physical sketches more easily evoke affective and aesthetic considerations in the learning tasks than static mathematical tools.

7.2. Limitations of the Study

I summarize the limitations in the study in three categories: i) the design of the dynamic sketches; ii) the number of study participants; and 3) the challenge of integrating the dynamic with the static concepts. The first two issues are so important that they can raise questions of broader validity of the study results. The third limitation is important to consider in developing curriculum materials, for instance, materials that use dynamic graphs for teaching university-level statistics courses.

7.2.1. The Design of the Dynamic Sketches

First, a possible design limitation is that the sketches may have given participants a false impression that the mean and standard deviation always co-varied when data points are dragged on the horizontal line. However, as Boris was able to prove, dragging data points may vary the mean value but standard deviation remains unchanged. One way of helping students to understand this is to design a task that first asks students to predict "What will happen to standard deviation and the mean if you select all the data points and drag them on the horizontal axis, to the right or to the left of the mean line?"

After their prediction, students could then check by selecting all the data points and dragging them on the horizontal axis, to the left or right side of the mean line.

Second, I used a small sample of numerical data points to design the dynamic sketches. However, as I explained in Chapter 4, I chose a small data set that so as to focus (Konold, 2007) participants on exploring the patterns of variability among the entities on the sketches. Based on my findings, I am satisfied that the study largely fulfilled the scope of its design. My study focus was on how students construct meanings about statistical variability in the static, and in the dynamic environments and whether the dynamic sketches contributed to their understanding. Moreover, I did not witness questions from my participants that would have required more data points. However, notwithstanding all the precautions that I took in the designing the sketches, limitations cannot be entirely ruled out. Future studies may be interested in increasing the number of data points and involving a small group of two to three students to work together to see if their interactions with the sketches and the discussions among them generate more mathematical signs.

7.2.2. The Number of Study Participants

I took a qualitative approach to the study and interviewed five participants in the final study. I ensured that I used complete data for all five participants on the same constructs. Overall, having complete and consistent data from all participants supported the validity and reliability of my findings. I believe that the measures that I took in planning, designing and implementing the study enabled me to adequately address all the issues that came to my attention. However, future studies can consider involving more participants, for instance, in small-group interviews rather than in one-on-one interviews in my case, to give another perspective of the learning outcomes.

7.2.3. Integrating the Dynamic with Static Concepts

Given that, in general, students are used to the textbook materials that are generally static, it stands to reason that my participants initially had some difficulties integrating dynamic concepts with the static ones. For example, the idea of the mean changing its value over time as data points were changed seemed problematic to my participants at the prediction stages. As well, the continuity of the normal distribution curve was challenging for most of them to predict, but later on, the tasks proved to be quite interesting for them. I argue that moving students from a predominantly static learning environment to a more dynamic one is a challenging task for both the students and the instructors, but the dynamic activities engage students in the learning tasks more than the static ones. Curriculum planning and design may consider incorporating dynamic sketches in the activities for teaching and learning abstract concepts in statistics and mathematics. I will come back to curriculum issues in my recommendations toward the end of this chapter.

7.3. Contributions to Research in the Mathematics/Statistics Education

The study has contributed to two locally designed sketches, the *dyMS Sketch* (pronounced as the *Dimes Sketch*), and the *gC Sketch* (pronounced as the *Geek* Sketch). The designs were based on statistical principles, and implemented using *Sketchpad* software. As I discussed in Chapters 5 and 6, the physical and dynamic as well as the visual nature of the sketches evoked participants' thinking about the notions of variability. My claim is that at the end of the tasks, participants were more able to explain, in their own words, patterns of change in standard deviation in relation to change in the mean, in connection with changes in data distribution. Formerly, participants relied on text book rules and formulas rather than on their own constructions of meanings of concepts. My findings suggest that the dynamic sketches have great potential to engage participants in the tasks and to challenge them to look for meanings of concepts.

7.3.1. Extending the Conceptual Understanding of Statistical Variability

Given the importance of variability in statistics, I believe my study contributes some novelty in integrating geometrical-computing tools to the teaching and learning of statistical variability in first-year university-level statistics. Geometrical approaches enable students to explore concepts using physical models and to develop meanings as they move toward formal mathematical meanings. For example, the physical sketches allow students to explore with graphs, to test conjectures, and to solve problems by integrating geometrical thinking (Jackiw, 1991, 1995; Goldenberg et al., 1998) with statistical thinking (Wild and Pfannkuch, 1999; Pfannkuch & Wild, 2004), as well as statistical reasoning (Cobb 1993; Moore, 1997; GAISE, 2005). Moreover, geometrical approaches can enhance students' understanding of the meaning of symbols, when they apply the symbols in the calculations. From the participants' statements as they solved tasks with the dynamic sketches, I developed a framework that I call a multi-variation reasoning (MVR).

7.3.2. Proposing a Framework for Multi-variation Reasoning.

In Section 3.2, I presented a covariational reasoning perspective based on Carlson, Jacobs, Coe & Hsu's (2002) work and noticed its similarity with Falcade, Laborde and Mariottti's (2007) work on reasoning about two co-varying entities. Carlson et al. (2002) define covariational reasoning as cognitive activities involved in coordinating two varying entities while attending to the ways in which the entities change in relation to each other. Their definition is consistent with Falcade et al.'s (2007) geometrical framework on "co-variation" (p. 3) that the authors implemented in a secondary classroom setting while teaching the concept of function in mathematics. Unlike the above researchers, however, I use covariational reasoning as a platform to develop a multi-variation reasoning framework in my participants' considerations of variability in statistics. Hence, my study extends the work of Carlson et al. and Falcade et al. on covariational reasoning in secondary school mathematics to multi-variation reasoning framework in the undergraduate statistics.

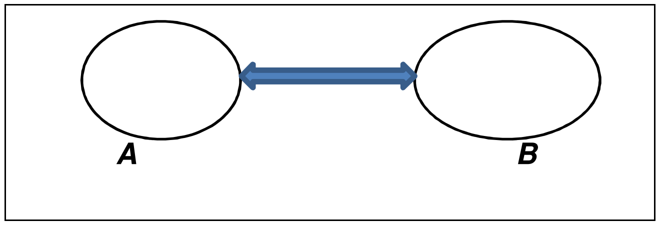


Figure 32. Covariational relation between two varying entities A and B (Carlson, Jacobs, Coe, Larsen, & Hsu, 2002; Falcade, Laborde, & Mariotti, 2007).

Both Carlson et al. (2002) and Falcade et al. (2007) explored co-variation reasoning between two functions A and B. The researchers' framework can be represented more generally as "*Change in entity 'A' responding to change in entity 'B*" (Figure 32).

Figures 33 and 34 respectively show three and four entities that change when participants solve tasks on the dyMS and gC. The entities that vary on the gC sketch are represented by the letters U, V and W. The proposed multi-variation framework describing participants' interactions is: "*Change in entity 'U' in relation to change in entity 'V', both responding to change in entity 'W'.*"

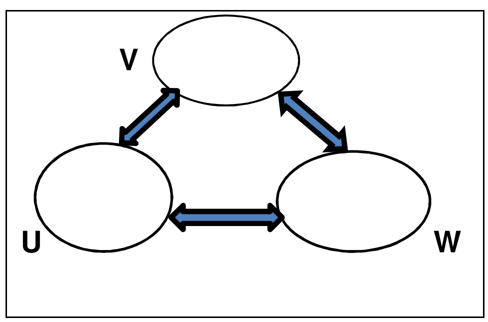


Figure 33. The multi-variation relationships among three varying entities U, V and W.

The statement below is given as an example of the statements that students may make as they interact with the dyMS sketch and interpret the changes in the sketch,

> *As the data points are dragged away from the mean line on the right side (U), the magnitudes of the standard deviation (V), and the mean (W) both increase.*

With respect to the activities with the gC sketch, I used four entities U, V, W, and Z (Fig. 34). The multi-variation framework is: '*Change in entity U, effecting changes in entities V, W, and Z."* For example, the statement, below can be considered under a multi-variation framework involving four entities:

> As the data points are dragged toward the mean line from the right side (U), the magnitudes of standard deviation (V) and the mean (W) decrease, whereas the height of the gC sketch increases.

Of course students' statements are expected to vary as they solve similar tasks, but common elements should include: i) key words that specify changes in *magnitude* of the entities (e.g. increase or decrease in an entity); ii) *direction* of movement of an entity from the mean-line (e.g. away from the mean line or toward the mean-line); and iii) statements should specify the *space* where the task is being solved (e.g., on the right side or left side of the mean line).

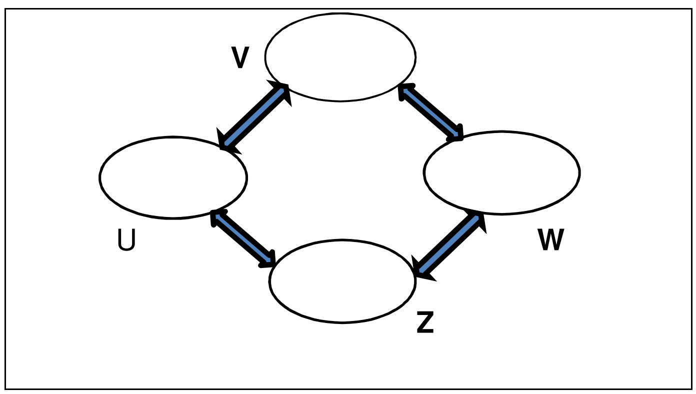


Figure 34. The multi-variation relationship involving four varying quantities U, V, W and Y used to analyze tasks on the gC sketch.

Building on Carlson et al.'s (2002) and Falcade et al.'s (2007) work and on the above proposals, I define a multi-variation reasoning (MVR) as follows: *"Statements/actions that involve coordinating changes in three or more entities, while attending to the specific ways that the entities change in magnitude, direction and location in relation to a reference point."*

Table 4 provides a summary of the different levels of a multi-variation reasoning framework. It is worth noting that, a multi-variation framework does not categorize the students or the learners; rather, it categorizes the statements that the students make during interactions with the sketches. Column one presents the levels of consideration of variability (CVs), one through five. Column two describes the possible activity at each CV level. Column three proposes some behaviour that can be observed in each level. I have used the notation CV (#) to specify one level from another (e.g. CV1 is level one and CV2 level 2). There are five levels in my proposed framework. Following are brief descriptions for each CV level.

CV1 is the level one consideration of variability. At this level, students' statements on the tasks are more general and do not specify the change in the first entity in response to change in the second entity. The term "change" could be a decrease or an increase in a given entity but it is not made explicit in students' statements. For example, Halen's statement, "if you change standard deviation, the mean is going to change too" [Halen, 10, p. 84] satisfies a CV1 level. **CV2** is the level two consideration of variability. It includes a statement of the *direction* of change in one entity responding to change in the second entity. **CV3** is the third level of consideration of variability that includes coordinating the magnitude of change in one entity in relation to the change in another entity. **CV3** is similar to Carlson et al.'s (2002) quantitative coordination. In fact **CV1, CV2**, and **CV3** are similar to Carlson et al.'s first three levels of covariational reasoning (i.e. considering—*change, direction*, and *magnitude* of one entity—in relation to change in the second entity).

In **CV4**, the fourth level of considerations of variability, a student is expected to relate the amount of change in one entity with changes in two different entities at the same time. **CV4** relates more to the tasks with the dyMS sketch. In **CV5**, the fifth level of consideration of variability, a student is expected to connect the amount of change in one entity with changes in three different entities on the sketch. **CV5** relates to the interactions with the gC sketch. The CV levels apply equally in the dyMS and the gC sketches, except CV5 that does not apply on the dyMS sketch.

Table 4. The Proposed Multi-variation Reasoning Framework

Consideration of Variability (CV) level	Descriptions of actions	Behavior
CV1	Coordinating the change in one entity with changes in another entity.	Statement and actions which show coordinating two entities, e.g., "*A changes when B changes*"
CV2	Coordinating the direction of change of one entity with change in another entity.	Statement and actions showing an awareness of the direction of change of one entity while considering changes in the second entity, *e.g., "A changes as B is moved to the right side of a reference point""*
CV3	Coordinating *the amount* of change of one entity with changes in the another entity	Statement and actions showing consideration of the amount of change of the in one entity while considering changes in the second entity. e.g., As *"entity A increases, entity B decreases as it is moved to the left side of a reference point."*
CV4	Coordinating the amount of change in one entity with the changes in two other entities	Statement and actions showing consideration of the amount of change in one entity while responding to changes in two other entities (Fig. 33). *The magnitudes of A and B both increased as entity C changed direction.*
CV5	Coordinating the amount of change in one entity with changes in three other entities	Statement and actions showing consideration of the amount of change in one entity while responding to changes in three other entities (Fig. 34). *e.g., The magnitudes of A and B both decreased as C changed in the direction d, such that D increased.*

If a student's statement on a dynamic task includes elements in the highest CV level in Table 4, then his or her interactions are placed at that level. For example, Boris' statement [Boris, 10] in Chapter 5, includes coordinating the amount of change in one entity with changes in three or more entities, hence his interaction can be categorized at CV level five (i.e. CV5) of consideration of variability.

7.4. Implications for the Curriculum

Beginning topics in introductory statistics can be considered as the foundation for understanding more advanced topics in statistics. Thus, students may benefit from integrating geometrical approaches that would help them understand the connections among concepts by, for example, exploring and testing conjectures. The two sketches designed can be modified and used as teaching and learning resources for developing conceptual understanding about standard deviation and its links with the mean and data distribution. The physical and dynamic representations may enable students develop greater coordination between symbolic and static representations, which are mostly found in textbooks, and the dynamic ones, which are not. I propose that the geometric approach to teaching variability would help students develop flexibility among different representations of concepts. In fact, as my study findings suggest, interactions with dynamic and physical geometric representations of concepts could enable students to develop more aggregate reasoning with data.

Integrating dynamic and geometric concepts can also enable students to explore concepts and to build their own knowledge on the basis of properties that are noticed on the sketches (as was the case in Gol Tabaghi's (2012) study of linear algebra concepts). At the start of the tasks, participants in my study recalled more procedural and symbolic approaches to the concepts of distribution, the mean and standard deviation, but they did not recall the concepts. My study agrees with Gol Tabaghi's (2012) recommendation that a balance between the two approaches—dynamic geometric in her instance and in my case dynamic analytical—would enable students to overcome certain learning difficulties. I suggest that the two approaches be used concurrently with students exploring the dynamic approaches through small projects and homework assignments.

My study participants also showed difficulties distinguishing between the terms distribution and the normal distribution as well as between standard deviation and the Z-*scores* in the standard Normal curve. In both cases, when asked about distribution, all of them except one considered the normal distribution curve. Similarly for standard deviation, my participants considered the Z-scores in the standard Normal curve. I suggest that the general construct of distribution be introduced to students through dynamic graphs for them to coordinate and construct the meanings of concepts such as the mean, standard deviation and other features of variability in the data set. Using more specific sketches such as the gC sketch, students could then focus on the Normal curve as a special case of distribution. As Gol Tabaghi's (2012) study proposes, the integration of dynamic and geometric concepts could provide students with concrete contexts for meaning making of abstract notions such as the standard deviation.

With respect to the assessment of tasks, I recommend, based on this study that students are assessed through small projects and course assignments. For example, students can be asked to explore properties of concepts such as standard deviation, through dynamic geometric graphs and to write reports (500-600 words) about their findings. Such reports have two benefits: i) they can encourage students to use the geometric tools for exploring mathematical/ statistical constructs; and ii) they can also encourage students to practice communicating their results with a larger audience, for example, in small groups discussions. I recommend that the curriculum material be designed following the Vygotskian socio-cultural context, with emphasis on exploring concepts, writing reports and communicating results with a larger audience. However, I also support students having a strong content base of the subject of mathematics and statistics. Emphasis on communication alone without mathematical content is not sufficient. It is akin to drilling students on the English lexicon for its own sake.

7.5. Implications for Future Studies

I have proposed in my study that the pivot signs may be more easily evoked in activities involving small-to-large-group discussions than in individual interactions. Future studies could shed more light on these questions: "Are the pivot signs independent of artifact signs? Does group size (the number of participants working together on mathematical tasks, using dynamic mathematics sketches) affect the generation of mathematical signs?" One suggestion is to organize a study in which participants work on tasks in small groups, but vary the group sizes, while others work on the same task individually.

With regard to the multi-variation framework, I recommend that future studies consider testing and validating my proposed levels of consideration of variability on statistical/mathematical tasks. Finally, the current study proposes that foundation statistics students are more likely to face challenges with statistical terminologies and statistical language in general, which impact negatively on their understanding of the meaning of statistical ideas. More studies, especially on variability, would clarify this claim. On the issue of statistical language, I recommend that more opportunity is given for students to practice writing and communicating mathematical/statistical concepts in small working groups and where possible, in larger groups, to consolidate ideas that they learn especially in large classes. Small group and presentations, facilitated by teaching assistants could help students develop analytical skills and build confidence in sharing statistical/mathematical concepts with peers.

7.6. Personal Reflection

This study has enabled me to appreciate more deeply the challenges that students can face with the meaning of statistical concepts and, by extension, with statistics as a discipline. If students are not comfortable with the foundation concepts, they are very unlikely to understand other related concepts. Statements such as "statistics is not my subject" that some students make about the subject could be linked to foundation concepts of statistics that they failed to get to grips with, conceptually. By interacting with the participants in my study and reading their transcripts, I was encouraged in my belief that taking this line of research has been worthwhile. The study has made me want to continue to extend the applications of dynamic, physical and more interactive tools for learning concepts in statistics. In the future, I would like to extend the study on the normal distribution by incorporating more concepts, given that the normal distribution attracts many applications as well as students' attention, and yet poses some challenges at the same time such as using the Normal curve to estimate probability of an event happening. I believe that the engagement, commitment, fun, and the sheer joy, among many other benefits that such an approach entails, is worth putting time on.

At a more personal level, the study took me through many theoretical and philosophical landscapes, which I had to navigate to find my own bearings and come back home. Along the way, I traversed through many lands, some known and others not so well known to me. I passed through the Vygotskian land and the Vygotskians gave me a pair of binoculars to help me navigate the semiotic terrain and join the Wild land. In the Wild land, the Wildians also handed me another pair of binoculars to navigate the rocky foundations and find my way home. I cherish the thoughts and wisdom of all those who pointed me to my way home. And now I am Home!

References

Angel, M., & Gibbs, A. (2013). At the Time of Writing: Digital Media, Gesture, and Handwriting, *Electronic Book Review.* Retrieved October 1, 2013 from*,* http://www.electronicbookreview.com/thread/electropoetics/gesture,

Artique, M. (2002). Learning mathematics in a CAS environment: The genesis of a reflection about instrumentation and the dialectics between technical and conceptual work. *International Journal for Computers in Mathematical Learning*, *7*(3), 245–274*.*

Arzarello F. (2000). Inside and Outside: Spaces, Times and Language in Proof Production', in: *Proceedings of PME XXIV*, Hiroshima, Japan, **1**, 23-38.

Arzarello, F., & Paola, D. (2007). Semiotic Games: The Role of the Teacher. In J. H. Woo, H. C. Lew, K.S. Park & D. Y. Seo (Eds.), *The Proceedings of the 31st Conference of the International Group for the Psychology of Mathematics Education*, *2*, 17–24.

Arzarello, F., Olivero, F., Paola, D. and Robutti, O. (2002). A cognitive analysis of dragging practises in Cabri environment. *ZDM, 34* (3), pp. 66-72.

Arzarello, F., Paola, D., Robutti, O., & Sabena, C. (2009). Gestures as semiotic resources in the mathematics classroom. *Educational Studies in Mathematics*, *70*, 77-109.

Atkin, A. (2013). Peirce's Theory of Signs, *The Stanford Encyclopedia of Philosophy*. In Edward N. Zalta (ed.). Retrieved August 22, 2013 from URL=<http://plato.stanford.edu/archives/sum2013/entries/peirce-semiotics/>.

Bakker, A. (2004). *Design research in statistics education: On symbolizing and computer tools* (Doctoral Dissertation). Utrecht, The Netherlands: CD Beta Press.

Bakker, A., & Gravemeijer, K. P.E. (2004). Learning to reason about distributions. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 147–168). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Barab, S. A., & Plucker, J. A. (2002). Smart people or smart contexts? Cognition, ability and talent development in an age of situated approaches to knowing and learning *Educational Psychologist*, *37*, 165–182.

Bartolini Bussi, M. G., & Mariotti, M. A. (2008). Semiotic mediation in the mathematics classroom: Artifacts and signs after a Vygotskian perspective. In L. English, M. Bartolini Bussi, G. Jones, R. Lesh, & D. Tirosh (Eds.), *Handbook of international research in mathematics education, second revised edition*. Mahwah: Lawrence Erlbaum.

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. 121–145). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Ben-Zvi, D., & Garfield, J. (2004). Research on reasoning about variability: A forward. *Statistics Education Research Journal*, *3*(2), 4-6.

Ben-Zvi, D. (2006). Scaffolding students' informal inference and argumentation. In A. J. Rossman & B. L. Chance (Eds.), *Proceedings of the seventh international conference on the teaching of statistics* (ICOTS-7), Salvador, Bahia, Brazil, [CD-ROM]. Voorburg, The Netherlands: International Statistical Institute.

Biehler, R. (1999). Discussion: Learning to think statistically and to cope with variation. *International Statistical Review, 67*(3), 259–262.

Biehler, R. (2001). "Girls (tend to) watch less television than boys" – Students' hypotheses and data exploration strategies in group comparison tasks. Talk presented at LOGOS #10, Mathematics Education Unit, Department of Mathematics, The University of Auckland, New Zealand.

Biehler, R., & Steinbring, H. (1991). Entdeckende Statistik, Stengel-und-Blätter, Boxplots: Konzepte, Begründungen und Erfahrungen eines Unterrichtsversuches [Statistics by discovery, stem-and-leaf, box plots: Basic conceptions, pedagogical rationale and experiences from a teaching experiment]. *Der Mathematikunterricht, 37* (6), 5-32.

Bransford, J., Brown, A. L., & Cocking, R. R. (Eds.) (2000). *How people learn: Brain, mind, experience, and school.* Washington, DC: National Academy Press.

Carlson, M., Jacobs, S., Coe, E., Larsen, S., & Hsu, E. (2002). Applying covariational reasoning while modeling dynamic events: A framework and a study. *Journal for Research in Mathematics Education,* (5), 352–378.

Chance, B., delMas, R., & Garfield, J. (2004). Reasoning about sampling distribution. In D. Ben-Zvi & J. Garfield (Eds.), *The Challenge of Developing Statistical Literacy, Reasoning, and Thinking* (pp. 295–318). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Cobb, G. W. (1992). Report of the joint ASA/MAA committee on undergraduate statistics. In the *American Statistical Association 1992 proceedings of the Section on Statistical Education*, (pp. 281–283). Alexandria, VA: American Statistical Association.

Cobb, G. W. (1993, July). Reconsidering statistics education: A national science foundation conference. *Journal of Statistics Education*, *1*(1), 1–28

Cobb, G., & Moore, D. (1997). Mathematics, statistics and teaching. *American Mathematical Monthly, 104*(9), 801–823.

delMas, R. (1997). A framework for the development of software for teaching statistical concepts. In J. B. Garfield & G. Burril (Eds.), *Research on the role of technology in teaching and learning statistics: Proceedings of the 1996 International Association of Statistics Education (IASE) Round Table Conference* (pp. 85-99). Voorburg, The Netherlands: International Statistical Institute.

delMas, R. C., Garfield, J., & Ooms, A. (2005, July). *Using assessment items to study students' difficulty with reading and interpreting graphical representations of distributions*. Presented at the fourth international research forum on statistical reasoning, thinking, and literacy (SRTL-4), July 6, 2005. Auckland, New Zealand.

delMas, R. C., & Liu, Y. (2005, May). Exploring students' conceptions of the standard deviation. *Statistics Education Research Journal*, *4*(1), 55–82

diSessa, A. A. (2007): An Interactional Analysis of Clinical Interviewing, *Cognition and Instruction*, *25*:4, 523-565.

Drijvers, P., Kieran, C., & Mariotti, M-A. (2010). Integrating technology into mathematics education; theoretical perspectives. In C. Hoyles & J.-B. Langrange (eds.), *Mathematics Education and Technology: Rethinking the terrain.* New York: Springer.

Ekol, G. (2011). Understanding and promoting mathematical modeling competencies: An applied perspective: In G. Kaiser, W. Blum, R. Borromeo-Ferri & G. Stillman (Eds.), *Trends in Teaching and Learning of Mathematical Modelling*, *vol. 1,* pp. 55-64. New York: Springer.

Erickson, T. (2002). Fifty Fathoms: Statistics demonstrations for deeper understanding. Oakland, CA: EEPS Media.

Falcade, R., Laborde, C., & Mariotti, M.A. (2007). Approaching functions: Cabri tools as instruments of semiotic mediation. *Educational Studies in Mathematics, 66,* 317–333.

Fong, G. T., Krantz, D. H., & Nisbett, R. E. (1986). The effects of statistical training on thinking about everyday problems. *Cognitive Psychology, 18,* 253–292.

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,* (2), 124–158.

GAISE (2005). Guidelines for assessment and instruction in statistics education (GAISE) college report. Alexandria, VA: American Statistical Association

Gal, I. (2002). Adults' statistical literacy: Meaning, components, responsibilities. *The International Statistical Review*, *70*(1), 1–25.

Garfield, J. (2000). Beyond testing and grading: New ways to use assessment to improve student learning. In T. L. Moore (Ed.) T*eaching statistics: Resources for undergraduate instructors* (pp. 201–208). Washington DC: Mathematical Association of America.

Garfield, J., & Ben-Zvi, D. (2008). Developing students' statistical seasoning: Connecting research and teaching practice, New York: Springer.

Garfield, J., delMas, R., & Chance, B. (2003). *Web-based assessment resource tools for improving statistical thinking.* Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Garfield, J., delMas, R., & Chance, B. L. (2007). Using students' informal notions of variability to develop an understanding of formal measures of variability. In M. C. Lovett & P. Shah (Eds.), *Thinking with data* (pp. 117–148). Mahwah, NJ: Lawrence Erlbaum Associates.

Ginsberg, H. (1981).The Clinical Interview in Psychological Research on Mathematical Thinking: Aims, Rationales, Techniques, *For the Learning of Mathematics*, *1*(3), 4-11.

Ginsburg, H. (1997). Entering the child's mind: The clinical interview in psychological research and practice. Cambridge, UK; New York: Cambridge University Press.

Gol Tabaghi, S. (2012). *Using Dynamic Geometry to Explore Linear Algebra Concepts: the Emergence of Mobile, Visual Thinking* (Doctoral dissertation).Vancouver: Simon Fraser University.

Goldenberg, E. P., & Cuoco, A. A. (1998). What is Dynamic Geometry? In R. Lehrer & D. Chazan (Eds.), *Designing Learning Environments for Developing Understanding of Geometry and Space* (pp.351-367), Mahwah, New Jersey: Lawrence Erlbaum Associates.

Goldenberg, P.E., Cuoco, A., & Mark, J. (1998). A role of geometry in general education. In R. Lehrer & D. Chazan (Eds.). *Designing learning space for developing understanding of geometry and space*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Goldenberg, E. P., Lewis, P., & O'Keefe, J. (1992). Dynamic representation and the development of an understanding of functions. In E. Dubinsky & G. Harel (Eds.), *The concept of function: Aspects of epistemology and pedagogy* (pp. 235-260). Washington DC: MAA.

Gould, R. (2004, November). Variability: One statistician's view. *Statistical Education Research Journal*, *3*(2), 7–16.

Guin, R., & Trouche, L. (1999). The complex process of converting tools into mathematical instruments: the case of calculators. *International Journal of Computer for Mathematical Learning,* 3, 195- 227.

Hammerman, J. K., & Rubin, A. (2004). Strategies for managing statistical complexity with new software tools, *Statistics Education Research Journal, 3*(2), *17-41.*

Hancock, C., Kaput, J. J., & Goldsmith, L. T. (1992). Authentic inquiry with data: Critical barriers to classroom implementation. *Educational Psychologist*, *27*(3), 7–364.

Hardiman, P., Well, A. D., & Pollatsek, A. (1984). Usefulness of a balance model in understanding the mean. *Journal of Educational Psychology*, *76*, 792–801.

Healy, L., & Sinclair, N. (2007). If this is our mathematics, what are our stories *International Journal of computers for*: *mathematical learning, 12*:3-21*?*

Hegedus, S. J. & Kaput, J. J. (2004). An introduction to the profound potential of connected algebra activities: Issues of representation, engagement and pedagogy. *The Proceedings of the 28th Conference of the International Group for the Psychology of Mathematics Education*, *3*, 129–136.

Hegedus, S., & Moreno-Armella, L. (2009). The transformative nature of "dynamic" educational technology. ZDM: The International Journal on Mathematics Education: Transforming Mathematics Education through the Use of Dynamic Mathematics Technologies, 41(4), 397-398.

Heid, M. K. (2005). Technology in Mathematics Education: Tapping into Visions of the Future. In W. J. Masalski and P. C. Elliot (Eds.), *Technology-Supported Mathematics Learning Environment*. Reston, VA: The National Council of Teachers of Mathematics.

Holzkamp, K. (1992). Die Fiktion administrativer Planbarkeit schulischer Lernprozesse [On the fiction of the possibility to administratively plan learning processes at school]. In K.-H. Braun & K. Wetzel (Eds.), *Lernwidersprüche und pädagogisches Handeln* (pp. 91–113). Marburg, Germany: Verlag Arbeit und Gesellschaft.

Hoyles, D., & Trouche, L. (2003). What can digital technology take from and bring to research in mathematics education? In A. J. Bishop, M. A. Clemens, C. Keitel, J. Kilpatrick & F. K. S. Leung (Eds.), *Second International Handbook of Mathematics Education* Vol. 1 (pp. .3-349), Dordrecht: Kluwer.

Jackiw, N. (1991). *The Geometer's Sketchpad.* Berkeley, CA: key Curriculum Press.

Jackiw, N. (1995). *The Geometer's Sketchpad, v3.0*. Berkeley, CA: key Curriculum Press.

Jacobsen, T. (2010). Beauty and the brain: culture, history and individual differences in aesthetic appreciation. *Journal of anatomy*, *216*(2), 184-191.

Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases.* New York: Cambridge University Press.

Konold, C. (1989a). Informal conceptions of probability. *Cognition and Instruction, 6*, 59–98.

Konold, C. (1989b). An outbreak of belief in independence? In the *eleventh annual meeting of the North American chapter of the international group for the psychology of mathematics education* (Vol. 2, pp. 203–209). New Brunswick, NJ: Center for Mathematics, Science and Computer Education.

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. (2003). Reasoning about data. In J. Kilpatrick, W. G. Martin & D.E. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 193–215). Reston, VA: National Council of Teachers of Mathematics (NCTM).

Konold, C., & Miller, C. D. (2005). *Tinker Plots: Dynamic data exploration* (Version 1.0) [Computer software]. Emeryville, CA: Key Curriculum Press.

Konold, C., & Pollatsek, A. (2002). Data analysis as the search for signals in noisy processes. *Journal for Research in Mathematics Education*, (4), 259–289.

Konold, C., Pollatsek, A., Well, A., & Gagnon, A. (1997). Students analyzing data: Research of critical barriers. In J. Garfield & G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics* (Proceedings of the 1996 International Association of Statistics Education round table conference, pp. 151–167). Voorburg, The Netherlands: International Statistical Institute

Lann, A., & Falk, R. (2003). What are the Clues for Intuitive Assessment of Variability? In C. Lee (Ed.), *Proceedings of the Third International Research Forum on Statistical Reasoning, Thinking and Literacy (SRTL-3).* [CDROM] Mount Pleasant, MI: Central Michigan University.

Lave, J., & Chaiklin, S. (Eds.). (1993). *Understanding practice: Perspectives on activity and context*. Cambridge University Press.

Lave J., and Wenger, E. (1991). *Situated learning: legitimate peripheral participation*. New York, NY: Cambridge University Press.

Lehrer, R., & Chazan, D. (Eds.). (1998). *Designing learning space for developing understanding of geometry and space*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Leung, A. (2008). Dragging in a Dynamic Geometry Environment through the Lens of Variation. *International Journal of Computers for Mathematical Learning*, (13) 135–157.

Luria, A. R. (1976). *Cognitive development its cultural and social foundations*. Cambridge, MA: Harvard University Press.

MacGillivray, H. (2004). Coherent and purposeful development in statistics across the education spectrum. In G. Burrill, & M. Camden (Eds.), *Curricular Development in Statistics Education: International Association for Statistical Education 2004 Roundtable.* Voorburg, The Netherlands: International Statistical Institute.

Mariotti, M. A. (2009). Artifacts and signs after a Vygotskian perspective: the role of the teacher. *ZDM Mathematics Education*, 41:427-440.

Mariotti, M. A. (2013). Introducing students to geometric theorems: how the teacher can exploit the semiotic potential of a DGS. *ZDM*, *45*, 441–452.

Mason, J. (2008). Being mathematical with and in front of learners: attention, awareness, and attitude as sources of differences between teacher educators, teachers and learners. In T. Wood and B. Jaworski, *International handbook of mathematics teacher education*: vol. 4. Rotterdam, the Netherlands: Sense Publishers, pp. 31-56.

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

Mokros, J., & Russell, S. J. (1995). Children's concepts of average and representativeness. *Journal for Research in Mathematics Education*, *26*(1), 20–39.

Moore, D. S. (1997). New pedagogy and new content: The case for statistics. *International Statistical Review, 65*, 123–165.

Moore, D. S. (1998). Statistics among the liberal arts. *Journal of the American Statistical Association, 93*, 1253–1259.

Moore, D. S. (2004). Foreword. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking,* (p. ix). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Moreno-Armella, L., Hegedus, S. J., & Kaput, J. J. (2008). From static to dynamic mathematics: Historical and representational perspectives. *Educational Studies in Mathematics*, *68*, 99-111

Nickerson, R. S. (1995). Can technology help teach for understanding? In D. N. Perkins, J. L. Schwartz, M. M. West, & M. S. Wiske (Eds.), *Software goes to school: Teaching for understanding with new technologies* (pp. 7-22). New York: Oxford University Press.

Olivero, F. (1999). 'Cabri-Géomètre as a mediator in the process of transition to proofs in open geometric situations', in: W. Maull & J. Sharp (eds.), *Proceedings of the 4th International Conference on Technology in Mathematics Teaching,* University of Plymouth, UK.

Packer, M. J., & Goicoechea, J. (2000). Sociocultural and constructivist theories of learning: Ontology, not just epistemology. Ontology, not just epistemology. *Educational Psychologist*, *35*, 227–241.

Petocz, P., & Reid, A. (2003). Relationships between students' experience of learning statistics and teaching statistics. S*tatistics Education Research Journal, 2*(1), 39-53.

Pfannkuch, M. (2005). Informal inferential reasoning: A case study. In K. Makar (Ed.), *Reasoning about distribution: a collection of research studies*. Proceedings of the fourth international research forum on statistical reasoning, thinking, and literacy (SRTL-4), University of Auckland, New Zealand, 2–7 July. Brisbane: University of Queensland, [CD-ROM].

Pfannkuch, M. &. Wild, C. J. (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, The Netherlands: Kluwer Academic Publishers.

Piaget, J. (1972). Intellectual evolution from adolescence to adulthood. *Human Development*, *1*(15), 1-12.

Pollatsek, A., Konold, C., Well, A. D., & Lima, S. (1984). Beliefs underlying random sampling. *Memory and Cognition*, *12*, 395–401.

Pollatsek, A., Lima, S., & Well, A. D. (1981). Concept or computation: Students' misconceptions of the mean. *Educational Studies in Mathematics*, *12*, 191–204.

Rabardel, P., & Samurçay, R. (2001). *From Artifact to Instrument-Mediated Learning, New challenges to research on learning,* International symposium organized by the Center for Activity Theory and Developmental Work Research, University of Helsinki, March 21-23.

Reading, C., & Shaughnessy, J. M. (2004). Reasoning about variation. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 201–226). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Reading, C., & Reid, J. (2005). Consideration of variation: A model for curriculum development. In G. Burrill & M. Camden (Eds.), *Curricular Development in Statistics Education: International Association for Statistical Education 2004 Roundtable* (pp. 36-53). Voorburg, The Netherlands: International Statistical Institute.

Reid, J., & Reading, C. (2004). Just a minute? The use of minute papers to investigate statistical thinking in research, teaching and learning. *Adults Learning Maths Newsletter* (21), 1-4.

Reid, J., & Reading, C. (2006). A hierarchy of tertiary students' consideration of Variation. In A. Rossman & B. Chance (Eds.), *Working Cooperatively in Statistics* Education. Proceedings of the Seventh International Conference on Teaching *Statistics*, Salvador, Brazil. Voorburg, The Netherlands: International Statistical Institute.

Reid, J., & Reading, C. (2008). Measuring the development of students' consideration of variation. *Statistics Education Research Journal, 7(1), 40-59.*

Rossman, A. J.(1996). *Workshop Statistics: Discovery with Data*. New York: Springer-Verlag.

Roth, W-M., & Lee, Y-J. (2007). "Vygotsky's neglected legacy": Cultural-historical activity theory. *Educational Research Review*, *77*(2), pp. 186–232.

Rumsey, D. J. (2002). Statistical literacy as a goal for introductory statistics courses. *Journal of Statistics Education, 10*(3), 6-13.

Saada-Robert, M. (1989). La microgenèse de la représentation d'un problème. *Psychologie française, 34*(2-3), 193-206.

Schunk, D. H. (1995). *Learning theories: An educational Perspective.* Upper Saddle River, NJ: Prentice Hall.

Seitz, J. A. (2000). The bodily basis of thought. New Ideas in Psychology: An *International Journal of Innovative Theory in Psychology*, *18*(1), 23-40.

Sfard, A. (1991). On the dual nature of mathematical conceptions: reflections on processes and objects as different sides of the same coin. *Educational Studies in Mathematics*, *22,* 1-36.

Sfard, A. (2007): When the rules of discourse change, but nobody tells you: Making sense of mathematics learning from a commognitive standpoint, *Journal of the Learning Science, 16:*4, 565-613.

Sfard, A. (2008). Thinking as communicating: human development, the growth of discourses, and mathematizing. New York: Cambridge University Press.

Shaughnessy, J. M. (1997). Missed opportunities in research on the teaching and learning of data and chance. In F. Biddulph & K. Carr (Eds.), *People in mathematics education* (Proceedings of the 20th annual meetings of the mathematics education research group of Australasia), (pp. 6–22), Rotorua, New Zealand: MERGA.

Shaughnessy, J. M., Watson, J. M., Moritz, J. B., & Reading C. (1999). *School mathematics students' acknowledgment of statistical variation: There's more to life than centres*. Paper presented at the research pre-session of the 77th annual meeting of the National Council of Teachers of Mathematics (NCTM), San Francisco, CA.

Shultz, J. (1986). The individual and the group: An examination of American folk ideology. *Anthropology & Education Quarterly*, *17*, 239–252.

Sinclair, N. (2001). The aesthetic is relevant. *For the learning of mathematics ion*, *21*(1)*,* 25-32.

Sinclair, N., de Freitas, E., & Ferrara, F. (2013). Virtual encounters: the murky and furtive world of mathematical inventiveness. *ZDM*, 1-14.

Sinclair, N. and Gol Tabaghi, S. (2010). Drawing space: mathematicians' kinetic conceptions of eigenvectors. *Educational Studies in Mathematics*, *74*(3), pp. 223-240.

Sinclair, N., Pimm, D., & Higginson, W. (2006). *Mathematics and the aesthetic: New approaches to an ancient affinity*. New York: Springer

Sinclair, N. and Yurita, V. (2008). To be or to become: how dynamic geometry changes discourse. *Research in Mathematics Education*, *10*(2)*,* 99-111.

Sinclair, N., Zazkis, R., and Liljedhal, P. (2002). Number worlds: Visual and experimental access to elementary number theory concepts. *International Journal of Computers for Mathematical Learning*, *8,* 235-263.

Sivan, E. (1986). Motivation in social constructivist theory. *Educational Psychologist*, *21*(3), 209-2.

Skinner, B. F. (1953). *Science and human behaviour*. New York: Free Press.

Slauson, L. V. (2008). *Students' conceptual understanding of variability*. (Doctoral dissertation). The Ohio State University: Retrieved December 16, 2012, from http://www.stat.auckland.ac.nz/~iase/publications/dissertations/08.Slauson.Dissertation.pdf.

Stavey, R., & Tirosh, D. (1996) Intuitive rules in science and mathematics: the case of 'more of A --more of B. *International Journal of Science Education*, *18*(6): 653-667.

Stigler, S. M. (1999). Statistics on the table: The history of statistical concepts and methods. Harvard University Press.

Strauss, S. (1987). Educational-developmental psychology: Some case studies. In L. Liben (Ed.), *Development and learning: Conflict or convergence?* (pp. 1–157). Hillsdale, NJ: Erlbaum.

Thompson, P. W. (1994). Students, functions, and the undergraduate curriculum. In E. Dubinsky, A. H. Schoenfeld, & J.J. Kaput (Eds.), *Research in College Mathematics Education Educational, I: Issues in Mathematics Education,* (Vol. 4, pp. 21-44). Providence, RI: American Mathematical Society.

Trigueros, M., & Ursini, S. (1999). Does the understanding of variable evolve through schooling? In O. Zaslavsky (Ed.), *Proceedings of the 23rd International Conference of PME, vol. 4* (pp. 273-280). Haifa, Israel: Israel Institute of Technology.

Utts, J. (2003).What educated citizens should know about statistics and probability. *The American Statistician, 57*(2), 74–79.

Verillon, P. and Rabardel, P. (1995). Cognition and artifacts: a contribution to the study of thought in relation to instrumented activity. *European Journal of Psychology of Education*, *9* (3), 77-101.

Vygotsky, L. (1978). Mind in society: The development of higher psychological processes. Cambridge, MA: Harvard University Press.

Vygotsky, L. S. (1981). The genesis of higher mental functions. In J.V. Wertsch (Ed.), *The concept of activity in Soviet psychology* (pp.147-1880. Armonk, NY: Sharpe.

Vygotsky, L. S. (1986). *Thought and language*, Translation newly revised and edited by Alex Kozulin. The MIT Press: Cambridge.

Watson, J. M. (2005). Developing an awareness of distribution. In K. Makar (Ed.), *Reasoning about distribution: A collection of research studies*. Proceedings of the fourth international research forum on statistical reasoning, thinking, and literacy (SRTL-4), University of Auckland, New Zealand, 2–7 July, 2005. Brisbane: University of Queensland.

Watson, J., Collis, K., Callingham, R., & Moritz, J. (1995). A model for assessing higher order thinking in statistics. *Educational Research and Evaluation, 1*, 247-275.

Wertsch, J. V., & Addison Stone, C. (1985). The concept of internalization in Vygotsky's account of the genesis of higher mental functions. In J. V. Wertsch (Ed.), *Culture, communication and cognition: Vygotskian perspectives* (pp. 162–166). New York: Cambridge University Press.

Wiggins, G., & McTighe, J. (1998). *Understanding by design*. Alexandria, VA: Association for Supervision and Curriculum Development.

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

Zawojewski, J. S., & Shaughnessy, J. M. (2000). Mean and median: Are they really so easy? *Mathematics Teaching in the Middle School*, *5*(7), 436–440.

Appendix A.

Interview Task 1: Exploring Participants’ Constructs of the Features of Variability in the Static Environment

What comes to your mind when you hear these words?

i) Distance

ii) Centre

iii) Deviation

iv) Distribution

v) Mean

vi) Standard deviation

Appendix B.

Interview Task 2:
Exploring Participants' Constructs of Variability with the dyMS Sketch

There are six numerical data points marked on the horizontal number line and labelled B, C, F, E, A and D. The vertical purple line shows the mean-line of the six data points. The length from each data point to the mean-line constitutes one side of a square formed with that point and the mean-line. Each square is identified with its respective data point, for example, "square B" is formed with data point B as one of its four corners. Six different squares are similarly constructed and all the squares touch the mean-line at two points: i) at the mean-point (all six squares); and ii) at one other point either below or above the horizontal line. You are to select any one of the data points, say, 'D' using the mouse pointer, and to drag it on the horizontal axis away from, or toward the mean-line. As you drag the point, square '*D*' will change according to the direction of dragging and the standard deviation of the data set will also change with dragging the point. You are to describe the changes in the square and the standard deviation as you drag the point. However, before you drag any of the data points, predict, and justify your prediction, how the square, and standard deviation will change as you drag the point.

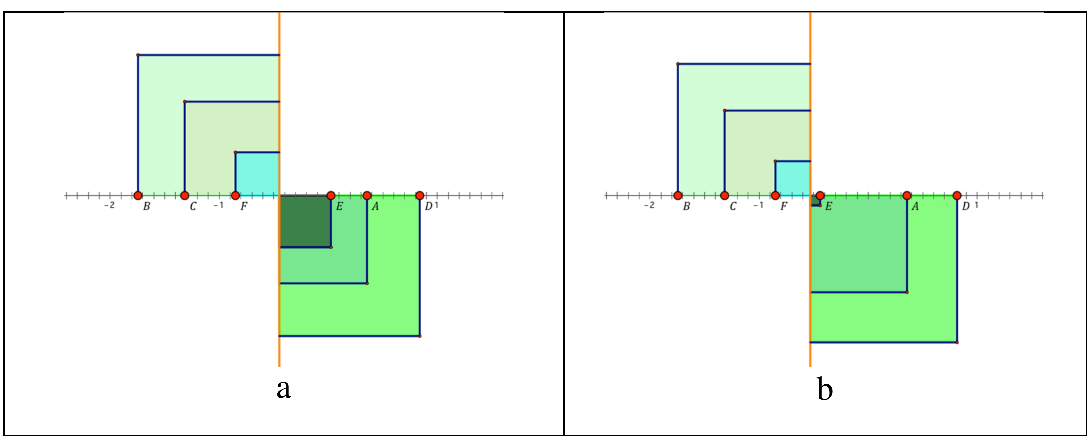


Figure B1. .The original design of dyMS sketch. a) dyMS sketch before dragging data point; b) dyMS sketch after dragging data point E toward the mean line from the right side and square area E gets much smaller.

Appendix C.

Interview Task 3:
Exploring Participants' Constructs of Variability with the gC Sketch

There are six numerical data points are marked on the horizontal number line and labelled B, C, F, E, A and D. The vertical purple line is the mean line for the six data points. A Gaussian curve very close to the horizontal line is fitted on the six data points and its peak as well the standard deviation of the data set will change as you drag a data point toward the mean-line or away from it on either side. You are to select one of the six data points using the mouse pointer and to drag it on the horizontal axis away from or toward the mean line. As you drag the point, the curve peak will change according to the direction of dragging. You are to describe the changes that you notice in the peak and the standard deviation as drag the points. Before you carry out the task, predict and justify your prediction, how the curve peak and standard deviation of the data set will change.

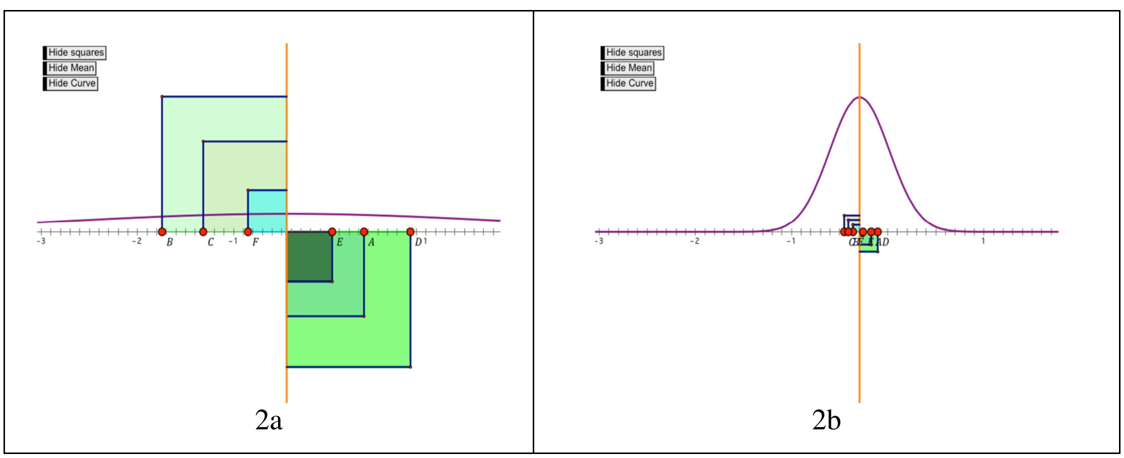


Figure C1. The original design of the gC Sketch. 2a) The gC sketch before data points are dragged much closer to the mean line; 2b) the gC sketch after data points have been dragged much closer to the mean line.