Fall 11-15-2018

LEARNING STATISTICS THROUGH GUIDED BLOCK PLAY: A PRE-CURRICULUM IN STATISTICAL LITERACY

Robert P. Giebitz
University of New Mexico - Main Campus

Robert Giebitz
Candidate

Organization, Information & Learning Sciences
Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

Patricia Boverie, Chairperson

Nick Flor

Karl Benedict

Sylvia Celedón-Pattichis

LEARNING STATISTICS THROUGH GUIDED BLOCK PLAY

A PRE-CURRICULUM IN STATISTICAL LITERACY

By

ROBERT GIEBITZ

B.S., Basic Sciences, New Mexico Institute of Mining and Technology, 1980
M.S., Management/Human Resources Management, Florida Institute of Technology, 1989

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
Organization, Information & Learning Sciences

The University of New Mexico
Albuquerque, New Mexico

December, 2018

Dedication

Knowledge is as wings to man’s life, and a ladder for his ascent. Its acquisition is incumbent upon everyone.

Endeavor to the utmost of thy powers to establish the word of truth with eloquence and wisdom and to dispel falsehood from the face of the earth.

– Bahá’u’lláh

Acknowledgements

My heartfelt gratitude to:

My committee chair Patsy for her continual support, guidance, and encouragement.

Sylvia, Nick, and Karl for serving on my committee and giving great feedback.

Bob for many cups of coffee and a ton of great music.

Lani, Greg, Ruth, and Cheryl each of whom had a special gift along the way.

Chris and Linda who cleared many obstacles and kept me on track.

Doris and Nora, partners in making the invisible visible with statistical tools.

Linda, Linda, and Theo, my Dalian friends who helped me get started.

My students at Dong Cai Da Xue who reminded me it was time to go back to school.

Especially my wife Amalia for years of unwavering support and encouragement.

LEARNING STATISTICS THROUGH GUIDED BLOCK PLAY

A PRE-CURRICULUM IN STATISTICAL LITERACY

By

Robert Giebitz

B.S., Basic Sciences, New Mexico Institute of Mining and Technology, 1980
M.S., Management/Human Resources Management, Florida Institute of Technology, 1989
Ph.D., Organization, Information & Learning Sciences, University of New Mexico, 2018

ABSTRACT

Learning to use data to investigate the world and make decisions has become an essential skill for all citizens. Play and curiosity are powerful motivators for learning. Inquiry – the process of asking questions and seeking answers – can engage the natural curiosity of young learners and motivate early learning. Recent research in statistics education has shown that children as young as 4 and 5 years old can learn to collect, organize, and interpret data they acquire through observation, counting, and measuring in a process of guided inquiry. Guided block play has been used for over 100 years to enable children to interact with mathematical structures paving the way for abstract understanding. Jerome Bruner conjectured that playing with a concept in concrete form prepares the mind for later abstract understanding and can begin at any age. Interaction with an embodied concept engages sensorimotor faculties and initiates neuronal activity that leads to useable knowledge grounded in experience. The frequency distribution is a core concept of statistics. Simple wooden cubes can be arranged on a ruler in the form of an embodied frequency distribution. This multiple case study explores how interaction with concrete representations of data structures in guided block play

can engage learners in grades K-2 and lay a foundation for understanding a data set as an aggregate with emergent properties of shape, spread, and center. Activity Theory provides a flexible theoretical framework for describing the interactions and explaining the outcomes of a series of exploratory tutorial sessions. It is further conjectured that this early experience with embodied learning enjoyed in the first years of formal schooling may prevent statistics anxiety and misconceptions in later years.

Table of Contents

List of Figures

List of Tables

Chapter 1: The Statistical Literacy Imperative

The ability to think and reason with data has become an essential life skill. Personal decisions concerning health, finance, and consumer choices; media claims; employment skills; preparation for science and engineering careers – all demand a basic understanding of data. According to a recent report commissioned by the American Statistical Association, "in an increasingly data-driven world, statistical literacy is becoming an essential competency, not only for researchers conducting formal statistical analyses, but for informed citizens making everyday decisions based on data. Whether following media coverage of current events, making financial decisions, or assessing health risks, the ability to process statistical information is critical for navigating modern society" (Franklin et al., 2015, p. 1). The rapidly growing fields of data science and analytics have opened up new career opportunities as well as raised deep concerns about the possibilities and potential for good or ill. Both the opportunities and the threats point to a common imperative: citizens need to be data literate and develop skills of statistical inquiry.

Data Literacy and Statistical Literacy

In 2015, the Oceans of Data Institute (ODI) convened a panel of experts from business, government, and education to define data literacy and outline the essential knowledge, skills, and behaviors needed in the emerging field of "Big Data" and how schools might develop data literacy in K-16 classrooms. They endorsed a definition of data literacy:

> The data-literate individual understands, explains, and documents the utility and limitations of data by becoming a critical consumer of data, controlling his/her personal data trail, finding meaning in data, and taking action based on data. The data-literate individual can identify, collect, evaluate, analyze, interpret, present, and protect data. (Oceans of Data Institute, 2015, p. 2)

ODI, together with IBM and the panel, launched a Global Data Literacy Initiative with a mission to promote data literacy as a global imperative. The report notes, "our current education systems have not been equipped to produce either the workforce or the citizenry with the skills, knowledge, and judgment to make wise use of the data streams that our technologies are delivering" (p. 14). So what are data?

In this study, data are the numerical outcome of a counting or measuring process. But data can also be in the form of text, images, places, or times. In all cases, data are socially constructed. Data are not “little nuggets of truth” (Best, 2004, p. xii) found in nature; they are rather the product of human activity. Best likens statistics to jewels: whereas gemstones are found in nature, jewels are created by people, “selected, cut, polished, and placed in settings to be viewed from particular angles” (pp. xii-xiii). Statistics are products of people's choices and compromises and they are colored by the biases, beliefs and values of those who decide what data to collect, how to go about it, and how to interpret them. As “Big Data” gets bigger and more embedded in our everyday lives, those who create the algorithms as well as those who are affected by them need a foundation in statistical literacy to understand the threats and the opportunities presented by a data-rich environment.

Statistical literacy empowers citizens “to make sense of real world messages containing statistical elements or arguments” (Gal, 2002, p. 4). Gal identified two components of statistical literacy in adults:

> (a) people's ability to interpret and critically evaluate statistical information, data-related arguments, or stochastic phenomena, which they may encounter in diverse contexts, and when relevant,
>
> (b) their ability to discuss or communicate their reactions to such statistical information, such as their understanding of the meaning of the information, their opinions about the implications of this information, or their concerns regarding the acceptability of given conclusions. (Gal, 2002, pp. 2-3).

Having the knowledge, skills, and dispositions to understand and to communicate using data and statistical arguments is the essence of statistical literacy. Watson (2006) proposed a set of interrelated core concepts of statistical literacy: “preliminary ideas of sampling, representation, summary, chance, inference, and variation” (p. viii) as a basis for decision making and evaluating statistical claims.

Pfannkuch & Wild (2004) developed a conceptual framework for statistical thinking in empirical inquiry from a historical perspective[i]. Statistical thinking involves understanding the “big ideas” behind statistical investigations. These include the ubiquity of variation, how and why to choose an analytical method, the nature of sampling and how inferences are drawn from samples to populations, how and why designed experiments are used to establish causation, and an understanding of context and underlying assumptions in a statistical investigation. Pfannkuch and Wild see statistical thinking as three interacting components: process thinking, understanding variation, and using data to guide actions. Statistical thinking is a mindset, a thought process needed for data-driven decision-making.

Gould (2017) maintained that data literacy and statistical literacy are synonymous. He proposed expanding the definition of statistical literacy in recognition of the dramatic rise of data science and the need for data scientists to be statistically literate. He proposed the definition of statistical literacy should contain at least the following elements:

- understanding who collects data about us, why they collect it, how they collect it;
- knowing how to analyze and interpret data from random and non-random samples;
- understanding issues of data privacy and ownership;
- knowing how to create basic descriptive representations of data to answer questions about real-life processes;
- understanding the importance of the provenance of data;
- understanding how data are stored;
- understanding how representations in computers can vary and why data must sometimes be altered before analysis; and
- understanding some aspects of predictive modeling. (Gould, 2017, p. 22).

This expanded notion of statistical literacy recognizes the dramatic increase in the impact of data on the lives of ordinary citizens and their need to have the knowledge, skills, and dispositions to access, store, analyze, interpret, and control data that pertains to their lives; to communicate and evaluate statistical arguments and evidence; and thereby to secure and maintain “a powerful voice in a democratic society” (p. 25).

Data-Driven Decision-Making and Dialogue

Data-driven decision-making is a structured process for generating useful knowledge from information derived from data, then using that knowledge to gain insight and understanding of the underlying processes from which the data were derived. Figure 1.1 illustrates this statistical inquiry process. Dialogue is the critical factor in using knowledge to make decisions. Dialogue is “a shared inquiry, a way of thinking and reflecting together … a living experience of inquiry within and between people” (Isaacs, 1999, p. 9). Dialogue exposes underlying assumptions to critical scrutiny (Bohm, 2004). Statistically literate participants in dialogue know how to use data to achieve new insights through collective inquiry.

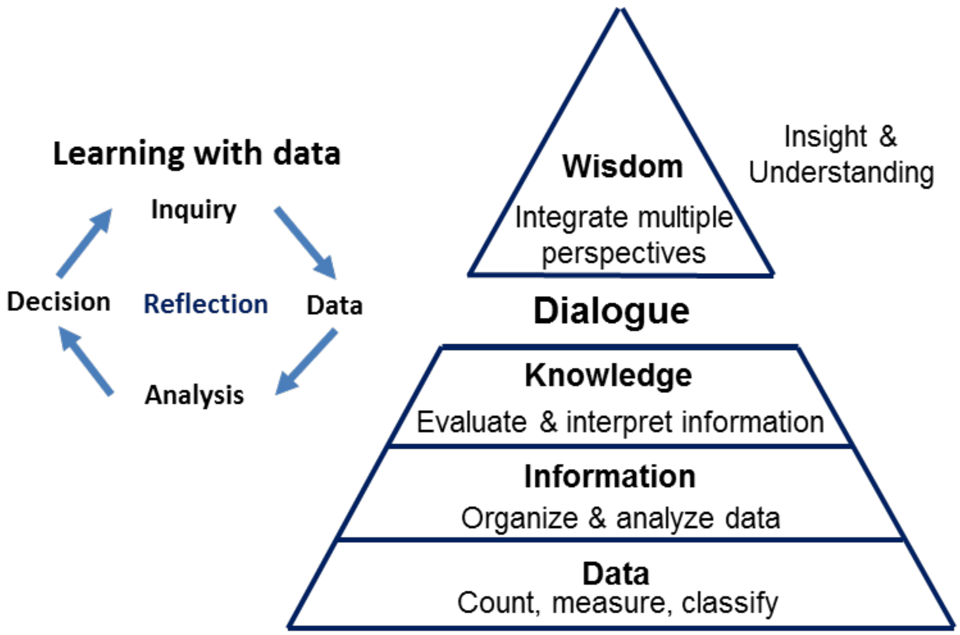


Figure 1.1. Data-driven decision-making entails counting or measuring, organizing the data to facilitate analysis, interpreting the information to generate knowledge, and finally deriving insight and understanding through dialogue to support wise decision-making.

Predictable biases are embedded in how human beings think (Tversky & Kahneman, 1974). Through dialogue, the impact of these biases can be minimized. Diversity of thought can provide new insights and access collective wisdom. “Intelligence emerges as the system connects to itself” (Isaacs, 1999, p. 63). Dialogue plays a vital role in reconciling personal and social perceptions of reality (Freire, 1970). Presenting data as pictures to accompany the numbers makes them more “digestible” in the process of collective inquiry and decision making.

Data Visualization

Visual techniques can be especially effective in rendering large data sets into a form the human mind can easily grasp. Florence Nightingale was an accomplished statistician[ii]. She was skilled at using visual displays of datasets to communicate. Her prowess in rendering statistical data into compelling visual images enabled her to convince the British Parliament to support her efforts to redesign army hospitals to eliminate sources of infection that were killing the soldiers. The polar area diagrams she devised during the Crimean War showed that deaths in hospitals from preventable causes outnumbered battlefield deaths ten to one (Figure 1.2). She used data to transform the healthcare system. Data visualization tools are not just graphic illustrations but are integral to the process of statistical thinking and good decision-making.

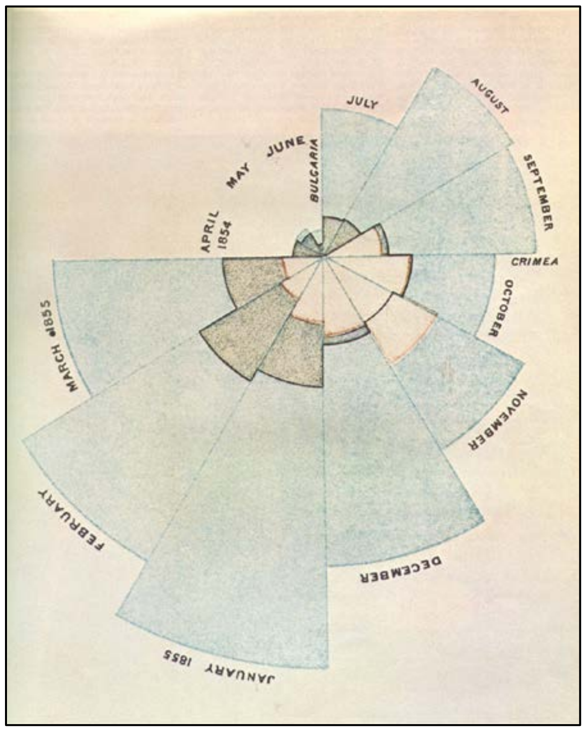


Figure 1.2. This polar area diagram, or "Nightingale Rose," shows deaths of soldiers each month from April 1854 to March 1855. The small inner sectors indicate battlefield deaths; the larger outer sectors indicate deaths due to preventable causes in hospitals.

Just as Nightingale's skill with data saved lives, lack of such skill can cost lives. Late at night, on January 27, 1987, flight managers were wrestling with the decision to launch or not launch the Space Shuttle Challenger on a cold morning in Florida. The engineers who designed and tested the booster rockets staunchly opposed the launch; they felt the rocket engines had not been adequately tested at low temperatures. They faxed tables and graphs and diagrams to the decision makers. But the data they sent did not clearly support their position; the data were either irrelevant or poorly organized. The extraneous documents the engineers faxed to the decision makers obscured the essential message behind the relevant data. The arguments of the engineers were unconvincing and were ultimately rejected. Consequently, seventy-three seconds into the flight, a critical failure allowed hot gases to escape past the O-rings. The escaping gases melted the casing on the rocket engine and ignited the fuel in an uncontrolled explosion. Seven astronauts died. The O-rings failed because they lost their resiliency due to low temperatures. Engineers and managers "agreed they had insufficient quantitative data to support an argument against the launch, were unable to frame basic questions of covariation among field variables, and thus unable to see the relevance of routinely gathered field data to the issues they debated before the Challenger launch" (Lighthall, 1991, p. 63). Participants in the fateful discussion leading up to the launch decision could not quantify the relationship between O-ring temperature and O-ring erosion, a simple statistical analysis for which data had been available for months and was available at the time of launch. Lighthall concluded "these failures of thought and perception were not from a lack of sophisticated expertise but from lack of simple, elementary, understandings and methods" (p. 73). Deficits in statistical literacy were a major contributing cause of the Challenger disaster.

Lighthall (1991) attributed the Challenger disaster to a systemic deficiency or "professional narrowness" and gaps in engineering education. Tufte (1997) claimed, "the consequences resulted directly from the *quality* of methods used in displaying and assessing quantitative evidence" (p. 5). In his analysis, Tufte focused on data visualization and the link between precise thinking, clear data displays, and rigorous scientific reasoning and analysis. "If displays of data are to be truthful and revealing," Tufte affirmed, "then the design logic of the display must reflect the intellectual logic of the analysis" (p. 31). Figure 1.4 shows Tufte's rendering of the data available before the launch. Tufte and Lighthall both affirmed the importance of statistics in engineering and management education. The ability to "see", to think, and to communicate statistically is critical in a complex decision process. Understanding the links between data quality, analytic methods, logic, and common sense is essential for effective data-driven decision-making.

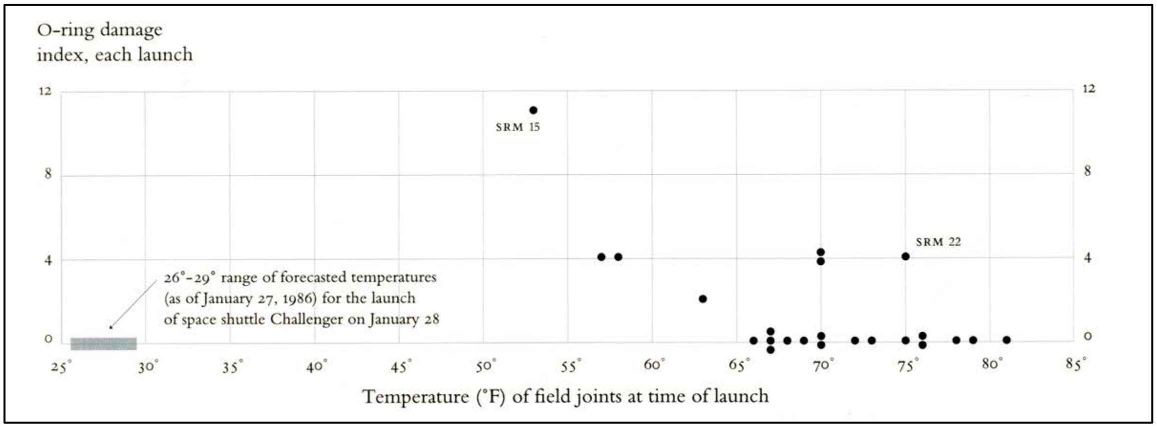


Figure 1.3. Edward Tufte's (1997) rendering of Challenger field data from previous missions. This scatterplot shows the relationship between O-ring damage and temperature at the time of launch. The data show clearly that a launch below 30 degrees would likely result in catastrophic O-ring damage.

Recent research has shown that children as young as 5 years old can collect and represent data visually (Leavy & Hourigan, 2018). These "representations are cognitive tools that give meaning to discovering, communicating and reasoning with data" (Soledad, 2018, p. 239). If children in kindergarten can begin to acquire the skills of statistical literacy and data visualization, why are these skills lacking in engineers and managers who have almost certainly had some instruction in statistics in their professional training?

The Challenge of Developing Statistical Literacy, Reasoning and Thinking[1]

Statistics education has largely failed to develop the ability of students to think and reason statistically (Delmas, Garfield, Ooms, & Chance, 2007). Statistical reasoning is "the way people reason with statistical ideas and make sense of statistical information" such as making sense of statistical summaries, interpreting and representing data sets (Ben-Zvi & Garfield, 2004, p. 7). It involves understanding statistical processes and interpreting statistical results. Introductory courses in statistics tend to focus on computations and rote procedures; students generally do not learn to think statistically. Many adults suffer from *statistics anxiety* and *negative attitudes toward statistics* – psychological dispositions that interfere with learning (Chew & Dillon, 2014; Garfield & Ahlgren, 1988; Onwuegbuzie & Wilson, 2003; Schau,

[1] This is the title of a book by Dani Ben-Zvi and Joan Garfield (Ben-Zvi & Garfield, 2004b).

2008; Schau & Stevens, 1995). Misconceptions students have at the beginning of a statistics course persist after formal instruction (Delmas et al., 2007; Hannigan, Gill, & Leavy, 2013; Huck, 2009). Even among researchers and editors of peer-reviewed journals, statistical errors and misconceptions are a persistent problem (Cohen, 1994; Sohn, 1991; Wilkinson, 1999).

Tversky & Kahneman (1971) investigated the erroneous intuitions about probability held by professional researchers. For example, researchers who otherwise have a high degree of statistical competence commonly consider a random sample to be representative of the population. This leads to an exaggerated belief in the likelihood that a replication of a study will confirm a previous finding. Tversky & Kahneman found that these erroneous intuitions are shared by naïve subjects as well as trained scientists. There is a widespread misconception within the research community that replication data should show a significance level close to the original study. Paradoxically, the same data can lead to opposite conclusions depending on whether it is viewed as an independent replication study or merged with the data of the original study. Statistical intuitions contain a cognitive or perceptual bias and this operates in the untrained as well as in the highly trained researcher. Although this has been known for decades, its impact has gone largely unchecked.

A “reproducibility crisis” is shaking the scientific community; many prominent published studies cannot be replicated (Baker & Penny, 2016). Underlying this crisis is a failure of statistics education at the highest levels (*Nature*, 2017). The problem is not a lack of statistics courses in the curriculum – students learn how to “do” statistics – but rather a failure to learn how to *think* statistically. Statistical literacy programs have increased worldwide in recent decades (Pullinger, 2013; UNESCO, 2006; United Nations Economic Commission for Europe, 2012)[iii]. But the magnitude of the task far outstrips available resources. Innovative strategies are needed to cultivate statistical literacy worldwide.

For decades, statistics education has been shifting emphasis from mathematical theory to data analysis (Tukey, 1962). There is growing awareness of statistical thinking as fundamentally different from mathematical thinking (Cobb & Moore, 1997; Wild, Utts, & Horton, 2018). Yet introductory courses still emphasize computation and hypothesis testing over engagement with real data. In response to deficiencies and incoherence in statistics curricula, the *Guidelines for Assessment and Instruction in Statistics Education (GAISE)* were developed to focus on statistical literacy and the "big ideas"[iv] of statistics like understanding variation[v] and data distributions (Franklin et al., 2005) . The *GAISE PreK-12 Report* offers a framework for statistical problem-solving consisting of four components: (1) formulate a question, (2) collect data, (3) analyze the data, and (4) interpret the results. In addition, *GAISE* makes six recommendations for tertiary level statistics education: "(1) emphasize statistical literacy and develop statistical thinking; (2) use real data; (3) stress conceptual understanding rather than mere knowledge of procedures; (4) foster active learning in the classroom; (5) use technology for developing conceptual understanding and analyzing data; and (6) integrate assessments that are aligned with course goals to improve as well as evaluate student learning" (p. 44). These guidelines represent decades of research. Yet academia lags industry in teaching and learning statistics. By looking at private industry, we can gain a better sense of how statistics creates value in the world through practical use.

The 1990s was a time of transformation in U.S. manufacturing. W. Edwards Deming was at the forefront of these sweeping changes. He was a physicist, engineer, statistician, and management consultant. During the reconstruction of Japan after WWII, Deming brought knowledge of statistical methods for controlling manufacturing processes (Shewhart, 1931). The systematic application of statistical quality control together with an enabling management philosophy (*kaizen*) was a key factor in transforming the Japanese economy. In less than three decades, Japan went from a producer of junk to a world leader in high quality automobiles and electronics. Western management had relegated quality to an ancillary function of production relying on mass inspection rather than on statistical methods. By the time Deming was recognized in the U.S. in the 1980's, Japanese manufacturers had captured a substantial share of the world market and that share was rapidly growing. Only then did Western companies wake up to the need for statistical methods to improve quality and reduce costs.

Deming was a staunch and relentless critic of Western management. He saw how common Western management practices systematically strip workers of their dignity and deprive the organization of the benefit of their creative capacities. Management, not the workers, he emphatically declared, is responsible for waste and low quality in manufacturing. To improve the efficiency and effectiveness of the manufacturing system, management must provide the workers with the tools, the training, the environment, and the autonomy to make decisions, solve problems and improve processes. Workers as well as management must know how to gather and use data to make decisions; they must be statistically literate.

Deming (1986) observed that executives often evaluate information on “averages only” without due regard for stratification or dispersion in the data. “The central problem in management, leadership, and production,” Deming wrote, “is failure to understand the nature and interpretation of variation” (p. 465). After Western management discovered that Deming was a force behind the surge in Japanese quality, his services were much in demand. Deming adapted the Red Bead Game (Figure 1.4) (The Deming Institute, 1980) – a playful enactment of statistical illiteracy – to teach management about variation in human systems. In this game, workshop participants take on various roles: workers, inspectors, supervisors,

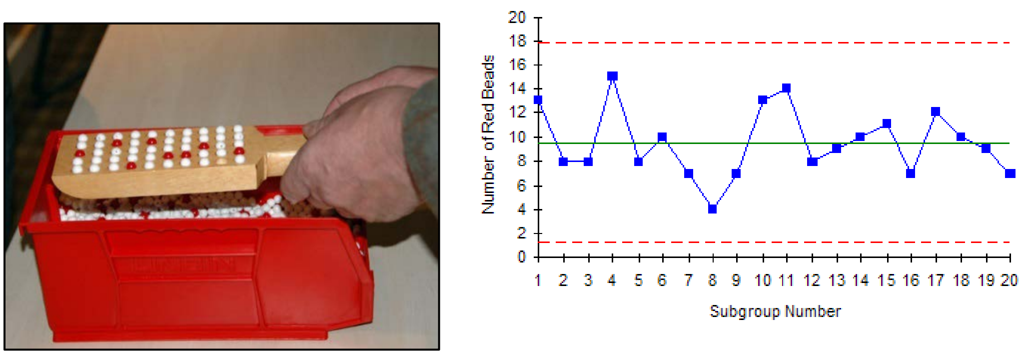


Figure 1.4. The Red Bead Game demonstrates how ignorance of variation in a system leads to poor decision-making and erroneous conclusions about cause and effect. The number of red beads shown in the chart varies randomly around the centerline.

and managers. The workers scoop the beads out of the bin and the inspector counts and records the red beads (defects). This is analogous to working in a production system with multiple interacting factors that influence process outcomes, such as equipment design and maintenance, ergonomics, lighting, noise, air quality, temperature, raw materials, training, and supervision. However, to improve outcomes, the *system* must be improved, and the design of the system is in the hands of management. No matter how the “worker” in the game manipulates the paddle to draw the beads out of the bowl, the outcome will always be a random draw. Efforts to decrease the number of red beads can only result in frustration. Only by changing the mix of beads in the bowl – the system – can the outcome change. There is a vital link between statistical literacy and effective leadership. Understanding how variation manifests in business processes of all kinds is essential for effective data-driven decision-making.

The example of manufacturing can be extended to education. Paulo Freire (1970) and Deming (1986) both recognized a mismatch between the process of human learning and how educational systems are designed and operated. Deming affirmed that as in work, joy must be a central feature of the learning experience. Freire (1970) maintained that education must free people from the strictures of existing social conventions and cultivate an awareness of the power that lies latent within each of us to transform social structures. Freire saw that educational systems tend to preserve unjust power and economic arrangements rather than liberate the tremendous creative potential latent in a population. He advocated a *problem-posing* education where dialogue is the central collective act that unveils reality, rather than a *problem-solving* education where problems are framed according to the prevailing assumptions and priorities of power elites. For example, rather than asking “How do we make cars more efficient?” we might ask “How do we make private ownership of cars unnecessary?” This is not just a different question, it is a fundamental shift in how the question is viewed.

Statistics education does not need to wait until adulthood. Research suggests that students would benefit from an early exposure to distributions and their features such as spread, center, and shape and the relationships among them (Garfield & Ben Zvi, 2007). Initial exposure to these concepts should be informal and these concepts should be treated in an integral fashion rather than as isolated topics. The “averages only” thinking common among adults may stem from the fact that they learned about averages years before learning about measures of central tendency (Watson, 2011). American Statistical Association (ASA) guidelines suggest using frequency distributions to describe variability and using manipulatives, for example, using cubes to represent data points. The ASA urges, “the foundations of statistical literacy must begin in the elementary grades Pre-K through grade 5” (Franklin et al., 2015, p. 18). These foundations should include a conceptual understanding of a data distribution and an ability to describe its key features and represent a dataset with tables, graphs, and numerical summaries. The National Research Council (2012) concluded, "deep, well-integrated learning develops gradually and takes time, but it can be started early: recent evidence indicates that even preschool and early elementary students can make meaningful progress in conceptual organization, reasoning, problem solving, representation, and communication in well-chosen topic areas in science, mathematics, and language arts" (p. 9). By looking at how young children learn statistics we might better envision a learning trajectory for older learners based on natural patterns of learning. This study explored these patterns in young children.

Transformation of Society

The transformation of industry offers lessons for the transformation of society. Freire (1970) maintained that every human being is capable of critical thought and of contributing to the betterment of the community and extending our collective understanding of the world. Leadership entails stewardship of a process of inquiry in which dialogue and collective reflection lead to action, further reflection on its consequences, and so on in a continual process of learning. To change the systemic structures of society, we need new ways of thinking – new ways of *framing* problems, not just solving problems that have already been framed according to old ways of thinking. Excessive emphasis on *solving* problems limits the creativity needed to see them in a new way and may simply reinforce prevailing modes of thought, wasteful patterns of behavior, and unjust social structures.

In addition to statistical literacy needed for continual improvement, lasting change calls for a systems perspective (Figure 1.5). Understanding *leverage* – maximum change with minimum effort – is a key principle of systems thinking. Change at the level of events or patterns of behavior without changing the underlying system will not have lasting results. For example, a vehicular homicide by a drunken driver is an event. The common practice of drinking to excess at parties then driving home is a pattern of behavior. Mothers Against Drunk Driving (MADD) has worked at the level of systemic structure to change attitudes through public advertising campaigns, tougher sobriety laws, increased enforcement, and a national legal drinking age of 21 (Loewit-Phillips & Goldbas, 2013). These efforts have paid off. The "National Highway Traffic Safety Association (NHTSA, 2004) revealed that deaths due to alcohol related causes had fallen from 30,000 in 1980 to 16,694 in 2002" (p. 62). Data driven decision making and systems thinking is a potent combination for sustainable systemic change.

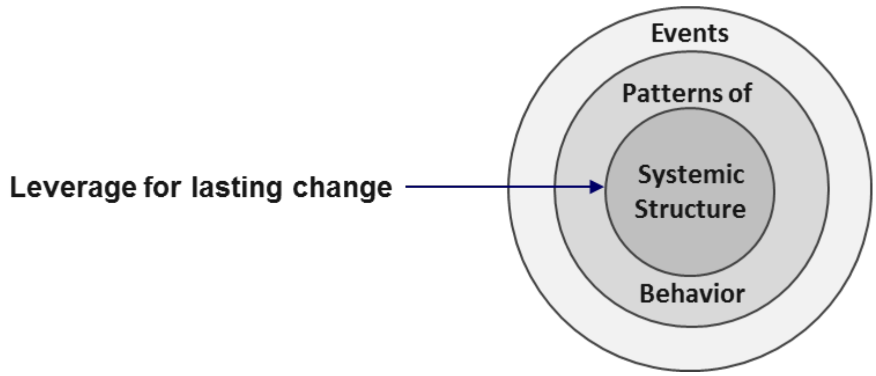


Figure 1.5. Leverage for sustainable change is found at the level of systemic structure, not in reacting to events or trying to change patterns of behavior directly (Senge, 1990), whether that system is a manufacturing enterprise, an educational system, or a global society.

Like Freire, Deming (1986) was a champion of universal education and universal participation in democratic processes. He believed that continual learning was essential throughout life. He was an outspoken and relentless critic of Western management practices that undermine the right of people to find joy and fulfillment in work. He decried the "appalling underuse, misuse, and abuse of skills and knowledge in the army of employed people in all ranks in all industries" (Neave, 1990, p. 10). Deming identified four essential capabilities of management: understanding patterns of interconnectedness within an organization or society (systems thinking); understanding human nature (psychology); understanding statistical variation (statistical literacy); and understanding the learning process (learning sciences).

The most outstanding reflection of Deming's work in Japan is the Toyota Production System (Ohno, 1978). Its chief architect Taiichi Ohno sought to create an organization of thinking people, a collaborative learning environment where increased production was achieved through developing the capabilities of people (Liker & Meier, 2007). The Toyota Production System was built through continual, systematic application of statistical techniques that reduce variation in manufacturing and other organizational processes while harnessing the capacity of every person in the organization to think critically, creatively, and collaboratively; to act on the outcome of their investigations; and to improve their individual and collective lives. Ohno's objective was to develop people to the fullest and to have them, in their turn, mentor others (Nakane & Hall, 2002). Ohno believed a leader continually challenges people to *think*. Ohno envisioned a culture of autonomous, thinking people. He set a pattern of inquiry and collaboration where each person was expected to use his or her own eyes and hands to see and touch and come to know the world as it is, including the world of automotive manufacturing, and figure out how to make it better.[vi] At Toyota during the time of Ohno, when a worker needed help to solve a problem, the supervisor or manager would go to the place of the problem, see with his own eyes, and get his hands dirty. Ohno is quoted as saying, "Don't look with your eyes, look with your feet. Don't think with your head, think with your hands" (Miller, 2008, p. 2). Ohno realized that being physical in the context of the problem was essential to understanding; no amount of processing information from a distance could get at the core of the problem. This practice of learning through direct contact with the world using all one's senses is not only a crucial aspect of solving problems and improving processes in the workplace, it can also be employed in learning fundamental statistical concepts and can begin in the earliest years of primary education (Franklin et al., 2015).

Learning through Movement, Play, and the Use of the Hands

Movement, perception, and learning are inextricably linked; the hand plays a central role in the emergence of symbolic thought (Wilson, 1998). Freedom of movement is the essence of our sense of agency. Self-movement structures our knowledge of the world; it is a way of knowing; "our tactile-kinesthetic bodies are epistemological gateways" (Sheets-Johnstone, 2011, p. xxv). Movement underpins conceptual understanding; cognition is grounded in sensorimotor experience (Barsalou, 2008; Clark, 1997). As the primate hand co-evolved with the brain, new possibilities emerged through interactions between "visual, tactile, and proprioceptive feedback on the same action system" (Merlin Donald quoted in Wilson, 1998, p. 46-47). Experience, knowledge, thought, emotion, and practice are all one process (Bohm, 2004).

Bruner (1964) conjectured that if children were afforded the opportunity to play with an abstract concept in some sort of concrete form, it would prepare their minds for later conceptual understanding. Vygotsky (1978) recognized the enormous influence of play in the development of the child. In this “realm of spontaneity and freedom”, the relationship between perception, motivation, and meaning achieves greater coherence. Play induces internal transformations in the child; it creates a *zone of proximal development* where budding mental processes mature. Through play, abstract thought emerges.

Vygotsky (1978) recognized that interaction with others, particularly cooperative interaction with an adult or more capable peer, is essential for the development and learning of the child. Learning occurs through participation with others in a meaningful context (Lave & Wenger, 1991). Choice and learner autonomy are essential elements of the learning process. “Learning awakens a variety of internal developmental processes that are able to operate only when the child is interacting with people in his environment and in cooperation with his peers” (Vygotsky, 1978, p. 90).

Block play has a long history in education (Fröbel, 1887; Hewitt, 2001; Read, 1992). Block play can be both playful and purposive. In order to master the structure of knowledge, subject matter must be intentionally organized to enable students to follow a learning trajectory “from qualitative understanding to more precise quantitative understanding” (National Research Council (U.S.), 2005, p. 15). The organization and delivery of the content must align with *how people learn* (National Research Council (US), 1999). Blocks have been used for over a century to teach mathematics to young children (Cuisenaire & Gattegno, 1962; Dienes, 1964; Gattegno, 1961; Montessori, 1912). Dienes sought to stimulate preverbal mathematical thinking using mathematically structured physical materials and so gain insight into the process of progressing from preverbal to symbolic forms of thought. However, there has not been a comparable effort using blocks to teach statistical thinking.

Exploring a Pre-curriculum in Statistical Literacy

This study explored how statistical literacy might begin with block play. It explored how the manipulation of blocks with appropriate scaffolding (Wood, Bruner, & Ross, 1976) might lay a foundation for developing statistical skills. Dewey (1910) observed, “the native and unspoiled attitude of childhood, marked by ardent curiosity, fertile imagination, and love of experimental inquiry, is near, very near, to the attitude of the scientific mind” (p. iii).

Following Bruner's (1964) suggestion that concepts be introduced first as playthings to prepare the mind for later conceptual understanding, and the recommendation of the ASA that a foundation for statistical literacy begin in the early grades, this study explored the capabilities of children ages five to eight. A pilot study was conducted to develop and test an apparatus designed to introduce a preliterate child to a frequency distribution. She had not yet learned computational skills but was able to read and write numbers from one to twenty and to read the scale on a ruler. The pilot study showed that a kindergarten student can engage an embodied data structure in a systematic and joyful way, find numerical statistics without computation, and create a variety of data representations by following simple procedures under the guidance of a tutor. The pilot study also identified analytic themes for the main study.

Children are capable of much more than adults generally recognize (Powell, Francisco, & Maher, 2003). Recent research has called into question prevailing notions of “readiness” and “developmentally appropriate” (National Research Council (US), 2001). Recent studies have also found “strong evidence that children … have the ability to abstract well beyond what is ordinarily observed” and the National Research Council calls into question “simplistic conceptualizations of developmentally appropriate practice that do not recognize the newly understood competencies of very young children” (p. 5).

In this exploratory multiple case study (Merriam, 1998; Stake, 2006; Yin, 2014), children played with an embodied frequency distribution (blocks on a ruler) and found measures of shape, spread, and center and created X-plots, marble plots, and box plots. They found the average by balancing the ruler on a fulcrum. Then through direct observation, counting, and manipulation of the blocks, they found minimum, maximum, median, range, first and third quartiles, and interquartile range. In a more complex activity, two of the students found the mean absolute deviation (conceptually equivalent to the standard deviation). Neither literacy nor mathematics was required. Understanding how children engage with and learn these concepts may shed light on how adults learn and provide some insight into how the anxiety, negative attitudes, and misconceptions that often afflict adults might be overcome or avoided entirely. Bruner suggested that learning a new concept begins with *enactive* representation, transitions to *iconic* representation, and finally to *symbolic* representation. *Driscoll* asks, “Might not adults as well as children, pass through the same sequence of enactive to symbolic representation when they learn a subject for which they have no prior experience?” (Driscoll, 2000, p. 226).

Merriam (1998) presents a three-level framework for a case study inquiry. A problem statement frames the logic of a case study. The problem statement leads to a purpose statement that shows how the phenomenon of interest (statistical variation and distribution) relates to the larger problem (the need to develop statistical literacy throughout society). The purpose of the study gives rise to the research questions. This framework is shown in Figure 1.6.

Problem Statement: Statistical literacy is an essential capability for informed citizenship and scientific inquiry. Yet people generally do not learn to think and reason statistically. Misconceptions and anxiety commonly interfere with the teaching and learning of statistics in adulthood.

Purpose Statement: The purpose of this study is to explore how facilitated playful engagement with a frequency distribution embodied in physical objects might provide a sensorimotor grounding for an understanding of variation and distribution.

Research Questions:

How might play with an embodied frequency distribution in the form of blocks arranged on a ruler, under the guidance of a tutor, facilitate learning statistical concepts and skills beginning in the first years of formal education?

What sequences of tasks and activities might comprise a learning trajectory toward conceptual understanding of data distribution and variation?

How might these findings inform the design of learning experiences for adults and instigate innovations in statistics education at all ages?

Figure 1.6. A Three-Level Framework for a Case Study Inquiry (Merriam, 1998).

Chapter 2: Literature Review

In this chapter, I present John Dewey's theory of logic, inquiry, and knowledge as a context for the use of statistics. Both common sense inquiry and scientific inquiry have biological and sociocultural roots. Learning requires agency, participation, imagination, and reflection. Learning is not a phenomenon pertaining only to a brain, but rather to a whole tactile-kinesthetic organism interacting with a physical, social, cultural, and historical environment. Cognitive systems are distributed, situated, and embodied. The human hand plays a crucial role in the learning process as it has in the evolutionary history of *hominids.* Coordinated actions of the hand on objects such as blocks together with the associated neural circuitry create metaphors from which conceptual systems are built, a process facilitated when pursued in a state of play or flow. Principles of embodied learning have been applied to teaching and learning mathematics but rarely to statistics. Statistics is one of the most problematic subjects for adults to learn yet one of the most useful. The natural learning processes of children have been subjected to fewer disturbances than those of adults and can serve as a window into how learning occurs at all ages. Activity Theory provides an apt conceptual framework for studying the systemic nature of learning.

Inquiry, Knowledge, and Logic

Dewey's (1938b) theory of inquiry provides a foundation for a cognitive science grounded in biology, evolution, and culture. His theory of logic differs from that of the Greeks who elevated Reason above practical experience. Dewey points out that the ruling class of ancient Greece had a stake in elevating abstraction above concrete forms creating a two-class divide: those who *know* and those who *do*. The leisure class enjoyed both ontological and epistemological privilege over the "practical" class of craftsmen and artisans. Dewey concluded that these cultural conditions prevented "the utilization of the immense potentialities for attainment of knowledge" afforded by the artisan and working classes. In this, Dewey saw a paradox: these capabilities of craftsmen and artisans were precisely what was needed to develop the observational and measuring instruments of experimental methods for generating knowledge about the world – a knowledge that could modify existing conditions "instead of being subordinated to a scheme of uses and enjoyments controlled by given socio-cultural conditions" (p. 59). The use of knowledge to sustain societal structures of privilege was not unique to the Greeks, but it was the Greeks who formalized a theory of logic and an epistemology that supports such structures and has survived into the 21st century. Dewey maintained that the rules of logic derive from an inquiry process based on human experience; logic does not exist independently of human consciousness in a realm of Pure Reason as Kant (1998/1787) suggested.

Dewey (1938b) saw how a formulation of logic inherited from ancient times constrained scientific inquiry and, conversely, how breaking away from such a conception of logic would free up the tremendous untapped creative potential in a population and foster the *democratization of knowledge*. He sought to bring the theory of logic more in line with scientific practice and to stimulate broad participation in scientific discourse. Dewey understood that inquiry must satisfy logical requirements to reach valid conclusions, but these logical forms have their origin in the process of inquiry, not vice versa. Logic provides constraints on the process of inquiry to help ensure that its conclusions are warranted and useful in further inquiry.

The Process of Inquiry Has Biological Roots – Logic is Rooted in Culture

Statistical literacy extends our powers of inquiry. Dewey (1938b) maintained that our powers of inquiry emerged from biological patterns of behavior and relations between people in a specific cultural context. Biological factors – eyes, ears, hands, brains – are essential constituents of inquiry, but a consequence of the mind-body duality is that these physical factors are "shunted off" to a separate domain and treated as a separate metaphysical or epistemological problem. Thus, logical theory developed on the rational side of this artificial divide reinforcing and legitimating the social arrangements of power and privilege from which it arose. The prevailing notion that Reason accounts for the appearance of the process of inquiry in humans is, in Dewey's view, an invocation of the supernatural. Life processes – including processes of inquiry – are enacted by the environment as much as by the organism – they are an integrated whole.

Vygotsky (1978) agreed that inquiry and logic have biological and sociocultural roots and are not a product of disembodied Reason. He identified two distinct, yet interweaving, lines of human development: elementary processes of biological origin, and higher psychological functions of sociocultural origin. Vygotsky advocated combining experimental cognitive psychology with neurology and physiology in a historical context of human society to develop a sociocultural theory of higher mental processes. He recognized that historical context is critical to understanding both the development of individual capacities and emergence of those capacities within a social context. Logic is subject to the influences of culture, including its dominant philosophy and values.

Dewey (1938b) observed that Western thought has largely been blind to its own underlying philosophical assumptions allowing logical theory to serve metaphysical and epistemological preconceptions and a particular set of cultural values. Kuhn (1962) drew on the evidence of history and contemporary scientific practice to refute the assertion that science is a values-free enterprise. Freire (1970) saw that the structures of society and prevailing modes of thought perpetuate mechanisms of oppression and injustice. Within these societal structures, the privileged enclose themselves in "circles of certainty" insulated from doubt while the marginalized live in a "culture of silence" that limits freedom, autonomy, and the power of choice and is maintained chiefly by the educational system. The "oppressor consciousness" uses science and technology to advance a materialistic conception of reality and cast the masses as objects whose purpose is defined by believers in the dominant culture. True education is a process of inquiry that builds trust, encourages self-examination, and fosters reflective participation leading to collective action. According to Freire, powers of inquiry are the birth right of all people; developing these powers should not be a special privilege of a few. Inquiry is liberating and empowering; knowledge is both individually and socially transformative.

Knowledge is the Outcome of an Inquiry Process

Powers of inquiry, including statistical inquiry, ought to be integral to a culturally responsive education not a specialty of "experts." All people need to be able to assess and critique the outcome of inquiries that inform public policies and affect their daily lives and aspirations. Different cultures have had diverse conceptions of knowledge as well as diverse modes of inquiry. Dewey (1938b) reasoned, "Since every special case of knowledge is constituted as the outcome of some special inquiry, the conception of knowledge as such can only be a generalization of the properties discovered to belong to conclusions which are outcomes of inquiry. *Knowledge*, as an abstract term, is a name for the product of competent inquiries" (p. 8). Inquiry is a continuous process:

> The attainment of settled beliefs is a progressive matter; there is no belief so settled as not to be exposed to further inquiry. It is the convergent and cumulative effect of continued inquiry that defines knowledge in its general meaning. In scientific inquiry, the criterion of what is taken to be settled, or to be knowledge, is being so settled that it is available as a resource in further inquiry; not being settled in such a way as not to be subject to revision in further inquiry (Dewey, 1938b, pp. 8-9).

Dewey (1938b) developed a set of guiding logical principles that provide direction to the inquiry process and principles for testing its conclusions.

1. Logic is a progressive discipline – logic rests upon analysis of methods of inquiry that have produced and continue to produce results; logic has no final formulation.
2. The subject matter of logic is determined operationally – methods of inquiry are operations performed on things or symbols; logical forms are the conditions that inquiry must meet.
3. Logical forms are postulational – inquiry must satisfy requirements that are not prior to and independent of inquiry, but rather postulates discovered in the course of inquiry which further inquiry must satisfy in order to yield warranted assertions.
4. Logic is a naturalistic theory – there is a continuity between operations of inquiry, biological operations and physical operations.
5. Logic is a social discipline – every inquiry grows out of a background of culture.
6. Logic is autonomous – it precludes the assumption of a prior definition of knowledge which determines the character of inquiry; knowledge is defined in terms of inquiry, not vice versa (Dewey, 1938b, pp. 14-21).

Dewey conceptualized inquiry as an organic process of knowledge creation underlying both common sense and science. It evolved through interaction of biological factors operating in a physical and social environment. His conception of inquiry avoids mere speculation, invoking the supernatural, or privileging the metaphysical assumptions of a particular class or culture. Inquiry liberates us from superstition, misconceptions, erroneous beliefs, and false assumptions. It both strengthens our sense of agency and equips us to engage with others in knowledge creation; to engage in dialogue on the implications of that knowledge and its application in research, decision-making, and collective action. Statistical inquiry is a more recent evolution of this process contributing powerful quantitative methods and the use of data to arrive at knowledge claims. These new capabilities expand our sense of agency and capacity for knowledge creation.

Science and Common Sense Share Common Ground

There is a "fundamental unity of the structure of inquiry in common sense and science" (Dewey, 1938b, p. 79). Whereas common sense may vary across cultures, science can be a bridge between cultures. Common language and other symbol systems derive their meaning from group experience and interests, habits and customs, and institutions, not from examined relationships among the elements of the system. Scientific language, on the other hand, builds on relationships among the elements of the system and seeks an internal coherence that reaches beyond cultural distinctions. Scientific communities establish norms of language and symbols: "The ideal of scientific-language is construction of a system in which meanings are related to one another in inference and discourse and where the symbols are such as to indicate the relation" (pp. 50-51). Science demands a higher standard of semantic clarity and precision, including tools and language to assess and express uncertainty. Uncertainty is expressed in stochastic language – the language of statistics and probability.

Scientific inquiries seek factual knowledge, natural laws, and theories. The domain of common sense inquiry, on the other hand, is "problems of use and enjoyment" (Dewey, 1938b, p. 61). Scientific inquiry grows out of problems and methods of common sense inquiry while it refines and expands the domain of common sense; it culminates in collective action and discovery. The consequences of collective action then serve as a resource for further inquiry and so on through continuing cycles of learning. Dewey notes that despite great strides in the development of experimental science, there has been little corresponding advance in human relationships. "Morals and the problems of social control are hardly touched. Beliefs, conceptions, customs and institutions, whose rise antedated the modern period, still have possession of the field … The paths of communication between common sense and science are as yet largely one-way lanes. Science takes its departure from common sense, but the return road into common sense is devious and blocked by existing social conditions" (p. 77). These social conditions include structural barriers to entry into scientific and technical fields by members of non-dominant cultures effectively maintaining their marginalized status and impeding the progress of science. Cultivating statistical inquiry skills throughout a society is one strategy to break these social barriers.

Science is Packaged in a Cultural Wrapper

Inquiry is situated – it is organic to social and environmental conditions; it is not an act of an isolated organism (Dewey, 1938b). Subjectivist assumptions about the organism, the environment, and their interaction as three independent factors destroys the link between inquiry as reflective thought and as scientific method. Their integration is more fundamental than their distinction. Similarly, culture is not created by an individual but rather emerges from the interactions among individuals. A natural pattern of inquiry is foreshadowed by a spatiotemporal pattern of life. "In a proper conception of experience," Dewey maintains, "inference, reasoning and conceptual structures are as experiential as is observation … the fixed separation between the former and the latter has no warrant beyond an episode in the history of culture" (p. 38). Dewey wondered how "the development of organic behavior into controlled inquiry brings about the differentiation and cooperation of observational and conceptual operations" (p. 39). He saw that relationships between people and the creation of shared meaning within a culture are essential to this process. "Problems which induce inquiry grow out of the relations of fellow beings to one another, and the organs for dealing with these relations are not only the eye and ear, but the meanings which have developed in the course of living, together with the ways of forming and transmitting culture with all its constituents of tools, arts, institutions, traditions and customary beliefs" (p. 42). A wider embrace of these constituents of culture would enrich all the domains of science. When the tools of statistical inquiry are more widely shared across cultures, they might become both culturally and conceptually more accessible. This would entail considerable rethinking of how statistics is taught.

Indigenous communities, given their numbers, their historical systematic exclusion from scientific discourses and the unique contribution they would make to such discourses, merit special attention (United Nations General Assembly, 2015). Kidman, Abrams, and McRae (2011) found that Māori learners were disengaged from science learning, even when taught in their native language and in an environment that reflected their cultural values – the use of Indigenous language does not automatically evoke an Indigenous worldview or reflect Indigenous knowledge. They found that criteria for making curricula and pedagogical practices more culturally-responsive and relevant to students are determined not *by* students but *on behalf of* students. These unequal knowledge-power relations contributed to Indigenous students deciding to disengage from the sciences taught in formal education. In many cases, culturally responsive curriculum is no more than "tokenistic activities designed to 'celebrate' ethnic diversity" (p. 204). The system of rules that govern the ways knowledge is selectively drawn from a field and then constructed, circulated, regulated, evaluated and reproduced for pedagogical transmission and assessment is a product of a Western worldview. Kidman et al. observed that the knowledge produced by scientists and the discourses that animate the various fields of science are repackaged for consumption in schools thus losing their vitality and potential appeal to students. This re-packaging is driven as much my ideological concerns as by educational priorities (echoes of Freire). Duschl (2008) identified this "rhetoric of conclusions" with an antiquated *content-process* curriculum orientation and contrasted it with the *discovery-inquiry* approach developed during the science curriculum reform movement in the U.S. during the 1950s and 1960s. Bruner (1960) was Director of that effort "to examine the fundamental processes involved in imparting to young students a sense of the substance and method of science" (p. vii). This problem of the "middle language," as Bruner called it, is that textbooks "talk about the conclusions in a field of intellectual inquiry rather than centering upon the inquiry itself" (p. 14). Similarly, in much of statistics education, the focus is on drawing conclusions using statistical techniques and not on the process of statistical inquiry. There is wide agreement within the statistics education research community that statistics education should begin with inquiry (Arnold, Confrey, Jones, Lee, & Pfannkuch, 2018; J. Garfield & Ben Zvi, 2007) and that the inquiry should be relevant to the sociocultural context of the learners.

As the learning sciences have advanced in recent decades the distinction between *discipline-based* and *learning-based* epistemic frameworks have become "critically important for situating school science learning, knowing, and inquiry" (Duschl, 2008, p. 274). Freire (1970) saw the educational system as a major instrument for maintaining a “culture of silence” while the dominating forces in society use science and technology as instruments of oppression. Such structural arrangements cast science as “a gift bestowed by those who consider themselves knowledgeable upon those whom they consider to know nothing” (p. 53). After over a decade leading the U.S. effort to redesign science curriculum, Bruner (1971) concluded that the underlying issue in education is not so much curriculum redesign as it is empowerment of the disenfranchised.

Inquiry-Based Teaching and Learning

The term “inquiry-based” has had a wide range of meanings. In a meta-analysis of the literature, Furtak, Seidel, Iverson, & Briggs (2012) found that researchers have been inconsistent in operationalizing the inquiry construct. They examined experimental and quasi-experimental studies on the effectiveness of inquiry-based science teaching published between 1996 and 2006 where effect sizes were derived from pre-post assessments of two-group designs. Based on this analysis, they proposed a two-dimensional framework comprising a *guidance* dimension and a *cognitive* dimension. They defined inquiry in terms of 1) cognitive and social activities of the student, and 2) guidance provided to the student. They adopted the three domains of Duschl's (2008) *conceptualization of inquiry-based teaching and learning*: 1) *conceptual structures and cognitive processes* (*e.g*., facts, theories, principles), 2) *epistemic frameworks* (knowing how scientific knowledge is generated), and 3) *social interactions* such as processes of collaboration and communication by which knowledge is constructed (Grandy & Duschl, 2007). Furtak et al. (2012) added a *procedural* sub-category under Duschl’s epistemic domain. Within the *procedural* category, they placed activities such as asking research questions, designing experiments, executing procedures, and creating data representations. For the *guidance* dimension, the authors adopted a continuum of directedness across the essential features of inquiry (National Research Council, 1996). They found a positive effect of inquiry-based science teaching on student learning, particularly when students were engaged in the epistemic domain of inquiry, but also in the combined procedural, epistemic, and social domains (p. 324). They found higher effect sizes for teacher-led activities and identified a need for more refined models of inquiry-based teaching and learning capable of resolving the conflicting findings of previous studies. A pre-post research methodology enables the researcher to compute effect sizes but reveals very little about *why* an intervention has the observed effects. The present study uses *microgenetic* methods to examine what happens in between the “bookends” of pre-post assessments.

Statistical literacy, as a gateway to science and engineering fields, is subject to the same constraints blocking access to its knowledge-enabling potential. A fundamentally different approach to teaching and learning statistics might contribute to facilitating broader access to the tools and skills of collective knowledge creation. However, real progress in creating culturally responsive learning environments might also call for professional development efforts toward greater awareness of the dominant epistemological orientation embodied in mainstream Western science and the need to develop capacity in both teachers and students to navigate multiple epistemologies (Bang & Medin, 2010). This study explored this two-fold perspective considering methods that are in harmony with how people naturally learn and recognition of the cultural factors that affect development of the capabilities of scientific inquiry. In the following sections, we will explore the anthropological and evolutionary antecedents of human learning and cognition to gain a broad perspective on how to foster inquiry skills and make statistics education more culturally responsive.

Movement, Cognition, and Thinking with the Hands

Human powers of inquiry evolved from more primitive powers of movement and sensation. “Self-generated movement is the foundation of thought and willed action … a critical element at the core of all learning” (Wilson, 1998, p. 291). Movement “makes possible all perceptual categorization” (Oliver Sacks quoted in Wilson, 1998, p. 208). “Movement … is the generative source of our notions of space and time. … our tactile-kinesthetic bodies are epistemological gateways” (Sheets-Johnstone, 2011, pp. xvii-xxv). Contemporary cognitivist science “inordinately favors brains to the exclusion of the animated realities of living creatures” (p. xxix). Such a perspective devalues the importance of movement in the process of concept formation. However, *encephalocentrism* – the idea that the brain is the seat of all consciousness, sensation, and learning (Crivellato & Ribatti, 2007) – is being challenged by contemporary neuroscience and philosophy of mind.

Kinesthetic Consciousness and Symbolic Thought Evolved through Movement

Emergent kinetic capabilities arising from changes in hominid morphology in the transition from Neanderthals to anatomically modern humans afforded new possibilities for analogical thinking since “analogical thinking is foundationally structured in corporeal representation” (Sheets-Johnstone, 2011, p. 13). Animate bodies are semantic templates and there is a “dynamic congruency” between symbolic behavior and its referent. For example, stone tools are not *symbols*, but are rather analogues of body parts – teeth, arms, fists – not necessarily an outcome of reflective acts, but “embedded in pre-reflective corporeal experience” (p. 16). “Animate form,” Sheets-Johnstone maintains, “is the proper starting place for paleoanthropological reconstructions” (p. 21). There is a kind of “bigger brains” bias in traditional Western anthropological accounts. Sheets-Johnstone disputes a tenet of Stephen Jay Gould’s theory of punctuated equilibrium that considers the “uniqueness” of *Homo sapiens* to be essentially mental and brain-based. According to Sheets-Johnstone, Gould maintained that “in order to illustrate a concept” one moves the body. However, she points out that “if movement can *illustrate* a concept, then might not movement *generate* a concept in the first place” (p. 26). Cartmill, Beilock, & Goldin-Meadow (2012) found that gestures “can instill new ideas in learners – creating thought in addition to reflecting it. … Gesture’s ability to represent action offers a way to ground abstract ideas in concrete actions. … Representing action in gesture embeds embodied information into mental representations of action” (p. 134). Sheets-Johnstone (2011) again affirms the potential for movement to create a firm grounding for concept formation: “movement possibilities and dispositions delimit one’s conceptual possibilities and dispositions” (p. 26). Thinking is modeled on a tactile-kinesthetic body; movement is not simply a change in bodily position in response to some mental directive. “Animation is at the very origin of consciousness” (p. 128).

Vygotsky (1978) observed in the child a kind of ontogenic recapitulation of this anthropologic perspective. He concluded, “the child’s system of activity is determined at each specific stage both by the child’s degree of organic development and by his or her degree of mastery in the use of tools” (p. 21). The intentional use of the hands – the tactile, perceptual, baric experience with physical objects in coordinated action – may be as necessary to building a conceptual framework as it is to building a physical structure. If proprioception is an *epistemological gateway*, then by accessing proprioceptive channels in the body, we are also opening channels of learning.

Consciousness is not found in matter, but rather is a dimension of moving animate forms; it evolved along with living, moving creatures. Mechanistic accounts of mind that reduce living bodies to neurophysiological matter inside of heads misconstrue the nature of living creatures. Sheets-Johnstone (2011) contends that the evolutionary history of proprioception paralleled the emergence of consciousness. Cognitivist reductions of consciousness to neurophysiological states and computational models are misguided (Nagel, 1974). Cognition can be understood "as *embodied action* … within the context of evolutionary theory" (Varela, Thompson, & Rosch, 2016, p. lxvi). Behavior is not regulated by some sort of "mental code" operating behind the scenes that has somehow arisen independent of movement. The close coupling between action, perception, and cognition that develops in infancy is retained throughout adulthood and serves as the matrix within which reason, memory, emotion, language, and other higher mental functions manifest (Thelen, 2000), including statistical reasoning.

We need to think about intelligent behavior holistically. It is time "to abandon the idea of neat dividing lines between perception, cognition, and action; to abandon the idea of an executive center where the brain carries out high level reasoning; and most of all, to abandon research methods that artificially divorce thought from embodied action-taking. … Treating cognition as pure problem solving invites us to abstract away from the very body and the very world in which our brains evolved to guide us" (Clark, 1997, pp. xii-xiii). "Minds evolved to make things happen," Clark declares, they are not "disembodied logical reasoning devices. … Intelligence and understanding are rooted not in the presence and manipulation of explicitly, language-like data structures, but in something more earthy: the tuning of basic responses to a real world that enables an embodied organism to sense, act, and survive" (pp. 1- 4). Learning is not "an isolated process of information absorption," but is rather "a cultural and social process of engaging with the constantly changing world around us" (Thomas & Seely Brown, 2011, p. 47). Learning also draws on play and imagination. A robust theoretical model of the learning process needs to account for the embodied nature of cognition, social and cultural aspects of learning, as well as the interactions within the learning experience itself.

The Hand Played a Decisive Role in the Evolution of Cognition

Gestures both reflect and shape our mental representations and processes (Cartmill et al., 2012). Gesture "played an important role in the transition to symbolic thought and language in human evolution" (p. 129). The hand and its control mechanisms evolved "as prime movers in the organization of human cognitive architecture and operations" (Wilson, 1998, p. 286). From the evolutionary perspective, "the hominid hand and its growing repertoire of movements were integral to what was happening in behavioral, cultural, and cognitive evolution … the hand speaks to the brain as surely as the brain speaks to the hand" (pp. 290-291). Vygotsky (1978) maintained that to understand the system of activity in the learning process, we must include "systematic movement and perception, the brain and hands – in fact, the child's entire organism" (p. 21). Merlin Donald (cited in Wilson, 1998) notes, "hand control involves, for the first time in evolution, a coming together of visual, tactile, and proprioceptive feedback on the same action system" (p. 46-47).

The relationship between movement, perception, and cognition is essential to an understanding of human learning. Wilson (1998) declares that any theory of human intelligence must account for "the interdependence of hand and brain function, the historic origins of that relationship, [and] the impact of that history on development dynamics in modern humans" (p. 7). Wilson explored the role of the hand in the emergence of symbolic thought from three perspectives: the anthropological and evolutionary perspective, the biomechanical and physiological perspective, and the neurobehavioral and developmental perspective. The development and refinement of "the dynamic interactions of hand and brain" are reflected in "the unique character of human thought, growth and creativity" (p. 10). Peter C. Reynolds (cited in Wilson, 1998) postulates the cooperative tool manufacture by *Homo habilis* might have been a pre-condition for the development of language. Wilson reasons, "If language and the employment of the hands for tool manufacture and tool use co-evolved – effectively forging a new domain of hominid brain operations and mental potentials that we collectively refer to as 'human cognition' – then we should find analogous links, or reinforcing effects, between purposive hand use, language, and cognition in the individual histories of living people" (p. 34). Such links and reinforcing effects are apparent when manipulation of blocks under the guidance of a competent tutor facilitates learning complex mathematics at an early age (Albert, 2014; Cuisenaire & Gattegno, 1962; Dienes, 1964; Gattegno, 1961; Goutard, 1964). This study explores how such activity might also facilitate learning statistics at an early age.

The redesign of the hand in *Homo erectus* led to the reallocation of the brains circuitry and the possibility of "mythical thought" with its attendant explanatory and regulatory metaphors and the capacity for "a comprehensive modeling of the entire human universe" (Donald, 1991, p. 214, also quoted in Wilson, 1998, p. 59). Our conceptual system is fundamentally metaphorical derived from situated sensorimotor experience and emotions as well as from biological factors (Lakoff & Johnson, 1980). "What we call 'direct physical experience' is never merely a matter of having a body of a certain sort; rather every experience takes place within a vast background of cultural presuppositions. … Cultural assumptions, values, and attitudes are not a conceptual overlay … our culture is already present in the very experience itself" (p. 57). The cognitivist neuroscientist Marc Jeannerod calls the combination of the ventral visual system with the linguistic areas of the temporal lobes "the semantic processing system"– an information channel for the manipulation, identification, and transformation of objects (Wilson, 1998, p. 107). This channel connecting the occipital and temporal cortex "has access to and could thereafter influence the entire cognitive and behaviorally generative machinery of the brain" (p. 108). These considerations suggest that manipulation of physical objects is coextensive with the learning process. Hand and brain co-evolved. The capabilities of the hand afforded the brain new possibilities for representing the world. Linking foundational statistical concepts to hand movements and manipulation of physical objects might leverage this semantic processing system to stimulate and intensify deep learning. In this study, children engage with an embodied frequency distribution and its measures of central tendency and dispersion through movement and gesture. The experience of children can help us understand the learning process of adults (Thomas & Seely Brown, (2011).

Cognition is Embodied, Situated, and Distributed

Cognition is embedded in a sociocultural context (Sfard & McClain, 2002, p. 157). Learning is integral to practice, and “knowing” generalities and abstractions does not necessarily enable their use in appropriate circumstances[vii]. Lave & Wenger (1991) present a view of situated learning where membership and participation in a community of practice provides the context, as well as essential content of learning. Clark (1997) maintains that “internal representations the mind uses to guide actions may thus be best understood as action-and-context-specific control structures rather than as passive recapitulations of external reality” (p. 51). Hutchins (1995) proposed that cognition is *distributed* among various actors and artifacts. He points out that the cognitive properties of groups cannot be predicted “from a knowledge of the properties of the individuals in the group. The emphasis on finding and describing ‘knowledge structures’ that are somewhere ‘inside’ the individual encourages us to overlook the fact that human cognition is always situated in a complex sociocultural world and cannot be unaffected by it” (p. xiii). Hutchins (2014) describes cognition in terms of a cognitive system in which the individual participates (reminiscent of a Chinese worldview where *participation* takes precedence over *agency*). An essential element of this system is the invention, crafting, and use of tools, especially hand tools. In a *zone of proximal development*, learner, tutor, and artifacts comprise an *activity system* situated in a complex learning environment (Vygotsky, 1978). In such a system, guided manipulation of blocks coupled with play and imagination provide a powerful way to teach fundamental statistical ideas like variation and distribution.

Piaget's (1970) *genetic epistemology* seeks to explain knowledge in terms of the emergence of thought structures during development. Piaget describes the relationship between logical mathematical structures, language, and sensorimotor activities. Sensorimotor intelligence has its own logic – a logic of action, or logic of schemes. Coordination among these schemes and actions gives rise to concepts. Piaget maintained that these foundations of logical mathematical structures are not innate but are rather the result of development within a sociocultural context. Genetic epistemology “deals with both the formation and meaning of knowledge … there is a parallelism between the progress made in the logical and rational organization of knowledge and the corresponding formative psychological processes”

(Piaget, 1970, pp. 12-13). Logical and mathematical knowledge is abstracted not from an object as empiricists claim, but from the action *on* the object – a transformative operation. *Reflective abstraction* is “based not on individual actions but on coordinated actions. … The formation of logical and mathematical structures in human thinking cannot be explained by language alone but has its roots in the general coordination of actions” (pp. 18-19). Similarly, coordinated actions on a set of manipulatives designed to represent a frequency distribution might facilitate an understanding of variation. Pea (1993) noted that knowledge embodied in artifacts becomes accessible to new learners “through observations of use by other humans and attempts to imitate it, through playful discovery of its affordances in solitary activity, and through guided participation in its use by more knowledgeable others” (p. 54). Intelligence is expressed in activity – it is an achievement, not a possession. This study employed guided participation and playful engagement in coordinated action on a set of blocks to build a foundation for conceptual understanding of a dataset or frequency distribution.

Abrahamson & Lindgren (2014) differentiate two epistemic systems: the *primitive* characterized by immediate “doing”, and the *formal* characterized by mediated “thinking” – a distinction they compare to Kahneman's (2011) System 1 (effortless intuition) vs. System 2 (deliberate reasoning). The work of the educator is to guide learners “from immersive action to structured reflection” (p. 359). Abrahamson & Lindgren ask, “How do we select, create, and facilitate physical interactions that give rise to conceptual reasoning and thinking that is aligned with desired classroom learning outcomes?” (p. 364). They identify three areas of pedagogical design: activities, materials, and facilitation. Activities may not necessarily lead directly and immediately to a complex learning outcome but should be engaging and involve simple tasks. Interaction with materials should involve action-feedback loops that allow students to “gradually develop new perceptuomotor schemas that enable them to effectively control objects in service of the more sophisticated task objective” (p. 365). Students in this study developed such perceptuomotor schema as they learned to pick up a set of blocks arranged on a ruler (a metaphor for a frequency distribution) and balance it on a fulcrum. The tutor provided scaffolding and real-time feedback to help the learner enact “functional metaphors” leading to conceptual understanding. “Our ordinary conceptual system,” Lakoff & Johnson (1980) maintain, “in terms of which we both think and act, is metaphorical in nature" (p. 3). "Metaphors are rooted in physical and cultural experience … A metaphor can serve as a vehicle for understanding a concept only by virtue of its experiential basis" (p. 18). Thus, an arrangement of blocks on a ruler can be thought of as a "functional metaphor" for a dataset that can aid in teaching the big ideas of statistics, such as variation and the aggregate properties of a frequency distribution.

Goutard's (1964) work with Cuisenaire rods teaching mathematics to children around the age of 6 offers insight into the learning process of children playing with blocks to learn fundamental ideas of statistics. She proposed three *pedagogical phases of scientific activity*: the *empirical phase*, the *systematization phase*, and *mastery of structures*. In the empirical phase, children manipulate Cuisenaire rods to reveal facts; they develop technique rather than acquire rational knowledge. It is important in this phase that the activity remain a game. In the *systematization* phase, the children organize the facts. In moving from the empirical to the rational, the children use inductive reasoning to discover properties such as the commutative and the associative. In the third phase, *mastery of structures,* perception is synthesized and structured; the mind no longer needs props: "The role of the material is to provide the mind with the experiences from which it will elaborate its own structures" (p. 26). Just as Goutard used concrete materials to provide an intuitive basis for mathematical experience, the tutor in the present study used concrete materials to provide a basis for statistical experience. With respect to method, she claimed, "What is needed is to encourage without hampering, to draw conclusions from the children's activity without directing and channeling their creations" (p. 141). The greatest obstacle to the mastery by children of mathematical [and statistical] structures is the attitudes, habits, and preconceived ideas of adults. In Goutard's view, the role of the teacher is to foster discovery and creativity in the child without imposing preconceived limits on the child's capability. In this study, adopting Goutard's optimism with respect to the capabilities of children, and to test the limits of what might be possible with respect to learning statistics at an early age, and to further demonstrate the independence of statistical concepts from mathematical operations, 5- and 6-year-olds found an approximation of the mean absolute deviation (conceptually equivalent to the standard deviation) through guided manipulation of blocks – a notable feat.

Play

Huizinga (1950) maintained that play is a primary category of life – it does not exist to serve some other life function. It is a distinct form of thought essential to the emergence of culture. Play is the rich soil from which civilization emerged. When a society devalues play, it loses something essential to its vitality: “As a civilization becomes more complex, more variegated and more overladen, and as the technique of production and social life itself become more finely organized, the old cultural soil is gradually smothered under a rank layer of ideas, systems of thought and knowledge, doctrines, rules and regulations, moralities and conventions which have all lost touch with play” (p. 75). Play has rhythm and harmony; it is captivating. Play is the synthesis of strict rules and genuine freedom. Play foreshadows logic and inquiry. Wilson (1998) noted, “the spirit of play, of joyful or just curious experimentation and exploration, comes to us, just as the hand itself comes to us, as a powerful organizer of learning and growth” (p. 292).

“The influence of play on a child’s development is enormous” (Vygotsky, 1978, p. 96). Play organizes higher brain functions and “creates an arena for social interaction and learning” (Brown, 2009, p. 49). Manual play and object manipulation are influential in the acquisition of language and the development of cognitive skills. In play, perception, motivation, and meaning align to bring about internal transformations and create a *zone of proximal development* (Vygotsky, 1978). “Through play the child achieves a functional definition of concepts or objects, and words become parts of a thing” (p. 99).

> As Vygotsky suggests, play, itself, mediates the learning of children. … They learn to understand the meanings of the world as they play with their representations of the world. They build concepts of mathematics and science as well as language, including literacy. We believe that the concepts begun in play not only are the basis for scientific concepts but eventually become part of these concepts (Goodman & Goodman, 1990, p. 228).

Play enables children to master ideas and perform complex actions more easily (Broström, 1999). Thus, as a child manipulates blocks in play within this sphere of activity where the tutor is attentive to the perception and motivation of the child, a foundation for more complex conceptual understanding of the statistical structures represented by the blocks develops.

Vygotsky (1978) suggested, “We might trace the development of arithmetic skills in young children by making them manipulate objects and apply methods either suggested to them or ‘invented’ by them” (p. 74). He thought that in this way we might not only observe the outcome of a learning experience, but also infer the underlying psychological structure. Playful manipulation of physical objects opens up possibilities for learning mathematics at a much higher level than most adults would have thought possible (Cuisenaire & Gattegno, 1962; Dienes, 1964; Goutard, 1964; Montessori, 1912). On the other hand, Davydov (2008) warned, “If school instruction runs counter to the development of the child’s own mathematical intuition, i.e., counter to a more adequate unfolding of the mathematical structures, then we can conclude that instruction does not so much develop the child’s mathematical thinking as hinder its development” (p. 37). Similarly, instruction that runs counter to an intuitive grasp of statistical concepts and structures can hinder the development of statistical thinking. The difficulties adults have learning statistics, statistics anxiety, negative attitudes toward statistics, and widespread misconceptions even among the statistically educated all point to a legacy of learning “hindrances” that might be avoided. Rather than diagnosing and remediating deficient school instruction, why not prevent statistics anxiety, misconceptions, and bad attitudes by discovering how to develop statistical thinking and reasoning in children starting in the first years of formal education? Davydov underscores the importance of intuition in learning mathematics and it is no less important in learning statistics. Bruner (1960) affirmed that students must develop intuition along with analytic skills.

> Intuition [is] the intellectual technique of arriving at plausible but tentative formulations without going through the analytic steps by which such formulations would be found to be valid or invalid conclusions. Intuitive thinking, the training of hunches, is a much-neglected and essential feature of productive thinking not only in formal academic disciplines but also in everyday life. The shrewd guess, the fertile hypothesis, the courageous leap to a tentative conclusion – these are the most

> valuable coin of the thinker at work, whatever his line of work. Can school children be led to master this gift? (Bruner, 1960, pp. 13-14).

Bruner lamented, “Unfortunately, the formalism of school learning has somehow devalued intuition” (p. 58). Kahneman (2011) cautioned that the statistical intuitions of adults are notoriously incorrect. Perhaps developing such intuitions from the beginning of formal education in play-based learning might avert difficulties commonly seen in older learners.

Play-based Learning

Play is multifaceted. Conceptualizing play as a single entity has led to a profusion of theoretical perspectives (Fleer, 2009). Hutt (1989, cited in Fleer) identified fourteen categories of child behavior labeled “play” in the research literature; these categories accounted for almost all of children’s activities and behaviors. There is clearly no consensus on what researchers mean by *play*. Fleer (2011) advocates rethinking the concept of play in cultural-historical terms drawing on the work of Vygotsky and Davydov. Vygotsky (1978) rejected the idea that play is the child’s work, but rather play is a “leading activity” (p. 103), not the predominant activity of young children. Fleer maintains cognition and imagination develop together, “with imagination acting as the bridge between play and learning” (p. 224). “Play-based programs,” she contends, “can build children’s theoretical thinking in play, where imagination acts as the bridge between play as a leading activity and learning as a leading activity” (pp. 225-226). Fleer’s theory of *Conceptual Play* provides a framework for exploring the link between play and learning and it can enhance a model of a human activity system as shown in Figure 2.4.

Learning in play for children has its analogue in *flow learning* for adults (Csikszentmihalyi, 2014)[viii]. Although Fleer does not address this connection in her theory of Conceptual Play, Csikszentmihalyi’s conception of flow provides an additional enhancement of the activity system model in Figure 2.4. The conditions of the flow experience are like those found in the *zone of proximal development* when tutor and learner are synchronized in activity that carries both along an autotelic learning trajectory.

Playing with tangible representations of a concept does not, by itself, lead to conceptual understanding (Uttal et al., 1998). “Playing with an object may engage children’s interest but it may simultaneously make it hard for them to grasp the relation between the object and a concept or fact. … Children may enjoy playing with mathematics manipulatives, but doing so may not help them learn arithmetic" (p. 59). We cannot take for granted that a young child will understand an arrangement of wooden cubes on a ruler as a representation of a frequency distribution. And we should not make objects more colorful and attractive as this may detract from the ability of the child to see them as representations of something else. The plain wooden one-inch cubes used in this study to represent data points are almost identical to those Fröbel (2005/1826) used in the first kindergartens. Although play can be a catalyst for learning, it does not guarantee that a play experience will result in abstract understanding. Additional scaffolding is needed. The enhanced model of a human activity system shown in Figure 2.4 provides a theoretical framework for understanding a complex learning process from a broad systems perspective and for scaling up tutorial interactions to higher levels of complexity and diverse units of analysis such as classrooms, schools, and communities.

Learning through Block Play

The challenge for the teacher in using blocks as a didactic tool is to provide just enough structure and direction to achieve learning goals while allowing the learner a high degree of autonomy (Bruner, 1961; Goutard, 1964). The process of learning mathematical structures needs to be made enjoyable to be effective (Dienes, 1964). Dienes thought both the ethical as well as the intellectual development of children could be enhanced by eliminating rewards and punishments and cultivating instead their intrinsic motivation through interest in the task itself.[ix] Children can enjoy making embodiments of mathematical structures if we "put in their hands material so designed that through controlled manipulation certain mathematical relationships will become clear. … Concrete material can enormously accelerate the learning process" (p. 43). Dienes set up conditions in which children learned fractional exponents through block play. He observed that as children played with the blocks under appropriate guiding constraints, they began to play with the *properties* of the blocks, not just the blocks themselves. The young learners transformed the blocks into mathematical symbols. When children learn the underlying structure of powers and roots through block play, they recognize a logarithm as the same structure in another form. Similarly, if children learn the underlying structure of a frequency distribution through block play they might come to recognize a similar structure in any dataset.

By making multiple embodiments, learners might come to recognize a frequency distribution as an aggregate, not just a collection of individual data points, and with appropriate scaffolding, come to understand the underlying structure and the abstract idea of variation. As multiple embodiments vary over the full range of their possibilities, the abstraction is stripped of its non-essential or idiosyncratic features and the abstraction becomes a plaything (Dienes, 1964). Through multiple interactions, an abstraction such as a frequency distribution becomes a conceptual tool available to the learner in a broad range of practical applications. The National Research Council recommends:

- Using multiple and varied representations of concepts and tasks, such as diagrams, numerical and mathematical representations, and simulations, combined with activities and guidance that support mapping across the varied representations.
- Engaging learners in challenging tasks, while also supporting them with guidance, feedback, and encouragement to reflect on their own learning processes and the status of their understanding (National Research Council (U.S.), 2012, p. 9).

Vygotsky (1978) observed that in young children, concepts are based on concrete recollections not logical structures – the child's thinking is dependent on memory. Likewise, in the development of visual concepts their internal representations "are based on recall of concrete instances and do not yet possess the character of an abstraction. … Memory rather than abstract thought is the definitive characteristic of the early stages of cognitive development" (p. 50). More recent studies, however, indicate "that children, when they have accumulated substantial knowledge, have the ability to abstract well beyond what is ordinarily observed" (National Research Council (US), 2001, p. 5). Dienes (1964) maintained, "Symbol-manipulation in mathematics is all too often utterly meaningless simply because there is no corresponding transformation of images" (p. 105). As in mathematics, much of statistics education consists of nothing but rule-bound symbol manipulation. Meaning arises from multiple embodied interactions, not from repetitive manipulations of symbols. "It will be the task of the future teacher of mathematics," Dienes wrote, "to put children into carefully selected situations in which their creative urges can be set free, so that they become masters and not slaves of mathematical symbolism" (p. 152). Similarly, it will be the work of the teacher of statistics to create the learning environment in which children become masters, not slaves, of statistical symbolism and its application in a process of inquiry.

Statistics Education

Understanding a data set as an aggregate (a distribution) with the emergent properties of shape, spread, and center is one of the "big ideas" of statistics. Conceptual systems are built up of metaphors, and metaphors are grounded in embodied experience in the world (Lakoff, 2015a, 2015b; Lakoff & Johnson, 1980). Building on the Neural Theory of Language (Feldman & Narayanan, 2004), Narayanan proposed a neural theory of metaphor that accounts for the linking of the sensorimotor system through neural circuitry to higher cortical areas giving rise to metaphorical thought. Primary conceptual metaphors are "learned unconsciously and automatically in childhood simply by functioning in the everyday world with a human body and brain" (pp. 256-257). Neuronal maps are physically embodied in our nervous system through *neural recruitment* between clusters of neurons, or *nodes.* "This neural learning mechanism produces a stable, conventional system of primary metaphors that tend to remain in place indefinitely within the conceptual system and are independent of language" (p. 256). This suggests that establishing foundational statistical metaphors (such as a dataset as an aggregate) early in a learning progression might avert the anxiety and conceptual confusion that often plagues adults. It further might address the limitations of statistical intuition identified by Tversky and Kahneman (Kahneman, 2011; Kahneman, Slovic, & Tversky, 1982; Tversky & Kahneman, 1974).

The Challenges of Teaching and Learning Statistics

Statistics is a fundamental method of inquiry (Ben-Zvi & Garfield, 2004, p. 4). However, it is commonly understood as simple statements of numerical facts or a collection of skills, procedures, and computations. Consequently, the teaching and learning of statistics often lacks coherence, engenders misconceptions, and is for many an unpleasant experience. *Statistics anxiety (SA)* and *negative attitudes toward statistics (NATS)* have become a specialized area of investigation in the field of statistics education. Researchers have developed over a dozen psychometric instruments to assess statistics anxiety (Gal, Ginsburg, & Schau, 1997; Onwuegbuzie & Wilson, 2003; Ramirez, Schau, & Emmioğlu, 2012). A common misconception is that statistics is a branch of mathematics. Although statistics often involves computation, mathematical formulas frequently impede, rather than facilitate, understanding of statistical ideas (Piaget, 1948; Rumsey, 2002; Wild & Pfannkuch, 1999). Statistics is a transdiscipline; it should be taught across the curriculum along with literacy, numeracy, and critical thinking (Garfield & Ben-Zvi, 2008; Watson, 2011). Courses in statistics commonly do not train students to reason statistically and to think critically with quantitative data but rather they present mathematical abstractions and technical jargon in a context that is not meaningful to students. Although students may learn to navigate software packages and follow statistical recipes[x], often they lack understanding of underlying concepts (Delmas et al., 2007). Critical gaps in statistical literacy are found not only among those who may have struggled through a course in statistics, but also among the statistically educated.

Many findings published in scientific journals have been called into question due in large measure to a lack of *conceptual* understanding of the statistical methods used in the studies[xi] (Baker & Penny, 2016; Cohen, 1994; Sohn, 1991; Wilkinson, 1999). This has led to a *reproducibility crisis* (Baker & Penny, 2016). Breiman (2001) explained:

> Hundreds, perhaps thousands of articles were published claiming proof of something or other because the coefficient was significant at the 5% level … The deficiencies in analysis occurred because the focus was on the model and not on the problem. … When a model is fit to data to draw quantitative conclusions, the conclusions are about the model's mechanism, and not about nature's mechanism. It follows that if the model is a poor emulation of nature, the conclusions may be wrong. … The linear regression model led to many erroneous conclusions that appeared in journal articles waving the 5% significance level without knowing whether the model fit the data. (Breiman, 2001, pp. 202-203).

Many published research findings contain unexamined assumptions or fundamental errors in sampling or analysis (Ioannidis, 2005). For example, the What Works Clearinghouse (WWC) in the U.S. Department of Education maintains an evidence base on the effectiveness of educational interventions. One study (Malouf & Taymans, 2016) found that only 33% of studies reported in peer-reviewed journals met WWC standards for research evidence.

There is widespread agreement that the statistics education community needs to put more emphasis on conceptual understanding rather than on computational techniques and software skills. The Assessment Resource Tools for Improving Statistical Thinking (ARTIST) project at the University of Minnesota (2006) is one such effort. The goal of ARTIST "is to help teachers assess statistical literacy, statistical reasoning, and statistical thinking in first courses of statistics". It is aimed primarily at college students. The Comprehensive Assessment of Outcomes in a first Statistics course (CAOS) was developed to assess statistical literacy, conceptual understanding and reasoning about variability (Delmas et al., 2007). LOCUS (Levels of Conceptual Understanding in Statistics) is a set of assessments developed in alignment with the GAISE framework to measure conceptual understanding of statistics in grades 6-12 (Whitaker, Foti, & Jacobbe, 2015). A validated assessment tool for the early primary grades has not yet been developed.

The GAISE Framework

The *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report* (Franklin et al., 2005) was a milestone in addressing the global challenge of fostering statistical literacy. These guidelines originally commissioned by the American Statistical Association (ASA) for introductory college courses were updated in 2016. Statistics education should begin with inquiry – a question that can be answered with data. Then collect data and use it to answer the question. The *GAISE Report* recommends, "Teach statistics as an investigative process of problem-solving and decision-making. … students should understand that statistics is a problem-solving and decision-making process that is fundamental to scientific inquiry and essential for making sound decisions" (p. 6). Statistics education needs to focus on statistical thinking and conceptual understanding rather than on symbol manipulation and computation. In addition, the GAISE framework recommends that foundations of statistical literacy include a conceptual understanding of a data distribution and an ability to describe its key features and represent a dataset with tables, graphs, and numerical summaries. This should begin in grades PreK-5. Garfield & Ben-Zvi (2008) identify four key capabilities of statistical literacy: 1) formulate a statistical question; 2) collect, organize, and display data; 3) perform simple analyses and interpret results; and 4) use basic statistics vocabulary and symbols. None of these activities necessarily requires mathematics; even analysis can be done without computation using graphical methods. In the discipline of statistics, variability and context are paramount, not computation.

> Across all levels and stages of the investigative process, statistics anticipates and accounts for variability in data. Whereas mathematics answers deterministic questions, statistics provides a coherent set of tools for dealing with “the omnipresence of variability” (Cobb and Moore, 1997)… The focus on variability distinguishes statistical content from mathematical content. For example, designing studies that control for variability, making use of distributions to describe variability, and drawing inferences about a population based on a sample in light of sampling variability all require content knowledge distinct from mathematics (Franklin et al., 2015, pp. 1-2).

Even statistical inference can be conducted without computation using boxplots to estimate parameters (Wild, Pfannkuch, Regan, & Horton, 2011). The box plot superimposed on a dot plot (Figure 2.1) shows the relationship between individual data points, the statistics that describe the distribution, and the aggregation of the data into quartiles. In this study, students created a similar juxtaposition of an X-plot and a box plot.

One of the core ideas of statistics is the emergence of features from an aggregate that are distinct from the properties and features of its individual elements. The properties of a dataset cannot be derived from the properties of its individual data points independently of their relationship with each other. A *frequency distribution* is an organizing structure that makes variability visible. (Bakker & Gravemeijer, 2004; Cobb, 2004; Garfield & Ben Zvi, 2007; Wild & Pfannkuch, 1999). Seeing a data aggregate holistically is the beginning of *distributional reasoning* (Biehler, Frischemeier, Reading, & Shaughnessy, 2018).

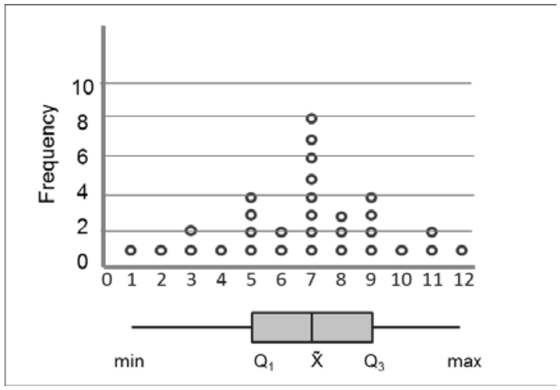
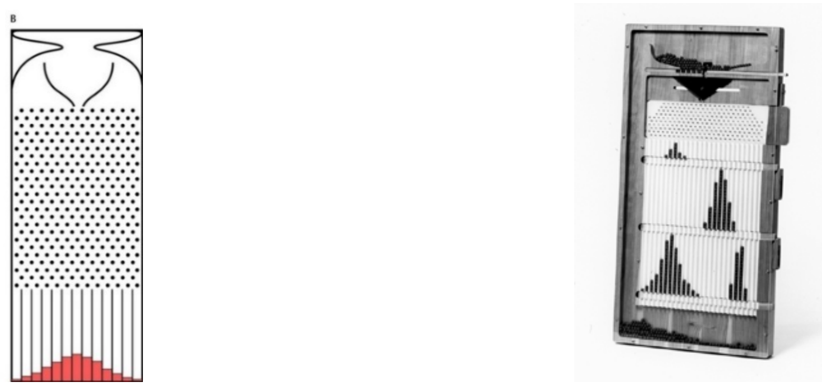


Figure 2.1. The dot plot shows each data point, the shape of the distribution and its spread and center. The box plot aggregates the data points into quartiles – one fourth of the data points are in each segment of the box plot. The box plot is more efficient than the dot plot at showing spread and center (median), but details of the shape are lost.

Learning Statistics with Manipulatives

The use of manipulatives to teach statistics is not a new idea (Kunert, Montag, & Pöhlmann, 2001). The Galton Board (Galton, 1894, p. 63), more commonly known as a *quincunx*, shows how a distribution pattern emerges when steel balls cascade down a grid of pins set in a wooden block (Figure 2.2). When a ball hits a pin, it has a 50/50 chance of falling to the right or to the left. This gives rise to the binomial distributions seen in Figure 2.2. The quincunx shows how a frequency distribution emerges from a series of single events. The bead bowl, the sampling box, the catapult, and the Red Bead Game (Deming, 1993) have all been used to teach principles of variation to adult learners.



a. Galton Board. b. A modern quincunx.

Figure 2.2. The quincunx demonstrates how a distribution pattern emerges from a series of single events. (Galton, 1894).

A Conceptual Framework for Building Statistical Literacy

Bruner's *modes of representation,* Goutard's *phases of scientific activity* and Piaget's *reflective abstraction* all begin with tactile-kinesthetic activity and proceed toward abstract understanding (Figure 2.3). These models are linear and oriented toward an individual learner. After ten years leading the curriculum reform movement, Bruner (1971) challenged the entrenched focus on the isolated learner that had come to dominate education reform. He saw that the process of education must draw on resources in the community and from other

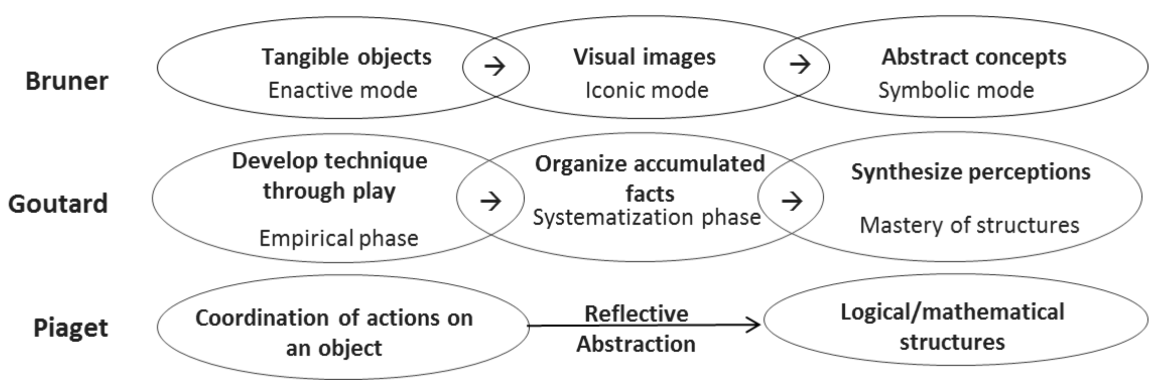


Figure 2.3. The learning progressions proposed by Bruner, Goutard, and Piaget begin with sensorimotor engagement with tools and artifacts and progress toward abstract conceptual understanding.

learners if it is to be relevant to the wellbeing of society. Ann Brown (1992) saw the systems perspective as essential for understanding learning environments: “Just as it is impossible to change one aspect of the system without creating perturbations in others, so too it is difficult to study any one aspect independently from the whole operating system” (Brown, 1992, p. 143). Activity Theory adopts this systems perspective.

Activity Theory

The model of an activity system (Engeström, 1999) includes more than just the learner, the teacher and the curriculum. It embraces the wider sociocultural context and so can accommodate ideas of learning as being situated, embodied, and distributed. The unit of analysis may be an individual learner, a dyad of learner and tutor, a learning group, a classroom, a school, or an entire community. Whether studying concept formation, skill development, or professional development of teachers, Activity Theory offers a comprehensive framework for designing, evaluating, and improving learning experiences and for understanding the process of concept formation and conceptual change. “With variations in topic and focus, the idea that concepts exist in distributed cultural practices and change through processes that extend beyond individual thinking is now widely accepted among researchers working on learning and conceptual change in a sociocultural tradition” (Hall & Jurow, 2015, p. 173). A dynamic system view of learning not only paves the way for discovering new learning strategies, it provides a framework for scaling them up to higher levels of complexity. Activity theory seeks to bridge the gap between theory and application, between conceptualization and intervention.

Engeström (1999) proposed an activity system of six elements. Figure 2.4 shows a model adapted from Engeström and augmented with Fleer's (2011) view of imagination as a bridge between play and learning and Csikszentmihalyi's (2014; Csikszentmihalyi & Bennett, 1971) conception of *flow*. Activity Theory developed within the psychology of play, learning, cognition, and child development (Engeström, Miettinen, & Punamaki, 1999). Activity Theory and the theory of Conceptual Play both have Vygotskian roots and recognize the vital role of history and culture in the learning process and in cognitive development.

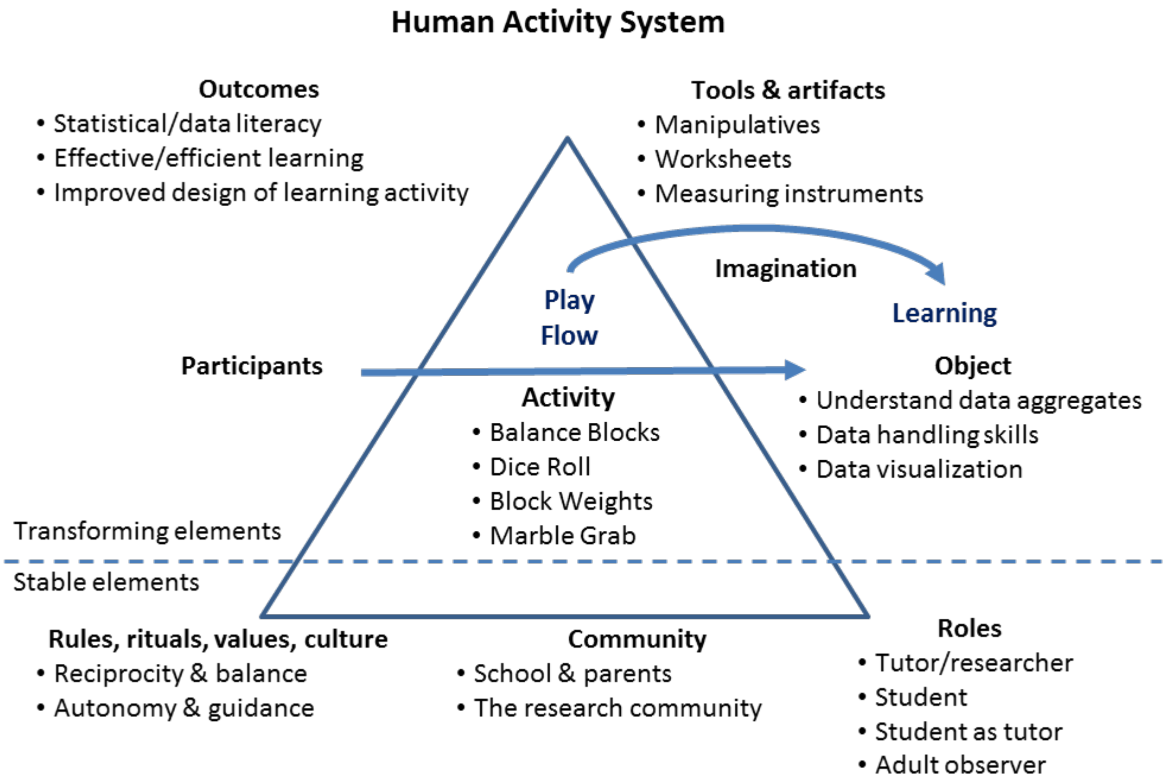


Figure 2.4. This model of a human activity system adapted from Engeström (1999), Fleer (2011), and Csikszentmihalyi (1999), gives a comprehensive view of the learning process.

In this study, participants engaged in learning activities (described in Chapter 3) to understand data aggregates and learn skills of data handling and data visualization using manipulatives, worksheets, and measuring instruments. Intended outcomes for the students included statistical literacy; intended outcomes for the tutor included more effective and efficient designs of learning activities. These elements of the learning system interacted during the learning sessions and were mutually transformative. In addition, more stable elements undergird these dynamics: a commitment to reciprocity and balance between learner autonomy and tutorial guidance in the learning interactions, defined roles, and community support.

Learning Trajectories

Learning trajectories are increasingly used in statistics education research and are useful in understanding the dynamics of the activity system in Figure 2.4. Their increased use parallels the growth in participatory research methods (Arnold et al., 2018). The concept of a learning trajectory combines a hypothesized psychological development progression with a sequence of instructional tasks designed to reveal the thinking and learning patterns of the learner and achieve a specific learning goal. Simon (1995) identified three components of a *hypothetical learning trajectory:* a learning goal, learning activities, and a hypothetical learning process. This hypothetical process anticipates "how the students' thinking and understanding will evolve in the context of the learning activities" (p. 136). This process may vary between students and for the same student at different times. It is unlikely that the hypothesized process will exactly match the actual learning process. The closer they match, the more efficient the learning. "The theory is that learning consistent with such natural developmental progressions is more effective, efficient, and generative for the student than learning that does not follow these paths" (Clements & Sarama, 2004, p. 84). The researcher must be flexible. There is a "symbiotic relationship" between "the development of a hypothetical learning process and the development of the learning activities" (Simon, 1995, p. 136). The simpler term "learning trajectories" is now used to refer to both the hypothesized trajectory and the enacted trajectory (Arnold et al., 2018).

Chapter Summary

Statistical literacy is an essential element of 21st century education. This chapter provided evidence and arguments to support the assertion that cognition is grounded in sensory-motor experience (Barsalou, Simmons, Barbey, & Wilson, 2003; Thelen & Smith, 1994). Learning occurs through playful discovery and guided participation (Dewey, 1938a) in a zone of proximal development (Vygotsky, 1978). Hand and brain co-evolved; the hand is central to the emergence of symbolic thought (Wilson, 1998). Self-movement structures our knowledge of the world (Sheets-Johnstone, 2011). Experience, knowledge, thought, emotion, and practice are all one process (Bohm, 2004). Statistics education has largely failed to impart conceptual understanding of statistical methods; innovative, culturally responsive approaches are needed. Play with an abstract concept in concrete form can prepare the mind for later conceptual understanding (Bruner, 1964). Activity Theory (Engeström, 1999) offers a powerful conceptual framework for designing, implementing, and assessing learning trajectories (Clements & Sarama, 2004). Children are capable of achievement at much higher levels than is generally recognized by adults (National Research Council, 2001; Goutard, 1964). By studying how children learn, we can gain insight into how adults learn (Thomas & Seely Brown, 2011).

Chapter 3: Methods

In this study, a tutor/researcher introduced statistical concepts to students in grades K-2 through guided play with blocks and other manipulatives. Students created datasets by rolling dice, weighing blocks and marbles with a digital scale, and counting handfuls of marbles. They organized and transposed data; described the shape of the dataset; created data visualizations; used symbols; and found measures of dispersion and central tendency. Six lessons were designed to introduce a frequency distribution as an aggregate with emergent properties of shape, spread, and center and guide the learner to an understanding of variation.

Study Design

Microgenetic methods allow for fine-grained analysis of learning processes in the study of pedagogical practices in learning trajectories (Chinn & Sherin, 2014). “The goal is not merely to identify factors that influence learning, but to understand how these factors mediate learning, step by step, as learning occurs” (p. 171). Learning is not a unitary phenomenon driven by “independent cogitations of the individual” but rather is *mediated* by cultural tools and people. It does not happen during “encapsulated moments” but rather learning occurs continuously in parallel on multiple fronts. With microgenetic methods “the aim is to see learning as it happens, and to understand the factors that engender it” (p. 180). However, Chinn and Sherin advise caution in making inferences about cognitive structures and processes.

Exploratory Case Study Methods

Educational research frequently uses *case study methods* (Merriam, 1998; Stake, 2006; Yin, 2014). Case study research designs are emergent and flexible, responding to changing conditions during the study and to the insights and intuitions of the researcher. They rely on the researcher as the primary instrument for data collection and analysis and therefore must include steps to minimize researcher bias. Case studies may be explanatory (causal), descriptive, or exploratory. Exploratory methods are appropriate for research aimed at discovery, insight, and understanding (Merriam, 1998) as in the present study. A case study research design relies on converging evidence from multiple sources to support theoretical propositions or to explain empirical observations and to guide data collection and analysis.

Case study research methods can accommodate multiple epistemological orientations. The unit of analysis may change because of discoveries during data collection (Yin, 2014).

Stake (2006) distinguishes “instrumental” case studies, where the objective is to go beyond the case (as in the present study), from “intrinsic” case studies where the main interest is the particular case under investigation. He emphasizes, “Good hard thinking about the relative importance of research questions will increase the relevance of observations” (p. 13), and “getting the research question and other content of the study right is as important as getting the methods right” (p. 17). It is important to be clear on “what concept or idea binds the cases together” (p. 23). Research questions form a *conceptual structure* for the research design and for interpreting findings. Care must be taken to ensure research questions fit the researcher’s intentions. For some questions, the knowledge base may be sparse and there may be no developed conceptual framework or hypotheses. This is a candidate for an exploratory study. The researcher should become familiar with a range of relevant types of theories, including theories of individual development, cognition, learning, and interpersonal interactions (theoretical triangulation). In this study, the researcher considered Dewey’s (1938b) theory of inquiry, Dienes’ (1964) theory of mathematical stages of development, Bruner’s (1964) modes of representation, Goutard’s (1964) phases of scientific activity, Vygotsky’s (1978) sociocultural theory and zone of proximal development, Piaget’s (1970) genetic epistemology and reflective abstraction, Engeström’s (1991) activity theory, embodied cognition (e.g., Barsalou, 2008; Barsalou et al., 2003; Varela et al., 2016), Hutchins' (1995, 2014) distributed cognition, Lave & Wenger's (1991) situated cognition, Thelen and Smith’s (1994) dynamic systems approach to the development of cognition and action, Lakoff's (2015b, 2015a) theory of conceptual metaphor, Feldman and Narayanan's (2004) neural theory of metaphor, and Fleer’s (2011) theory of conceptual play.

Stake (2006) notes, “generalizations, principles, or lessons learned from a case study may potentially apply to a variety of situations, far beyond any strict definition of the hypothetical population of ‘like-cases’ represented by the original case; … *analytic generalizations* may be based on either (a) corroborating, modifying, rejecting, or otherwise advancing theoretical concepts that you referenced in designing your case study or (b) new concepts that arose upon the completion of your case study” (p. 41). Case study research should aim for inferences beyond the case study findings themselves at the level of theory or policy. The present study has implications for bridging the gap between guided block play and statistical inquiry in the elementary grades; for scaling up tutorial learning sessions to classroom teaching; and for improving the statistical education of adults.

Multiple Case Studies

Multiple-case designs follow a replication logic. The research design reflects either an expectation of similar results for two or three cases (*literal replication*) or contrasting results for two or more sets of three to five cases (*theoretical replication*). The present study follows a literal replication logic for two types of cases: single participant and participant pairs. The added dimension of interaction between participants in pairs might provide additional insight into the learning process and how it might be scaled up to larger groups. The logic behind the expectation of contrasting results should be explicit and theoretically grounded. Yin (2014) illustrates a multiple-case study design procedure in which the initial step consists of theory development. The researcher treats each case independently seeking convergence of evidence on the conclusions of each case. This study is grounded in the theory of grounded cognition: cognition is embodied (Barsalou, 2008), situated (Vygotsky, 1978; Lave & Wenger, 1991) , and distributed (Hutchins, 1990, 2014).

In conducting a case study, there is a “continuous interaction between the theoretical issues being studied and the data being collected” (Yin, 2014, p. 72). Some of the information relevant to the study may not be apparent until the study is underway, so the formal protocol needs to be adaptable without compromising rigor. The researcher may need to interpret information as it emerges from the ongoing data collection process and adjust accordingly. Case study evidence may come from interviews, documents, direct observation, physical artifacts, archival records, or participant observation. Participant observation is prone to bias and the researcher needs to address this in the research design. Details of how researcher bias was addressed are presented in the section on trustworthiness later in this chapter.

Dynamics of the Learning Sessions

The researcher was also the tutor[2]. Consent of parents/guardians was obtained according to procedures approved by the Institutional Review Board. Assent of children volunteers was also obtained. At the start of the first session, the researcher asked each participant a few questions to establish rapport and get some background information. He asked what they like about school, if they like to play with blocks, their age, date of birth, and who their teacher is. Each session was videotaped; students gave their permission at the beginning of each session to start the camera. Students played with the blocks in free play until they were ready to try something new. In most cases, this was less than five minutes. To transition to the first lesson, the tutor asked if they wanted to learn a new way to play with the blocks.

The tutor followed a scaffolding process similar to that of Wood, Bruner, & Ross (1976) where young learners assembled interlocking blocks into a geometric shape. However, the task in their study was deterministic (there was only one correct way to assemble the blocks) whereas the present study allowed for greater learner autonomy in completing the task. Participants explored multiple embodiments of a dataset or frequency distribution – a *functional metaphor* (Lakoff & Johnson, 1980). They found minimum (min), maximum (max), and mode through direct observation; sample size (N), range, and interquartile range (IQR) by counting; median, first and third quartiles (Q1 and Q3) by manipulating the blocks, and average by balancing blocks on a ruler. Two students also found the mean absolute deviation through a three-stage procedure described later in this chapter.

Participants

Nine participants in grades K-2 and their parents responded to a call for volunteers at a small school in a major city in the U.S. Southwest. The population of the school was approximately 49% Hispanic, 36% White, 4% Black, 4% two or more races, 4% American Indian/Alaska Native, 1% Asian, and 1% Hawaiian Native/Pacific Islander. Approximately 65% of students were from “low-income” families[3]. The Head of School provided a letter of support to the

[2] In Case 2.3, a second-grade participant tutored her kindergarten friend.
[3] Source: https://www.greatschools.org/new-mexico/albuquerque/2022-International-School-At-Mesa-Del-Sol/#Race_ethnicity

Institutional Review Board (IRB) and asked K-2 teachers to send a letter composed by the researcher and approved by the IRB asking parents and children if they would like to participate in this study. Participants were expected to have prior knowledge[xii] of numbers from one to twenty and to recognize and write numerals and letters. Volunteers followed consenting procedures approved by the IRB. One boy left the study during his first session and one girl joined the study in the final week keeping the total number of participants at nine. Table 3.1 shows the grade level, sex, ethnicity, and age in decimal years of each participant on the day of their first session. Sessions were held from 3 April 2018 to 21 May 2018. Each student participated in from one to five one-hour sessions either as individuals or in pairs. Three sessions were held with pairs of students (Cases 2.1 to 2.3) and 25 sessions with individual students (Cases 1.1 to 1.7). A total of 28 sessions resulted in 23 hours of video, over 500 pages of transcripts, and 65 learning artifacts as shown in Table 3.1. Artifacts included pre-printed worksheets, data collection sheets, and data graphics. Some video was lost due to camera failure. In most cases, backup audio filled gaps in the video.

Table 3.1 *Summary of cases.* Each case consists of from one to six one-hour sessions. The first seven cases were one-on-one sessions and the last three were with pairs of students. [ethnicity codes: W=White, H=Hispanic, A=Asian]

Case No.	Code	Age	Grade	Sex	Ethnicity	Number of sessions	Hrs of video	Pages of transcript	Number of artifacts
Case 1.1	FB	6.3	K	F	W	6	5.3	147	12
Case 1.2	LC	6.8	1	F	H	4	3.4	63	11
Case 1.3	EM	7.3	1	M	H	5	3.9	114	15
Case 1.4	AS	7.7	2	F	A	1	0.7	12	3
Case 1.5	MT	7.7	2	F	W	5	3.9	62	10
Case 1.6	JL	7.9	2	M	W	1	0.6	12	2
Case 1.7	CB	8.1	2	F	H	3	2.4	52	7
Case 2.1	ASVB	7.7/8.3	2/2	F/M	A/W	1	0.9	13	1
Case 2.2	CBAS	8.1/7.7	2/2	F/F	H/A	1	0.8	22	1
Case 2.3	MTOT	7.7/6.2	2/K	F/F	W/W	1	0.9	20	3
TOTAL						**28**	**22.9**	**517**	**65**

Data Collection

Data sources included video, audio, field notes, analytic notes, worksheets, and the resarcher's reflections as a participant observer. Triangulation from these multiple data sources helped mitigate the effects of researcher bias. The primary sources of data were the video recordings and transcripts and the worksheets completed by the students (the artifacts).

Videotaping and Interaction Analysis

As video recording in education and learning sciences research has increased over the past 20 years (Derry et al., 2010), the National Science Foundation (NSF) recognized the need for guidelines to help ensure quality in video research methods (Goldman, Zahn, & Derry, 2014). The guidelines issued by the NSF (Derry, 2007) address four main categories of video-based research: 1) teaching and learning processes in formal settings such as classrooms, 2) peer-to-peer and adult-child interactions in informal settings, 3) video as a tool for learning, and 4) video as a tool for professional development. This study is in the second category: adult-child interactions in an informal setting. Observations can be biased by technical factors such as camera angle, field of view, lighting, and sound quality (Powell et al., 2003). Video data are also subject to researcher bias in judgment and perception. "It is impossible to include all potentially relevant aspects of an interaction, so that, in practice, the transcript emerges as an iteratively modified document that increasingly reflects the categories the analyst has found relevant to [the] analysis" (p. 48). These human and technical factors limit and shape the analytic possibilities. Interaction analysis is one way to address researcher bias in video-based research methods.

Interaction analysis is a method for investigating human interactions using audiovisual recordings (Jordan & Henderson, 1995). It is an empirically grounded method for understanding learning processes seen through the lens of situated engagement with people and things. Interaction analysis generally assumes that knowledge is socially constructed, organized, and used; that it is socially and culturally situated; that theories of knowledge and action must be held accountable to empirical evidence; and that audiovisual records can be reliable, empirically grounded sources of data. Participants in interaction analysis may corroborate or refute provisional findings of the researcher, offer alternative explanations, challenge assumptions, or entertain hypothetical learning trajectories.

Seven adult participants with experience in observing children in learning environments volunteered to participate in two interaction analysis sessions. Three had a background in Montessori education, one had a Ph.D. in clinical psychology, two were retired elementary teachers, and one was a retired specialist and trainer in early childhood education. They followed IRB-approved consenting procedures and met as one group of four and one group of three. Sessions lasted approximately 50 minutes and 90 minutes respectively. Each group viewed selected video clips along with the corresponding worksheets. They shared their observations and insights. The researcher sometimes stopped the video for an extended discussion. These sessions were audio recorded and transcribed and are summarized in Appendix B.

Analytic Rubrics

A rubric was prepared to aid in the interaction analysis (Figure 3.1). A rubric provides a clear and coherent set of learning criteria and clear descriptions of observable performance levels (Brookhart, 2013). Learning is not observable directly, so performance serves as an *indicator* of achieving target learning outcomes. An *analytic* rubric considers each criterion individually while a *holistic* rubric considers all learning criteria together. Rubrics are not evaluative; they are *descriptive.* Rubrics provide structure and coherence to observations. *General* rubrics, in contrast to *task-specific* rubrics, define criteria and describe performance that generalizes across tasks and focus on a target learning outcome. A rubric is a tool to help maintain focus on *learning* rather than on tasks. The act of creating a rubric helps to clarify both content and learning outcomes. A rubric can help keep process and product in balance in the design of learning experiences as well as in the analysis. In addition, the rubric can aid in maintaining balance between types of learning goals – sensorimotor, affective, cognitive, and metacognitive. A rubric matches observed performance to a description of anticipated performance: “The rubric description is the bridge between what you see … and the judgement of learning” (p. 22). Ideally, performance criteria are mutually exclusive and collectively exhaustive, definable over a range of performance levels, appropriate, and observable. The rubric in Figure 3.1 helped to clarify content and learning outcomes and provide some structure and coherence to the observations, however, it was not used in the interaction analysis due to limitations of time and availability of participants.

Performance Level / Learning Criteria	Factual knowledge	Procedural knowledge	Conceptual knowledge	Metacognition
Data sets				
Data collection & organization				
Data visualization				
Understanding variation				

Figure 3.1. A general analytic rubric for assessing performance.

Structure of Observed Learning Outcomes (SOLO) Taxonomy

The Structure of Observed Learning Outcomes (SOLO) taxonomy (Biggs & Collis, 1991) is a model of cognitive development and a touchstone for setting learning goals. This model consists of five levels of increasing conceptual complexity. It is simpler than Bloom's taxonomy, easier to use, and more aligned with how people think and learn (Hattie & Yates, 2014). The SOLO model has gained wide acceptance in research on the development of statistical reasoning, particularly reasoning about variation (Biehler *et al.*, 2018; Chick & Watson, 2001; Jones *et al.*, 2000; Jones, Langrall, Mooney, & Thornton, 2004). In the SOLO model, Biggs and Collis identify five modes of increasing conceptual complexity: *sensorimotor, ikonic, concrete-symbolic, formal,* and *post-formal.* Within each mode are three levels of response to new information: unistructural responses (U), multi-structural responses (M), and relational responses (R). These levels (U-M-R) form a cycle of cognitive growth. In addition, there is a *prestructural* level where the learner has not quite oriented to the learning task at hand, and an *extended abstract* level that is a transition to the next mode at a higher level of abstraction. Biehler (2018) contends that "the two modes most relevant to school-aged student reasoning are the ikonic mode (making use of imaging and imagination) and the more cognitively complex concrete-symbolic mode (operating with second-order symbol systems such as written language" (pp. 143-144). "The levels of the Biggs and Collis learning cycle have provided a powerful theoretical base for situating research on students' statistical reasoning from the elementary school years through college" (Jones *et al.*, 2004, p.

100). The current study suggests that the sensorimotor mode may play a more important role than is evident in current research on early statistical literacy. Figure 3.2 shows a learning progression of increasing levels of complexity based on lessons in the current study.

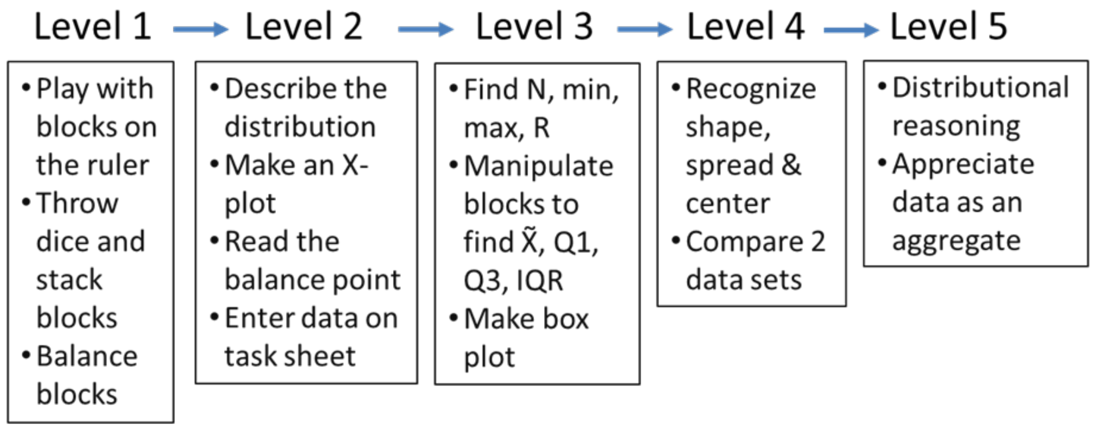


Figure 3.2. An example of a learning progression of increasing levels of complexity.

Data Analysis

The researcher transcribed the video of the learning sessions and the audio of the two interaction analysis sessions making observational and analytic notes in the transcript and adding field notes. He replayed the video while referencing the transcript and the artifacts making additional notes and corrections and a brief outline of each session. He then wrote a narrative summary of each session while referencing completed worksheets and reviewing the video at variable speed examining activity, facial expressions, gestures, and transitions. Finally, he added his reflections on the sessions. These reflections identified opportunities to improve the design of the manipulatives, the design of the learning activities, the tutor's effectiveness, and the dynamics of the learning interactions.

In examining data from a case study, Yin (2014) advocates "playing" with the data – watching for patterns, concepts, insights and creating diagrams, tables, matrices, memos, and other aids to thinking about the data. This should be done <u>before</u> formulating a general analytic strategy for cross-case analysis (Stake, 2006). "Data analysis consists of examining, categorizing, tabulating, testing, or otherwise recombining evidence, to produce empirically based findings" (Yin, 2014, p. 132). At the analytic stage of a case study, "much depends on a researcher's own style of rigorous empirical thinking, along with the sufficient presentation of evidence and careful consideration of alternative interpretations" (p. 133). "The main activity of cross-case analysis is reading the case reports and applying their findings of situated experience to the research questions …" (p. 47). Analysis of the pilot study provided analytic themes (Table 3.2) and coding categories for organizing the data in the cross-case analysis and synthesis.

Table 3.2. *Analytic themes developed from the pilot study.*

Theme 1: Where did learners seem to have difficulty?
Theme 2: What did learners find easy?
Theme 3: What did learners remember from one session to the next?
Theme 4: What skills did learners demonstrate?
Theme 5: What evidence of conceptual understanding did learners display?
Theme 6: What evidence of metacognitive awareness did leaners show?
Theme 7: How well was the balance between the learner's sense of autonomy and tutorial guidance maintained? How did this vary between tasks, sessions, and cases?
Theme 8: To what extent does evidence support elements of the composite theoretical framework? (Figure 2.3).
Theme 9: What new or unexpected elements emerged from the data?
Theme 10: Which elements of the theoretical framework were useful? Which were not?

Trustworthiness of the Study

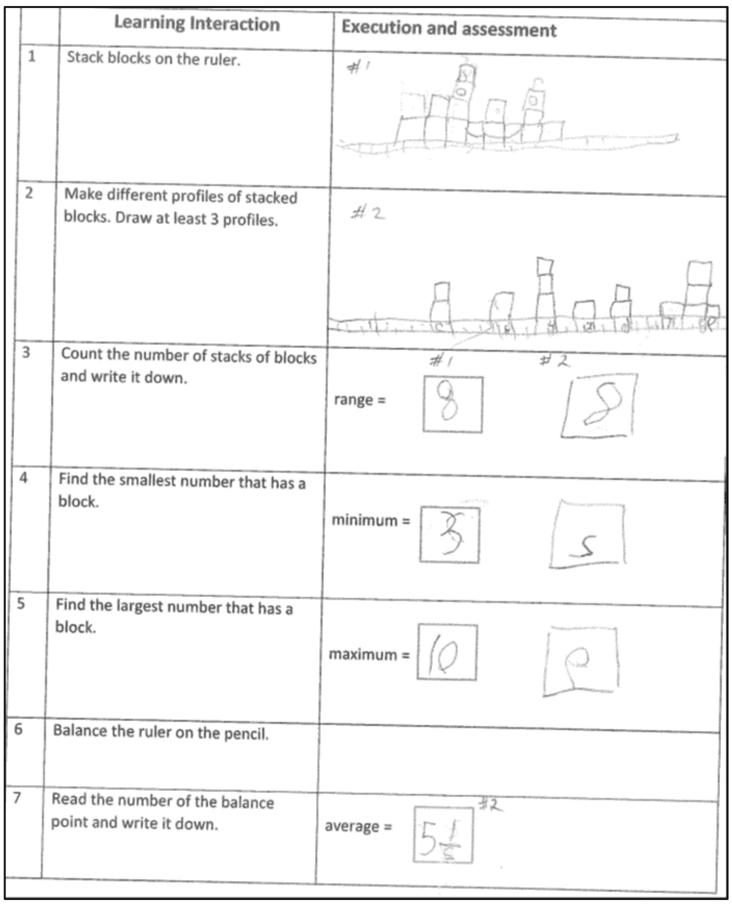
The quality and integrity of a qualitative study is articulated in terms of trustworthiness and transferability (Beaudry & Miller, 2016). Trustworthiness encompasses the idea of *validity* – the degree to which the overall approach and methods can be trusted to meet the objectives of the study; and *reliability* – the appearance of common patterns across multiple cases. Trustworthiness was addressed through several strategies: conducting a pilot study, coding the activities and interactions and identifying themes for the main study; describing the researcher's background, motivation, and philosophical orientation (positionality); triangulation of data (video, audio, worksheets, field notes, analytics memos); peer examination of selected videos (interaction analysis); verbatim excerpts from learning sessions (thick description); and recording researcher's reflections (Beaudry & Miller, 2016, pp. 52-53; Merriam, 1998, pp. 204-205).

Pilot Study

A pilot study was conducted to see if a five-year-old[xiii] might begin learning statistics through block play (Giebitz, 2015). The tutor guided her through a series of activities while providing encouragement and minimal scaffolding (Wood et al., 1976). Using blocks to represent data points and a ruler as a horizontal axis, she stacked blocks on the ruler to create an embodied frequency distribution. She was free to stack the blocks any way she chose. After a couple of turns stacking the blocks haphazardly, the tutor showed her how to align the blocks vertically and horizontally according to the markings on the ruler. The tutor was alert to how she responded to guidance and how she interacted with the materials, allowing her a high degree of autonomy and freedom to play. The apparatus went through several cycles of redesign, finally settling on three components shown in Figure 3.3: a 12-inch wooden ruler, toothpicks, and one-inch cubes made of hardwood (maple). The number of blocks varied from 18 to 28.

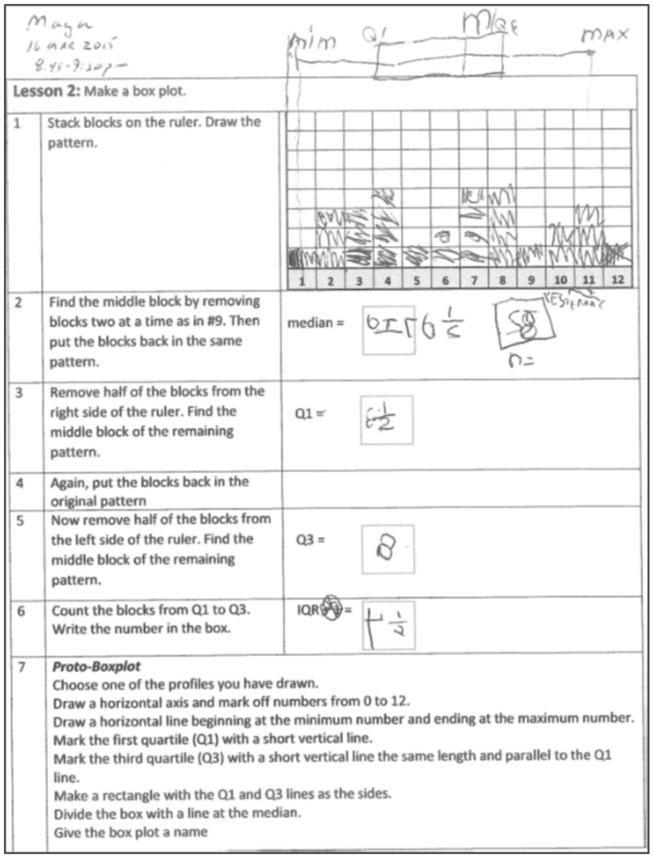
Figure 3.3. The participant called this configuration of blocks "The Crab." The tall stacks on the ends are the claws and the stack in the middle is the head. Fleer (2011) proposed that imagination is the bridge between play and learning.

Figure 3.4 shows a worksheet where the participant drew a picture of the blocks as she arranged them on the ruler then embellished several blocks by adding faces and antennae, reflecting the playful nature of the activity. Then she found the minimum (min), maximum (max), and range (R) following the tutor’s verbal guidance. She balanced the ruler on the fulcrum and read the average (5½) at the balance point. This was her first exposure to fractions. Figure 3.5 shows another worksheet where instead of drawing the blocks, she shaded boxes in a template. This was more efficient. Shaded boxes were later replaced by X’s inside the boxes, hence the *X-plot*. She found the median, first quartile (Q1), third quartile (Q3), and interquartile range (IQR). Then she made a *proto*-boxplot and labeled each point. In later sessions, she used dice to generate data for configuring the blocks on the ruler.



	Learning Interaction	Execution and assessment
1	Stack blocks on the ruler.	#1
2	Make different profiles of stacked blocks. Draw at least 3 profiles.	#2
3	Count the number of stacks of blocks and write it down.	#1 #2 range = 8 8
4	Find the smallest number that has a block.	minimum = 3 5
5	Find the largest number that has a block.	maximum = 10 9
6	Balance the ruler on the pencil.	
7	Read the number of the balance point and write it down.	average = $5\frac{1}{2}$ #2

Figure 3.4. This worksheet shows early examples of how the student arranged blocks on the ruler and recorded measures of dispersion and central tendency. This was her first encounter with fractions. Notice the embellishments of the top blocks.



Lesson 2: Make a box plot.		
1	Stack blocks on the ruler. Draw the pattern.	
2	Find the middle block by removing blocks two at a time as in #9. Then put the blocks back in the same pattern.	median =
3	Remove half of the blocks from the right side of the ruler. Find the middle block of the remaining pattern.	Q1 =
4	Again, put the blocks back in the original pattern	
5	Now remove half of the blocks from the left side of the ruler. Find the middle block of the remaining pattern.	Q3 = 8
6	Count the blocks from Q1 to Q3. Write the number in the box.	IQR =
7	***Proto-Boxplot*** Choose one of the profiles you have drawn. Draw a horizontal axis and mark off numbers from 0 to 12. Draw a horizontal line beginning at the minimum number and ending at the maximum number. Mark the first quartile (Q1) with a short vertical line. Mark the third quartile (Q3) with a short vertical line the same length and parallel to the Q1 line. Make a rectangle with the Q1 and Q3 lines as the sides. Divide the box with a line at the median. Give the box plot a name	

Figure 3.5. This worksheet provided a grid for reproducing an arrangement of blocks on the ruler and recording the five statistics needed to make a boxplot. In making the “proto-boxplot”, the student drew a line from min to max to represent the range. Then she made a small vertical line at the median. This structure shows pictorially simple measures of dispersion and central tendency. Then she added the box extending from Q1 to Q3 to complete the boxplot.

The participant found the mean absolute deviation (conceptually comparable to the standard deviation) in a three-stage process shown in Figure 3.6. First, she found the average by balancing the blocks on the ruler (6½). Then she made two columns listing the position of each block in the first column and its distance from the average in the second. Then she arranged the blocks on the ruler according to the new configuration in the second column. She then balanced the ruler on the fulcrum and read the mean absolute deviation directly from the ruler. She then recorded this value (1¾) on the worksheet. Although she did not demonstrate conceptual understanding of this measure of dispersion, this activity showed that such a measure can be demonstrated through block manipulation and thus lay a foundation for later conceptual understanding.

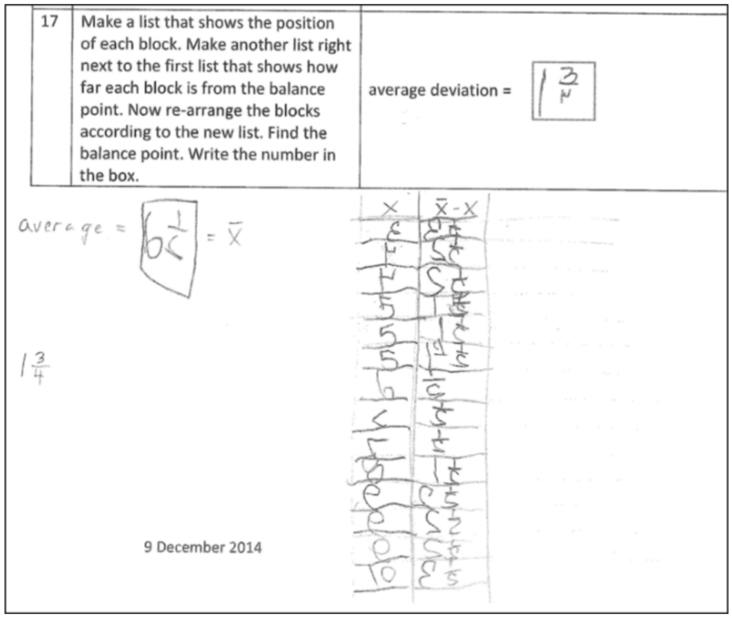


Figure 3.6. This worksheet shows the procedure for finding the mean absolute deviation. The tutor wrote the algebraic column headings and guided the student through the procedure.

Four sessions lasted from 20 to 50 minutes. Sessions were videotaped and transcribed then analyzed along with field notes, analytic memos, and worksheets. Initial coding of transcripts, field notes, memos, and documents was open ended following Saldaña's (2013) *pragmatic eclecticism* (p. 60). A combination of descriptive and process coding emerged after the first cycle of open-ended coding (Saldaña, 2013, pp. 87-101) which resulted in the following categories:

> Actions of learner: *sing/hum, stack blocks, remove blocks, draw, count, write number, play, off task, request clarification, ask for help, self-correction, ask for materials.*
>
> Actions of tutor: *verbal instructions, demonstrate, give feedback (encourage performance, encourage autonomy, redirect, affirm), give guidance, clarify instructions.*

A provisional view of the open coding suggested three second cycle axial codes: 1) *independent actions of the learner*, 2) *actions of the tutor*, and 3) *learner's response to* *actions of the tutor*. This suggested a theme (selective coding) of *balance between learner autonomy and tutorial guidance*. This finding is consistent with a meta-analysis of ten years of research in inquiry-based science teaching using a two-dimensional framework with a guidance dimension and a cognitive dimension (Furtak et al., 2012). This meta-analysis found that for the most part guidance enhances learning.

Research on guided statistical inquiry at the elementary level is in its early stages (Fielding-Wells, 2018; Makar, 2016). Validated assessments at this level have yet to be developed. Inquiry-based approaches require a higher level of teacher preparation than traditional methods (Barron & Darling-Hammond, 2008); teacher professional development is an importatnt area of future research (see Chapter 5). A report prepared by the Organisation for Economic Co-operation and Development (OECD, 2012) on preparing teachers for the 21st century stated, “if only traditional learning outcomes are assessed, then inquiry-based and traditional methods of instruction appear to yield similar results. The additional benefits from inquiry learning – namely, that it nurtures communication, collaboration, creativity and deep thinking – become apparent when the assessments try to determine how well the knowledge that has been acquired is applied and when they measure the quality of reasoning” (p. 40). The proposed learning outcomes of data literacy and statistical literacy were difficult to assess. A fair assessment would require demonstration of competence in statistical inquiry which was beyond the scope of this study.

Positionality: Researcher’s Background and Motivation

I worked as a manager in the automotive industry at a time of transition from outmoded industrial practices and attitudes to a manufacturing system patterned after the Toyota Production System (Liker & Meier, 2007; Ohno, 1978; Schonberger, 1986). Later, as Quality Manager in a mining and metals processing company, I applied what I had learned in manufacturing. I formed a team to implement a data-driven system of management[xiv]. My colleagues and I developed tools and methods for systematically improving work processes in collaboration with frontline workers. We eliminated common frustrations, improved operations, and reduced waste. We changed adversarial customer relationships into long-term collaborations in continual improvement of product quality, packaging, transportation, delivery, communication, and problem solving. We saw problems and failures as learning opportunities following Deming's admonition to eliminate blame from the work environment. We cultivated an organizational culture of collaborative learning through investigation, data, and dialogue.

An organization is a system of interrelated processes. Well-chosen performance measures provide insight into how well the system is performing. A carefully crafted set of performance measures can provide the feedback needed to monitor, control, and improve essential processes (Breyfogle, 2003; Kaplan & Norton, 1996) if the necessary data literacy and leadership is in place. An integrated, data-based view of the organization is essential, but it cannot replace expertise, tacit knowledge, and common sense. Data derive their meaning from human judgment; data don't tell stories, people do. As data are acquired, organized, analyzed, interpreted, and reported, the cost rises while the value of data decays with time. Minimizing the cycle time from data capture to its use can maximize its value. A well-designed dashboard presents timely critical information at a glance, like the dashboard of a car or the control panel of an aircraft, for rapid data-informed decision-making.

A statistically literate person can assess empirical evidence in numerical form to draw accurate conclusions about conditions and processes in the real world. Skills of statistical literacy include data collection planning; data acquisition and organization; the ability to evaluate a sample; measurement system analysis; the ability to conceptualize and interpret variability; the ability to use graphical and analytical tools and techniques to communicate effectively; the ability to interpret patterns revealed through the application of those tools and techniques; the skills of dialogue and discussion to engage in constructive and open consultation on a range of possible interpretations of the data and their implications. A statistically literate person can assess the strength of evidence for and make judgments about the conclusions of others based on statistical arguments, such as the claim of a cause-effect relationship. Collaborative interpretation of data is most effective.

Statistical literacy was central to my work in management, organizational and process improvement. The active learning and project-based methods used to teach statistics in industry were in stark contrast to the textbook teaching common in traditional classrooms. After leaving industry to teach at a university, I wondered how the more dynamic and practically-oriented methods used to teach statistics in industry might be adapted to a classroom of adults. This led me to wonder how adults might be better prepared to learn statistics considering the anxiety and misconceptions they often brought to the classroom. My training in quality management had instilled in me a strong orientation toward *preventing* problems rather than solving them after they had occurred. One strategy to develop statistical literacy is to begin with play and other active learning methods in kindergarten then introduce project-based and inquiry-based methods.

Whether we look at the inability of managers and engineers to communicate statistically, the legions of executive decision makers thinking in terms of “averages only,” the erroneous claims of researchers published in leading scientific journals, the uncritical reporting of statistical claims by the media, or the visceral aversion so many students have toward taking statistics courses – then contrast this with the sense of empowerment that comes with acquiring powerful tools to investigate the world – there is much to be gained by finding better ways to educate citizens in statistical literacy. This study is a contribution to that effort.

Philosophical Context

Learning occurs through the power of reflective thought manifested in action. John Dewey (1910) defined reflective thought as “active, persistent, and careful consideration of any belief or supposed form of knowledge in light of the grounds that support it” (p. 6) and its implications. Through the power of thought, human beings shape their world. Learning is transformative. Paulo Freire (1970) believed human beings “are authentic only when engaged in inquiry and creative transformation” (p. 65). Confucius affirmed that the investigation of things is at the core of personal and social transformation (Chan, 1963). Aristotle (350 BCE) maintained that the desire to know is intrinsic to human nature. Although the process of inquiry appears to be a universal human pursuit, it proceeds according to a particular view of the world.

A worldview is a combination of beliefs, assumptions, attitudes, values, and ideas that form a comprehensive model of reality (O’Sullivan, 2012, p. 164). Worldviews vary widely between cultures. The following comparison of three worldviews illustrates important differences in ways of thinking and knowing. Western worldviews are generally linear and reductionistic. Indigenous American philosophies generally hold “a circular worldview that connects everything and everyone in the world to everything and everyone else, where there is no distinction between the physical and metaphysical worlds and where ancestral knowledge guides contemporary practices” (Brayboy & McCarty, 2010, p. 190). In a traditional Chinese worldview, there are universal patterns in the heavens, within ourselves, and in human affairs. As we come to discern these patterns, we become increasingly attuned to the moral principles that guide our decisions and maintain balance and reciprocity in our relationships. Human beings “are interdependent with the world in which they reside, simultaneously shaping it and being shaped by it” (Lau & Ames, 1998, p. 20).

Western science generally sees “truth” as existing outside of time and space and independent of the individual. However, contemporary Western thought has also given rise to an exaggerated Individualism and sees the autonomous individual as “the locus of morality guided by individual conscience and private judgment, seeking no greater purpose than individual fulfillment” (Lange, 2012, pp. 200-201). This tension between truth and morality can give rise to “moral inversion” (Polanyi, 1958, p. 233) where a notion of truth, regardless of the strength of evidence to support it, assumes moral authority and methods developed from that particular worldview assume a privileged status. Thus, methodology and axiology are confounded – methodology is endowed with a moral quality of conducting one to the “truth” and the “truth” thus found is endowed with moral authority. “That form of science that emphasizes its own methodology exclusively implies the disappearance, even the expulsion, from scientific thought of all considerations based on value, perfection, harmony … subjective reality is replaced by efficiency and materialism …” (Anshen, 1986, p. xvi).

From the perspective of Native science, “truth is not a fixed point, but rather an ever-evolving point of balance, perpetually created and perpetually new” (Cajete, 2000, p. 19). There is a dynamic coherence between our consciousness, our perceptions, and patterns in the universe. “Native science at its highest levels of expression is a system of pathways for reaching this perpetually moving truth or ‘spirit’” (p. 19). Many paths lead up the mountain of true understanding; there is no single Indigenous worldview. However, there are some important differences between typical Indigenous and Western ways of knowing. Indigenous knowledge is situated or *emplaced*. “Cradled in the context of specific landscapes, knowledge is raised. The landscape – the places where teaching and learning take place – is not just a blank backdrop for the journey, but the locus of the power to move through a knowledge-seeking journey” (Brayboy & McCarty, 2010, p. 187). What counts as knowledge, how we use that knowledge, the place where that knowledge is gained and where it is used, and how the community benefits from using it are inseparable facets of Indigenous science (Cajete, 2000). “Pursuit of knowledge and application of knowledge is one process anchored in moral values” (Green, 1980, p. 207). Ontology, epistemology, methodology, and axiology are facets of a whole (Lincoln & Guba, 1985). The principles of reciprocity and balance demand that we maintain an attitude of respect toward all the elements of our investigation. The powers of reason and sense perception enable us to investigate reality and discover truth. But reason and intuition, the material and the spiritual, are inseparable.

From the Chinese perspective, “‘knowing’ is the unraveling and the coordinating of the patterns of continuity that emerge and persist in the natural, social, and cultural flux around us … always practical, contingent, and moral: it is a ‘doing’ rather than a state of mind. Further, ‘knowing’ is meliorative – it makes a situation better” (Lau & Ames, 1998, pp. 21-27). Thus, knowing is not so much about *agency* as it is about *participation* in a larger cultural and cosmic framework. Rather than “truth” in any definitive or abstract sense, Chinese knowing is more concerned with continuity and coherence within the flux of an ever-changing present – an *unfolding*. The *Great Learning*[xv] (Chan, 1963) is a synopsis of a way of life that seeks knowledge as integral to finding peace and balance within oneself and in human affairs. It is one of the Four Books of Confucian philosophy. At its heart is the idea of the investigation of things. By observing the patterns and learning the principles operating in the universe, we extend our knowledge and clarify our thinking. This provides a basis for cultivating our moral qualities and bringing order to our families, our institutions, and the world. The truths that matter most are moral truths; scientific investigation deepens our understanding of both the physical world and how we ought to conduct ourselves in it. This *Way of Learning* does not seek to impose a truth from above, but rather challenges us to articulate our moral understanding based on our own investigation of things. Science is inherently axiological; self-knowledge is transformative. Independent investigation of truth – the use of our rational faculty – is a moral injunction.

Scientific reasoning is one of the greatest achievements in human thought (Einstein & Infeld, 1938). It enables us to see beyond the limitations of our senses and to challenge intuitive conclusions that may be wrong. Western science tends to decontextualize knowledge in pursuit of broad generalizations and clearly defined abstractions; it values precision, repeatability, and verifiable causal mechanisms. But we have to remember, Heisenberg (1958) reminds us, “what we observe is not nature in itself but nature exposed to our method of questioning” (p. 26). The positivist perspective that dominated Western science in the 20th century values clarity and precision and the ability of theory to predict from observations in the material world. These values are generally accepted by the wider scientific community. But positivism views the non-material dimension of human experience as outside the purview of science, irrelevant, or non-existent. Yet Heisenberg and other architects of quantum theory viewed the dismissal of that which we can’t measure or empirically observe as an impediment to science. “As far as science is concerned,” Heisenberg warned, “if we may no longer speak or even think about the wider connections, we are without a compass and hence in danger of losing our way” (quoted in Wilbur, 1984, p. 38). Einstein repudiated the ascendency of reductionism and dualism in Western science and its materialist worldview. He asserted that “without the belief in the inner harmony of our world there would be no science” (quoted in Anshen, 1986, p. 13). Einstein elaborated, “there is a structural kinship between subject and object, an indwelling of one in the other, and the error of the empiricist is to denigrate *experience* into a reductionist ontology that atomizes time, space, causality, and substance” (p. 16).

The scientific enterprise to understand the fundamental principles that govern the operation of the physical universe advanced by Einstein and Heisenberg has a self-correcting mechanism. Einstein recognized that the empirical data available were contradictory <u>if</u> one accepted the prevailing notions of time and space (Einstein, 1905). His genius lay in part in his acceptance of <u>all</u> the data along with the contradictions it entailed but relinquishing the “common sense” assumptions. Einstein’s predictions about the bending of the light by the sun’s gravitational field were confirmed. His use of data to unveil an underlying mechanism of physical reality is a powerful example of data-driven inquiry followed by data-based confirmation. However, in the social sciences, data-driven inquiry takes on a different aspect.

There is a lingering adherence to a Unity of Science agenda that seeks to reduce social science to physics in the hope of achieving the prestige of physics – *physics envy* (Ghoshal, 2005). But Unity of Science is attained, not by reducing biology, psychology, and sociology to physics and chemistry, but by the structural uniformities of each (Bertalanffy, 1968, p. 87). In describing the effect of a mechanistic worldview on our society[xvi] Bertalanffy writes, "Practically, its consequences have been fatal to our civilization. The attitude that considers physical phenomena as the sole standard of reality has led to the mechanization of mankind and to the devaluation of higher values" (p. 88). The principles governing the interaction of human beings are fundamentally different from those governing the interactions of particles (Wilber, 1984). Economist Friedrich von Hayek warned against "the danger posed by scientific pretensions in the analysis of social phenomena" (Ghoshal, 2005, p. 79). Paradigms in the experimental sciences are self-correcting (Kuhn, 1962), whereas paradigms in economics and management are self-justifying, self-validating, and self-perpetuating (Ghoshal, 2005). Both experimental and social sciences make use of data in their knowledge claims. But the belief that a global population should be subjugated to impersonal market forces is precisely the kind of moral inversion Polanyi described (Yeager, 2004). Those affected by the consequences of those claims (everybody) need to have the skills of inquiry and statistical literacy to come to their own understanding of both their validity and their implications. The impact of data-driven decision-making in government, business, economics, management, and other social sciences, and its implications for the well-being of society underscore the need for a statistically literate citizenry.

Although the logical positivists and their successors failed to gain the allegiance of the architects of modern physics, their influence lingers. For example, the common reference to the logico-deductive method as *the* scientific method or enshrining randomized controlled trials as "*the golden standard*" regardless of context are persistent reminders of how science can be co-opted by special interest groups. From within the Western scientific philosophical tradition, Paul Feyerabend (1975) challenged the positivists who claimed a position of privilege with respect to declarations of what counts as knowledge relegating whatever fell outside tangible, testable propositions as “vaporous nonsense.” Feyerabend opposed ideological hegemony. He maintained that even empirical observations are theory-laden and are incommensurable across paradigms (Kuhn, 1962). Feyerabend was concerned about science falling into dogmatism or becoming a tool of a controlling elite, a concern shared by Freire. He was concerned about great works of science being recontextualized for political ends. “The ideas of these great thinkers [Mach, Boltzmann, Einstein and Bohr],” Feyerabend wrote, “were distorted beyond recognition by the rodents of neopositivism and the competing rodents of the church of ‘critical’ rationalism” (p. xviii). He recognized that “*there can be many different kinds of science*. People starting from different social backgrounds will approach the world in different ways and learn different things about it” (xx). He opposed “ideologies that use the name of science for cultural murder” (p. xxii). And he cites Einstein’s admonition for the scientist to keep his conceptual framework from adhering too closely to any particular epistemological system. All citizens need to be equipped with the tools and education to investigate reality for themselves, including the ability to evaluate and challenge knowledge claims based on statistical arguments.

Learning Activities

In the activity system described in this study, the tutor provided guidance and support for a student to engage in a series of lessons designed to develop an understanding of a dataset as an aggregate of counts or measurements with shape, spread, and center, and to develop an initial understanding of variation. The learning trajectory followed a course of increasing cognitive complexity. After the analysis, the researcher redesigned the activities in accordance with the findings. Activities were playful to facilitate learning and help ensure a positive attitude toward statistics.

Students sat at a table with a set of blocks, a ruler, and three round toothpicks to serve as a fulcrum. The tutor asked their permission to turn on the camera and begin the activity. The camera was set to capture the student’s interaction with the blocks, including facial expressions. Sometimes the researcher removed the camera from the tripod to get close-ups of the worksheet. In the first session, the tutor invited the student to play with the blocks and ruler for a few minutes before beginning the first lesson. Then the tutor showed the student how to arrange the blocks on the ruler in neat stacks, each stack aligned with a number on the ruler. The tutor then invited the student to play with the blocks but with this added constraint. Then the tutor showed the student how to balance the ruler on the fulcrum. This fulcrum was initially three toothpicks, then a couple of pencils, and finally two short pencils glued together. The student read the balance point on the ruler at the center of the fulcrum. The tutor then introduced the worksheet. As in Wood, Bruner, and Ross (1976), the tutor created an atmosphere of approval and encouragement. In transitioning from free play, the tutor gently redirected the student's attention to the required task. The student set the overall pace of the activities. The tutor generally deferred to the student, being mindful of the student's sense of autonomy and playfulness.

Lesson 1: Free Form Block Play (find statistics and create data displays) [xvii]

The student stacked 15 blocks on the ruler and made an X-plot[4]. Then the student counted the blocks to find N; found the *range* by counting from the *minimum* to the *maximum;* balanced the blocks on the fulcrum, read the average directly on the ruler, and recorded it in the worksheet. The student systematically removed blocks two at a time to find the median, repeated this operation with the bottom half then with the top half to find Q1 and Q3. Then counted from Q1 to Q3 to find IQR. The student used the five-number summary to create a box plot. The student then labeled each of the five points on the boxplot. Students did not know fractions so they learned both the meaning and notation of halves and fourths.

Lesson 2: Dice Roll (generate data, find statistics, make an X-plot)

The student rolled a pair of dice and arranged the blocks on the ruler according to the outcome, then proceeded as in Lesson 1 making an X-plot, finding the 10 statistics and making a boxplot.

Lesson 3: Block Weights (measurement and data collection)

The tutor made a three-column data collection sheet for recording the weights of blocks. The student weighed each block on a small digital scale. The tutor showed the student how to zero the scale and explained the meaning of grams by showing the student food labels. The

[4] The X-plot is an innovation of this study; conceptually, it is equivalent to a dot plot.

student recorded block weights to a tenth of a gram in the data collection sheet then rounded the weights to the nearest gram. The student arranged the blocks on the ruler according to the distribution of weights. Then the student found the statistics as before (except Q1, Q3, and IQR) and made an X-plot.

Lesson 4: Marble Weights (measurement and data collection)

The tutor made a two-column data collection sheet for recording the weights of marbles. The student weighed each marble on a small digital scale. The tutor showed the student how to zero the scale and explained the meaning of grams by showing the student food labels. The student put a small piece of paper of the scale so the marbles did not roll off. The tutor explained that the scale needed to be re-zeroed with the paper on it. The student recorded the weight in the data collection sheet. The tutor transformed the scale of weights onto the ruler and the student arranged the blocks on the ruler according to the distribution of marble weights. Then the student found the statistics as before and made a “marble plot” – another innovation of this study. Then the student transformed the scale back to grams.

Lesson 5: Marble Grab: (compare data sets)

The tutor set a bowl of marbles on the table and asked, “Which hand is bigger, the left or the right?” To answer this question, the tutor made a three-column data collection sheet. Columns were labeled “Trial”, “Left Hand”, and “Right Hand”. The student grabbed a handful of marbles and counted them and entered the number in the appropriate column. Then the student did the same with the other hand. After five trials, the tutor engaged the student in a conversation about how the marble counts could be different, why one hand seemed bigger on one trial while the other hand seemed bigger on the next. This exercise is like Deming’s Red Bead Game and it led to conversation about variation. The student found min, max, range, mean, and median for each hand.

Lesson 6: Find the Mean Absolute Deviation

The student arranged the blocks on the ruler either as a Free Form (Lesson 1) or a Dice Roll (Lesson 2) then balanced the blocks on the fulcrum and noted the average on the worksheet. The tutor provided a worksheet with two columns. In the first column, the student wrote the position of each block on the ruler. In the second column, the student wrote the distance of each block from the average. Then the student re-arranged the blocks on the ruler according to the second column and found the balance point. This was the mean absolute deviation and the student entered it on the worksheet.

Chapter 4: Findings

This study showed *how play with an embodied frequency distribution in the form of blocks arranged on a ruler, under the guidance of a tutor, might facilitate learning statistical concepts and skills beginning in the first years of formal education* (first research question). Students in grades K-2 under the guidance of a tutor demonstrated through block play procedures for finding mode, minimum, maximum, range, mean, median, first quartile, third quartile, interquartile range, and mean absolute deviation of a data set. They made visual representations of data distributions, including adaptations of the dot plot (X-plot and marble plot) and box plots. They built data sets using a data collection sheet to record measurements from a digital scale and dial caliper. They compared data sets of handfuls of marbles drawn from a bowl alternating left and right hands. These students showed that learning statistics can be enjoyable and does not require mathematics. Using measuring instruments exposed them to fractions and decimals in the normal course of data collection suggesting that mathematics education might begin with statistics.

Recent studies have explored introducing 4- to 6-year-old children to statistical inquiry in a holistic and coherent fashion in classroom settings (e.g., Fielding-Wells, 2018; Makar, 2016, 2018). The microgenetic methods used in this study suggest ways that one-on-one tutorial interactions might familiarize young learners with the conceptual structures used in statistical inquiry and reveal aspects of their thinking relevant to developing skills and conceptual understanding of fundamental statistical concepts. This study explored *sequences of tasks and activities that comprise a learning trajectory toward conceptual understanding of variation* (second research question). Analysis of these learning interactions suggested a redesign of the lessons described in Chapter 3. The redesigned lessons are described at the end of this chapter. This chapter also presents a synthesis of the analytic themes presented in Table 3.2 along with tutor reflections. Then it presents a brief description of each of the lessons described in Chapter 3, including excerpts from video transcripts, student artifacts, and photographs. The Chapter concludes with a summary of limitations and lessons learned. *How these findings might inform the design of learning experiences for adults and instigate innovations in statistics education for all ages* (third research question) is largely conjectural and is addressed in Chapter 5.

Analytic Themes and Tutor Reflections

This chapter explores the analytic themes shown in Table 3.2. The theoretical framework of embodied learning shown in Figure 2.3 was a useful starting point for designing this study, however, Activity Theory (Figure 2.4) proved to be a more comprehensive and useful model for describing the findings and for charting future studies. In this model, both the student(s) and the tutor are learners and participants. The element of play is made explicit by incorporating Fleer's theory of Conceptual Play into the model. The central theme emerging from the pilot study was the *balance between learner autonomy and tutorial guidance* while maintaining an element of play in the activities. The role of the tutor includes being attentive to activities and circumstances that lead to fatigue, boredom, or frustration and then to shift the dynamics of the interaction toward more playful learning interactions. The tutor's role is to maintain a trajectory toward the learning goals of embodied experience of the shape, spread, and center of the data set – the distribution of blocks on the ruler. In this context, the tutor asked the student what was fun, what was easy, what was hard, and what was remembered from previous sessions. In addition, themes of conceptual understanding and metacognitive awareness were examined. We will first explore the theme of what students remembered a week (in some cases more) after the session.

What do you remember from last time? (Participant names are pseudonyms.)

The first session began with building rapport and introducing the student to the materials: blocks, dice, ruler, pencils, marbles, and worksheets. Then with each session after the first, the tutor asked, "What do you remember from last time?" The following is a summary of student responses[5]. After the summary, excerpts from conversations with Fiona provide a detailed look at some of these interactions. Fiona was the youngest participant in the study. She was also the most talkative and demonstrated an active imagination.

Fiona-2 I put them in a special place on this ruler and I tried to balance them on the pencils. The Q1 was hard [22:40].

Layla-2 We were playing with blocks. We were playing with them so we could know how many they are. I remember balancing the blocks on the pencils.

[5] The number after the name is the number of the session. For example, Fiona-2 is Fiona's second session.

Layla-3 We were still balancing blocks. X-plot. That we were doing the X with the squiggly line. The max and the mode (that's all I remember). [When the tutor showed her a blank worksheet, she named all of the statistics. She was very articulate.]

Layla-4 That we did the X-pot and … the statistics; the blocks and the dice; the ruler and the pencils that are glued together.

Edgar-3 I don't know. [Edgar was more prone to get off track than were other students.]

Edgar-4 We were making our own stacks with these blocks. We were playing with the marbles at the end.

Mary-2 We weighed all the blocks. We tried to get the blocks to balance on the ruler. Some blocks weighed more than other blocks. [This may have been an opening to talk about variation.]

Mary-3 We balanced the blocks with pencils last time. I remember using the caliper? It was fun.

Mary-4 We put the blocks on the ruler and balanced the ruler on the pencils, but the pencils kept sliding apart.

Carla-2 Dice roll.

Carla-3 So, we, first, we put, we rolled the dice and we put the numbers that we got in here [pointing to the scale] and put a block on the number that we found. And if we got the same numbers we would put them back on. And when we were done with the rest and there was no more blocks, we put, we did this [slides the ruler] and we slided the pencils under. [Carla is gesturing throughout the explanation.] And then we found like the maximum and minimum and range and the Q1 and Q3 and stuff. And then we. After we did that [1:00], we started to weigh the blocks. [She rolls the dice, places a block on the ruler, and continues until all 15 blocks are on the ruler.]

For the most part, students remembered things they did (enactive mode) and their data representations (iconic mode), but they often did not remember terms like *median* and *average* and their respective symbols (symbolic mode). The following exchange from Fiona's fourth session shows that she remembered the structure she made with blocks and the imaginative story she told about it.

Tutor: Let's see what you remember from last time. What do you remember from last time? Or the time before that? Or the time before that? [playfully]

Fiona: I made a castle with stairs.

Tutor: You did?

Fiona: And an IN door and an OUT door.

Tutor: You made a castle with stairs and an IN door and an OUT door. Did you roll the dice?

Fiona: Um hmm.

Tutor: And what …

Fiona: No, I didn't roll the dice yesterday.

Tutor: You didn't! Well, did I show you how you can roll the dice and then put the blocks on?

Fiona: Yes. I know how to do that. [1:00]

Tutor: Let's see. (Transcript FB4).

Fiona then demonstrated the Dice Roll. If Fleer's theory of Conceptual Play is sound, then this combination of play and imagination is likely to lead to learning. If Dienes is correct, then Fiona will need to make multiple embodiments of a frequency distribution for conceptual understanding to develop. These questions were beyond the scope of the present study.

The following exchange is from Fiona's fifth session. She remembered the box plot and the statistics. But the Marble Grab was especially engaging.

Tutor: … Now, what do you remember from what we did last time?

Fiona: I remember that we did a box plot.

Tutor: Yeah. Tell me about it.

Fiona: And we did the marbles.

Tutor: We did a box plot.

Fiona: And the marbles.

Tutor: How would you like to do... [Fiona begins to speak.] Go ahead.

Fiona: With the grabbing marbles and then counting. [She gestures with her hands.]

Tutor: We did what? [1:00]

Fiona: We did the grabbing the marbles and then counting. [She repeats the gestures of grabbing and the counting.]

Tutor: We did grabbing the marbles and then counting them?

Fiona: How many marbles [The tutor talks over her then stops and lets her finish.]

Fiona: How many marbles we could hold. Like this. [She gestures as if holding a handful of marbles.]

Tutor: What else did we do?

Fiona: We had to figure out how those worked [pointing to the statistics on the blank worksheet]

Tutor: We had to do what?

Fiona: We had to figure out how those worked.

Tutor: Had to figure out how these worked?

Fiona: Yeah.

Tutor: We call these statistics.

Fiona: How the sta_. …Whatever they're called.

Tutor: Sta-tis-tics.

Fiona: Statixtix.

Tutor: Yes. Statistics.

Fiona: How to do the statixtix.

Tutor: Sta-tis-tics. [Fiona plays with her plastic iguana.] And what was the most fun of what we did last time?

Fiona: Balancing.

Tutor: You liked the balancing?

Fiona: Yeah. (Transcript FB5).

The original research question did not foresee the role the marbles would come to play in the lessons. The students seemed to relish thrusting their hands into the marbles and drawing out a handful.

The following exchange from the sixth and final session indicates that she did not remember even the simplest symbol *N*, but she did remember the *outlier* which intrigued her. Talking with her fingers in her mouth is a reminder of just how young these students are and the surprisingly advanced level at which they are able and willing to engage with statistical concepts and structures. She also completed the task of transposing numbers between marble weights and the numbers on the ruler.

Tutor: Can I turn on the camera?

Fiona: Ok. [Fiona picks up her water bottle in her teeth.]

Tutor: So, what do you remember from last time? [Tutor shows Fiona the marble plot she made in session 5 (Figure 4.13)].

Fiona: Mm. I did a marble plot and the box plot.

Tutor: You did a marble plot and a box plot.

Fiona: And there's an outlier [pointing to the outlier on the marble plot].

Tutor: Yeah. There’s an outlier.

Fiona: And I did the minimum, the maximum, the range, and the Q1 the Q3 and the IQR and the N. [She is reading these on the marble plot.] [1:00]

Tutor: Fiona, could I ask you to do something?

Fiona: What?

Tutor: Could you take your fingers out of your mouth? That way when I go back and listen again, then I'll be able to understand what you said. That will be very helpful to me. [She takes her fingers out of her mouth and repeats.]

Fiona: I did the minimum and the maximum and the N and the range and then the [pause] the X-squiggle [median] and then X-bar [average] and then the Q1 and then the Q3 and then the IQR.

Tutor: That's very good. Do you remember another word we used for X-squiggle?

Fiona: Mm. No. [She smiles and looks at the tutor.]

Tutor: We call that the *median*.

Fiona: Yes!

Tutor: But, you can call it the X-squiggle if you want to. Do you remember what we called the X-bar?

Fiona: Mm. X-bar. [Fiona looks up.]

Tutor: Aaavv… [prompting her]

Fiona: Average!

Tutor: Yes. Very good. [2:00]

Fiona: What's the N? See, just right there [pointing to the N on the worksheet].

Tutor: N stands for number of blocks.

Fiona: Number of blocks.

Tutor: Yes. Do you want to finish this up? [In the previous session, Fiona did not transpose the statistics back to grams; she recorded them based on the scale of the ruler from 1 to 12.]

Fiona: Yes

Tutor: Do you remember what else we did? We changed the scale.

Fiona: Yes.

Tutor: Remember?

Fiona: Yes, we changed it to twelve.

Tutor: Right

Fiona: 1,2,3,4,5,6,7,8,9,10,11,12.

Tutor: Right. And do you know what we have here? You put the min, we have to label things. So, the min is 4.4 what. What were we doing here? Oh, let's do that first.

Fiona: 4.4 is nothing. [playfully]

Tutor: Ok.

Fiona: Oh, that's 1. Four-point-four is 1. [The lightest marble weighed 4.4 grams which was transposed to 1 on the ruler.]

Tutor: Let's go back and write a title. Whenever we have a graph [3:00] or a marble plot or a box plot, we should put a title on it. What were we doing here?

Fiona: A marble plot.

Tutor: What were we doing with the marbles?

Fiona: We were um measuring them.

Tutor: Ok. What were we measuring?

Fiona: We were measuring how much they weigh. (Transcript FB6).

The tutor continued guiding Fiona in transposing the remaining statistics to grams. Fiona then wrote a title on the marble plot: "marble wates." This example illustrates that the ability to recall facts may be no indication of the cognitive level of engagement.

What did you enjoy most about the session?

At the end of each session, the students were asked what they enjoyed most, what they found easy and what was difficult. They enjoyed grabbing handfuls of marbles and counting them; rolling the dice and putting the blocks on the ruler and balancing them; collecting data using measuring instruments; making X-plots and finding statistics. The following summarizes their responses.

Layla: Putting the X's in the boxes; writing; putting the blocks and balancing them on pencils; rolling the dice; we counted these (pointing to statistics on the worksheet); the X-plot; writing the weights of the marbles; playing marbles; weighing them; the marble plot.

Edgar: Stacking the blocks; "I liked how I used the block without playing with the dices"; "I made a marble race" [making marble races was Edgar's own diversion, not part of the lesson] ; "I want to play sta-tis-tics … statistics with marbles"; "make my own stack"; "counting with the marbles"; "playing with the marbles"; "I really want to play music." [The room contained many musical instruments.]

Mary: Figuring out how much the blocks weigh; rolling the dice to put the blocks on the ruler; learning about X-plots; rounding the numbers; making a box plot; figuring out the mode; using the caliper; doing the mode and Q1 and Q3; IQR; the Marble Grab; making pictures with the blocks; balancing the blocks; making the X-plot; and the … [pause] "I think it starts with an 's' … statistics."

Carla: Weighing the blocks; balancing the blocks.

The following conversations show Fiona's responses through all six sessions.

Fiona session 1

Tutor: … What did you like most about the session today?

Fiona: I liked [cough] stacking blocks as high as I could. Uh, are all of those …

Tutor: Was that your favorite part?

Fiona: Uh. Yeah. Cause I like stacking. I guess you could do this. [She starts stacking blocks.] This kind of stack is pretty cool. I used to play with Legos. All the time I go like this. I went like this [She stacked three blocks on top of four blocks.] Oh, yeah. This is fun. (Transcript FB1).

Fiona session 2

Tutor: Well, we're out of time. [50:00] So what did you like about what we did today.

Fiona: I liked picking up the marbles and counting how many.

Tutor: Uh huh.

Fiona: Because I ... I have to go to the bathroom. [Fiona goes to the bathroom.]

Tutor: There you are. [53:12] Ok. Let's just take a couple of minutes and tell me what you liked most about today. What we did today.

Fiona: I liked when I stacked with the blocks on here.

Tutor: When you stacked the blocks on the ruler?

Fiona: Uh huh. When I stacked the blocks on the ruler.

Tutor: What did you like better, playing with the blocks or playing with the marbles?

Fiona: Marbles! (Transcript FB2).

Fiona session 3

Tutor: Let me just ask you a couple of things before you go. Tell me what you liked best about what we did today.

Fiona: I liked [pause] doing the marbles.

Tutor: You liked doing marbles?

Fiona: Yes.

Tutor: What was your favorite part?

Fiona: It was doing the marbles.

Tutor: Grabbing the marbles? Or counting the marbles?

Fiona: Yup, just the whole thing of the marbles.

Tutor: The whole thing of the marbles?

Fiona: Doing the marbles; that's what I said. I said doing the marbles.

Tutor: Doing the marbles; now we know what “doing the marbles” is. Do you want to do marbles next week too?

Fiona: Yes.

Tutor: Ok, well before we finish…

Fiona: Doing the marbles will mean the whole thing of grabbing the marbles and counting them.

Tutor: Ok, now we know what “doing the marbles” is.

Fiona: And showing you my ideas.

Tutor: You had some good ideas today.

Fiona: Yes. (Transcript FB3).

Fiona session 4

Tutor: Well, did you have fun today?

Fiona: Yes.

Tutor: What was the most fun?

Fiona: The dice. Rolling the dice [she rolls the dice] and then putting them on [she turns over the ruler so the numbers are facing up] the right number. [She places a block on the 2.]

Tutor: Uh huh. What else was fun for you.

Fiona: Marbles. Picking up the marbles. [She plunges both hands into the marbles.] And then counting them: one, two, three, four, five, six, seven.

Tutor: What else was fun for you?

Fiona: Mmm. Mm. (Transcript FB4).

Fiona session 5

Tutor: Did you have fun today?

Fiona: Yes.

Tutor: What was your favorite part?

Fiona: It was making the box plot.

Tutor: What did you like about making the box plot?

Fiona: It was challenging.

Tutor: Yes, it was, wasn't it? What else did you like?

Fiona: Making the marble plot.

Tutor: Yeah. What else did you like?

Fiona: Mmm. Weighing the marbles.

Tutor: Yeah. What else?

Fiona: That's it. (Transcript FB5).

Fiona session 6

Tutor: So, Fiona, what was the most fun we had today?

Fiona: It was with the marbles. No, it was measuring; that was so fun. I was like stack, stack, stack. Ok, that's enough; stack, stack, stack. Ok, that's enough; stack, stack, stack.

Tutor: Stack, stack, stack?

Fiona: Yes, stacking on there. [She touches the scale.]

Tutor: Oh, stacking the blocks on the scale? [Earlier in the session, she stacked the blocks on the scale reading the weight with each additional block until all 15 were on the scale.]

Fiona: Yes.

Tutor: That was fun?

Fiona: Um hmm.

Tutor: What else was fun?

Fiona: Oh, it was measuring how many inches stuff was. [She picks up a pencil.]

Tutor: Measuring how many inches stuff was?

Fiona: And this may be 12 [holding the pencil]. Wait.

Tutor: You mean with the caliper?

Fiona: Yes. Ok, let's measure this. Eight! This is eight.

Tutor: With this? With this? [Tutor hands her a dial caliper] That was fun?

Fiona: Yes. And the other one.

Tutor: And the other one?

Fiona: Yes.

Tutor: That would be this one. [Tutor hands her the Vernier caliper.]

Fiona: Yes. (Transcript FB6).

Fiona enjoyed making a box plot because it was challenging. Although she seemed to get fatigued making box plots in earlier sessions, by the sixth session she was enjoying the challenge of it. Student engagement can be engendered by either enjoyment or challenge.]

What was hard?

Fiona indicated in session 5 that making a box plot was here favorite part; it was fun because “it was challenging.” Difficulty can engender fatigue in one case and motivation in another. The following conversation illustrates that Fiona wondered why she was observing variation in the number of marbles she grabbed. This relates to the first research question, but rather than an arrangement of blocks on a ruler, variation in marble counts seems to be an effective way to introduce the concept of variation in a tactile-kinesthetic way.

Tutor: Let me ask you another question. What was hard?

Fiona: It was figuring out why [she picks up the data sheet] this was getting more than this [she points with her pencil to data on the worksheet]. [57:35]

Tutor: Why the numbers of marbles were different in your two hands?

Fiona: Yeah.

Tutor: Ok.

Fiona: Even though it was the same answer the next, the last time. [Fiona is wondering about the variation in the marble counts.]

*   *   *

Tutor: Was anything else hard for you?

Fiona: Figuring out what [she picks up the worksheet and looks at it closely] the X-bar, X-squiggle, Q1, Q3, IQR, and R, maximum, minimum, and the mode. [Before she finishes, her mother comes.] [58:36]

Tutor: Yeah. That was hard, wasn't it? You did it. Do you want to do it again?

Fiona: Yes. [58:44] (Transcript FB4).

Although Fiona acknowledged that finding the statistics was hard, she was able to consistently do so beginning in the first session.

Describe the shape of the distribution

To bring attention to the *shape* of the distribution, the feature most often neglected in evaluating a data set, the tutor asked the student about symmetry, gaps, and outliers. Then the tutor asked what the shape reminded the student of. Students saw such things as buildings, cars, boats, dolphins, and musical notes going up and down. The following interaction is an example of bringing the attention of a first-grade girl to the shape of the distribution after demonstrating what she remembered from the previous session.

Tutor: Can you show me what you remember from last time? [3:00]

Layla: Um.

Tutor: And we’ll do something a little bit different. Do you want to do something challenging today? [The tutor is referring to finding the mean absolute deviation which they do later in the session.]

Layla: Yeah.

Tutor: Do you? Ok. Um. So, go ahead and show me what we did last time. Rolling the dice and putting the blocks on the ruler and balancing it. Ok? And finding the statistics. And then we'll make a boxplot. And then we'll do something more challenging. Ok? [She nods] Does that sound like fun?

Layla: Yeah.

Tutor: You want to? [She nods and smiles] Alright. Go ahead. [She picks up the dice and proceeds with the Dice Roll exercise.] [3:38] I'll put these over here for you. Ok? [5:30] [Layla completes the rolling of the dice, looks up and smiles.] [6:36] Ok, what do you see there? [Figure 4.1 shows Layla's block arrangement.]

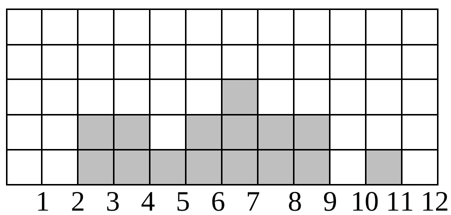


Figure 4.1. This arrangement of blocks on the ruler reminded Layla of musical notes going up and down.

Layla: Um. I don't know.

Tutor: What shape is that?

Layla: Like you know when you can have the … the music goes like up and down, up and down.

Tutor: Like the music. Like the music that goes up and down? You mean like the notes on the page [She nods her head]. Ok. [7:00] Yeah. And do you have any gaps?

Layla: Only one.

Tutor: Yeah. Do you have any outliers?

Layla: What are outliers again?

Tutor: An outlier is one that's way, way outside the others, far away. [She shakes her head.] No. Um. So, it looks like music? What else does it look like?

Layla: Um. That's it. (Transcript LC4).

The practice of first putting attention on the shape of the blocks emerged during the study. A note was added to the worksheet to ensure that shape was discussed before spread (range) or center (average and median).

Metacognition

Put simply, *metacognition* is “knowledge about and control of one's own learning” (Brown, 1992, p. 146). Fiona, a kindergarten student, demonstrated her awareness of the limits of her capabilities when she recognized she could not handle a dice roll of 11 because she only has ten fingers. This example illustrates the dynamics in a zone of proximal development (ZPD).

> Tutor: Do you want to roll the dice again? [33:45]
>
> Fiona: Ok. [She grabs the dice, shakes them, and throws them across the table.] Five plus six? I don’t have enough fingers for that. [34:00]
>
> Tutor: Oh no! Do you want to borrow one of mine?
>
> Fiona: Ok. [Smiling]
>
> Tutor: Here. You can borrow one of mine. Which one do you want to borrow?
>
> Fiona: Uhh. I don’t know.
>
> Tutor: This one? [Tutor holds up a finger.]
>
> Fiona: I guess.
>
> Tutor: Ok.
>
> Fiona: One, two, three, four, five, six, seven. Ok. Ten. [She extends a finger with each count.] Ok. Eleven. So, I’ll put a block on eleven. [She places a block on the ruler at 11]. [34:26]

The tutor recognized and acknowledged the student’s limitation and offered her a way forward with a spontaneous response to her need. Later in the same session, she began to question why she was getting so many 7’s.

> Fiona: [35:20] I want to roll some more. [She throws the dice.] Three plus four. I have enough fingers for that. So … 5 and 2. Ok. [She places a third block on the 7 and notices that the stack of 7’s is getting higher than the others.] Oh my God! Ok. I’m going to get ready for the next. [She shakes the dice in her cupped hands then throws the dice onto the table.] Weee! [One lands on the floor.] Ok. I don’t know what number I got right there.
>
> Tutor: You can roll again if you want. [She shakes the dice and throws them on the table a little less vigorously this time.]
>
> Fiona: Ok. Two [she accidently turns the die with her finger]. [36:00] Two [she rotates the die and the 4 is facing up] Where’d that 2 go? [She picks up the die and rotates it in her hand looking for the 2] I know I got a two [she finds the 2 and places the die on the table with the 2 facing up] Ok. Two and five. [She places the fourth block on the 7 then picks up the dice.] What if you don’t have enough to do? What if you’re tired of doing so many 7’s?
>
> Tutor: You could roll again, maybe you’ll get a different number.

Fiona: Ok. [She rolls a 5 and a 6.]

Tutor: That's a different number.

Fiona: Five and six. Ok, that'll be … [The video ends as she goes to put a block on 11. Figure 4.2 shows the final block arrangement.] [36:37] (Transcript FB1).

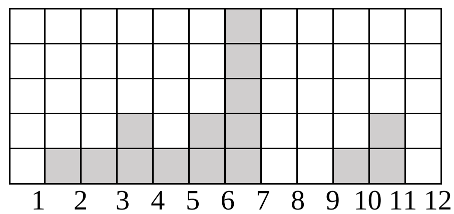


Figure 4.2. Fiona felt that she was getting too many 7's after the third 7. In a subsequent session when she got a preponderance of 5's, she referred to the 5 as "greedy" and "evil."

This was an opportunity to introduce some ideas about probability. Seven is the most likely outcome of rolling two dice. More could be done to integrate probability and statistics (stochastics) in future lessons.

Imagination

According to Fleer's theory of Conceptual Play, imagination is the bridge between play and learning. The following is a conversation about the shape of the distribution Fiona created. She periodically took control of the conversation with her question, "You know what?" followed by an interesting story, anecdote, or factoid. In the following exchange, she is describing a "castle" she built on the ruler with a door on either end – one for going in, and one for going out.

Fiona: And you know what?

Tutor: What?

Fiona: At the door of the castle it has … and it says "IN." That one says, "OUT."

Tutor: It has a sign?

Fiona: Yes.

Tutor: On the door?

Fiona: Yes. So, then the princess is like … but the out one is on the inside; that one is on the outside.

Tutor: The out one is on the inside, you mean the sign?

Fiona: Yeah.

Tutor: The out sign is on the inside of the out-door and the…

Fiona: IN sign

Tutor: …IN sign is on the outside of the in-door.

Fiona: Yes.

Tutor: Is that right?

Fiona: Yes.

Tutor: Ok, what else can you tell me about the shape?

The tutor tried to redirect FB's attention to the task but she continued explaining the logic of the signs on the doors of her castle.

Fiona: Because you can only come out when you're first starting inside.

Tutor: That's right.

Fiona: You can't go out when you're outside.

Tutor: That's right.

Fiona: And you can't go in when you're inside. And you can't go… and, but you can go in when you're outside.

Tutor: Ok, I get it. Can I ask you a question?

Fiona: Yeah.

Tutor: When you look at the shape of the distribution on the ruler, what else do you see?

In this exchange, the tutor is with Fiona in her imaginary castle but gently persists in re-directing her to an awareness of the important statistical concept that a distribution has a *shape*. Thus, learning a statistical concept in accordance with the intent of the first research question.

Balance between learner autonomy and tutorial guidance

The foregoing exchanges illustrate the back and forth of the tutor guiding the learner to a learning goal and the learner exercising her autonomy. In the pilot study, the balance between learner autonomy and tutorial guidance emerged as a central theme. Maintaining this balance proved to be a challenge. Video analysis revealed a need for greater metacognitive awareness on the part of the tutor to stay attuned to the learner. At times, the tutor seemed more focused on "covering the content" and making "progress" along a hypothetical learning trajectory than on the state and motivation of the learner. This was most apparent when finding the quartiles and making box plots. During one session, the tutor twice referred to Q1 as Q4. In another session, he instructed the student in the wrong procedure for finding Q1 and Q3. This finding suggests that the tutor needs a self-awareness check before beginning the session and perhaps a visual reminder during the session to stay more attuned to the learner and her learning needs than to achievement of a learning goal. Again, this underscores the importance of maintaining the balance between tutorial guidance and learner autonomy.

The tutor did, however, ask permission to start the camera at the start of each session. He offered the student a choice of lessons at transitions, such as blocks or marbles, Free Form or Dice Throw. The student always had the choice of ending the session at any time and enjoying free play. At the end of the session, the tutor asked if the student wanted to return the following week for another session. They invariably said yes. One student expressed a desire for more frequent sessions. Although the tutor did maintain an awareness of learner autonomy during most of the sessions, he periodically fell into an “achievement” mode rather than maintaining the play mode.

Improvements in mediating artifacts

Several improvements in the apparatus emerged during the study. The worksheets went through several revisions based on interactions between the tutor and students and between the students and the apparatus. The fulcrum evolved from three parallel toothpicks to a pair of pencils to a pair of short pencils glued together. The representation of the pattern of blocks evolved from a template with shaded boxes to the X-plot. The marble plot emerged from the learning sessions. The inclusion of an outlier marble was fortuitous – it was unplanned but introduced an important element of the shape and spread of the distribution. The *mode* was found to be a good way to begin finding the statistics since it brings the attention first to the shape of the distribution and is simple to determine by inspection. The worksheet was redesigned to bring attention first to the shape of the distribution and then to the mode as the first statistic.

Lessons in Play Learning with an Embodied Frequency Distribution

The following lessons revealed the capabilities of K-2 students to demonstrate procedural knowledge of foundational statistical concepts. They began with free play, then learned a few simple rules for playing with blocks to find statistics and make data visualizations. These lessons evolved during the study through the engagement of the students with the materials and the engagement of the tutor with the student. Students were introduced first to free play with blocks followed by Free Form block arrangements on the ruler followed by the Dice Roll. They were then introduced to Block Weights, Marble Weights, and Marble Grab, but not in a set order.

Free Play

The first session with each participant began with building rapport and free play with blocks as illustrated in the following exchange:

> Tutor: Tell me what you like about school.
>
> Layla: I like math after school.
>
> Tutor: You like what?
>
> Layla: Math.
>
> Tutor: What else?
>
> Layla: And um playgrounds.
>
> Tutor: Playgrounds? [2:00] And what else?
>
> Layla: And the classrooms.
>
> Tutor: What do you like about the classrooms?
>
> Layla: That they're all together so you can have um like that they're in the same room but in different [she coughs] doors. [she gestures with her hands].
>
> Tutor: Uh huh. OK. So, you have 2 classrooms right next to each other?
>
> Layla: Well like they're the same, but you have to go in the door to go to the next one.
>
> Tutor: Oh. Like this you mean? Like this door [Tutor is off camera] Um, what about blocks? Do you like to play with blocks? [Layla nods yes]. How do you like to play with blocks?
>
> Layla: Make shapes.
>
> Tutor: Really? [She nods yes] Can you show me? [Tutor slides the 15 blocks across the table toward her.]
>
> Layla: This way. [She has the blocks between her hands ordered into a 3x4 rectangle plus 3 loose blocks.] And … [3:00] a triangle [She manipulates the blocks.]
>
> Tutor: What's that?
>
> Layla: A triangle.
>
> Tutor: What else can you do? (Transcript LC1).

Layla then makes a rectangle with a triangle on top, “the shape of the sun”, and a teardrop. After seven and half minutes, the tutor introduced her to the Balance Blocks game. Figure 4.3 shows some of the students’ block creations.

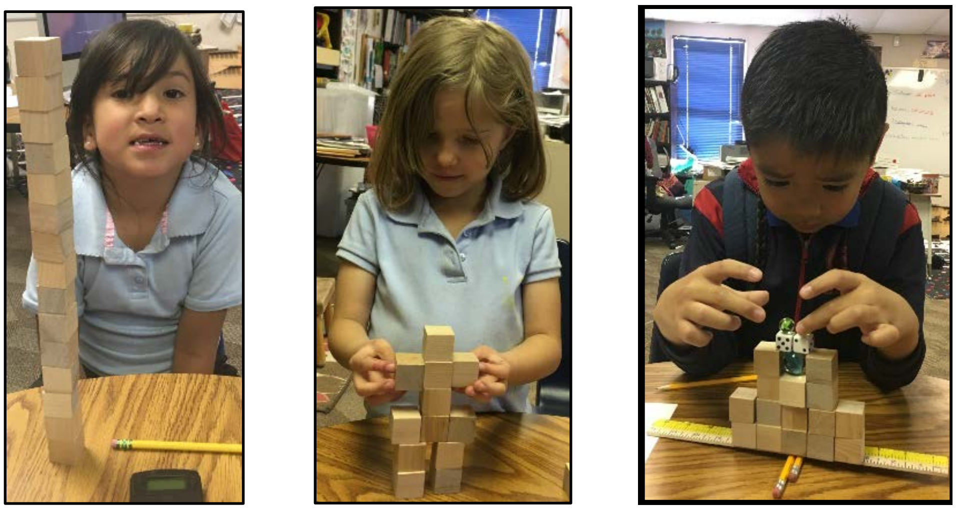


Figure 4.3. In free play, students played with the blocks, dice, pencils, ruler, and marbles. They stacked blocks in a single column, made representational figures and abstract creations.

Structured Play (Free Form)

After about five minutes of free play and conversation, the tutor introduces Edgar, a first grader, to the Balance Blocks game. Here, as in most cases, students’ free form block constructions were symmetrical.

Tutor: But I’m gonna show you something. A game I like to play with blocks. You make a stack like that. [Tutor slides the stack over to the ruler and begins placing the blocks on the ruler]. I’m going to place each block on the ruler over a number. So, we’ll put that on 4. That one on 3. You wanna help? [Edgar eagerly reaches over and starts placing the blocks on the ruler making a line one block high from 1 to 12.]

Edgar: I just have three left.

Tutor: Here’s what I do now [5:00]. I’m gonna put some blocks on top of other blocks just like that.

Edgar: So, you just know how to make your own stack. [inaudible]

Tutor: Yeah [He puts the 3 blocks he was holding on the stack then slaps his hands on the table.]

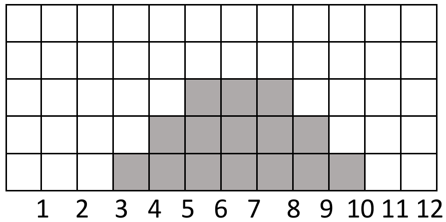


Figure 4.4. Edgar's first stack of blocks.

Edgar: Now do I just have to balance them on the ruler?

Tutor: Yeah. I gotta go get a couple of pencils.

Edgar: That's it?

Tutor: Yeah. I'll show you how I do it.

Edgar: Why do you have marbles here?

Tutor: Sorry? I'll show you how I do it. Then you can do it your way. See, this is the way that I do it. I put it so I can read the numbers. [Tutor slides the ruler with the blocks toward himself.] And then I just pick up one side and I can slide the pencils underneath [6:00]. Then I move it. Put the pencils together. [Tutor and blocks are off camera]. Then we get it to balance. And then. Let me show you. You can read the number where it balances. [The tutor slides the ruler balanced on the pencils over toward Edgar.]

Edgar: But do you just have to stack them more?

Tutor: I want you to read the number where it balances. You see right there? [pointing to the balance point]. What number is that?

Edgar: Seven.

Tutor: Yeah. It balances on the 7. So.

Edgar: So, can I make my own stack?

Tutor: Yup. [Tutor picks up the ruler with blocks and passes it over the Edgar. As he sets it down, the blocks fall off]. Upp! Make your own stack and see if you can balance it on the pencils. [Edgar stacks the blocks on the ruler.] [7:00]

Tutor: That's a pretty creative stack. [He lifts up the right side of the ruler and slides the pencils under it. He adjusts the ruler until it balances].

Edgar: There! [he declares triumphantly.] (Transcript EM1).

Although the blocks often fell off the ruler when the students picked it up to balance it, they never seemed to get frustrated. They just put the blocks back on and tried again until they succeeded. Then they showed satisfaction in their achievement.

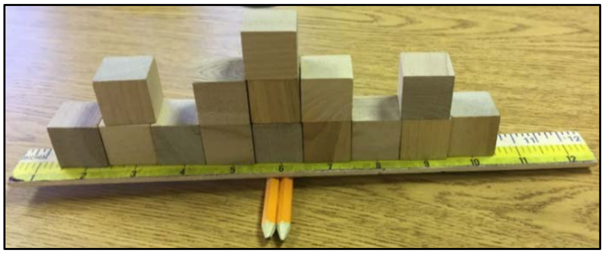


Figure 4.5. Students arranged the blocks on the ruler according to two simple rules: 1) align the blocks with the numbers on the ruler, and 2) stack the blocks in straight columns. Then they balanced the blocks on the ruler to find the average.

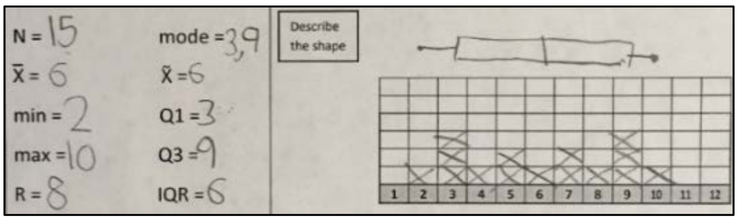


Figure 4.6. Students made an X-plot and found minimum (min), maximum (max), range (R), median, first quartile (Q1), third quartile (Q3), and interquartile range (IQR).

Figure 4.5 shows a typical block arrangement balanced at 6 on the ruler. These “Free Form” block arrangements tended to be symmetrical. Figure 4.6 shows a typical X-plot, the corresponding box plot and the ten statistics.

Dice Roll

In the Dice Roll, students rolled the dice and arranged the blocks on the ruler according to the dice roll. They described the shape in terms of symmetry, gaps, and the resemblance of the outline of the blocks to familiar things like buildings, stairs, cars, hands, and dolphins. They balanced the blocks to find the average (X-bar), made an X-plot, and found the statistics. Then thy made a boxplot and labeled the min, Q1, median, Q3, and max. Figure 4.7 shows a second grader engaged in the Dice Roll lesson. Figure 4.8 and Figure 4.9 show the X-plot, corresponding box plot and statistics for two of Layla’s Dice Roll lessons.

Figure 4.7. Carla makes an X-plot then balances the blocks on the ruler to find the average.

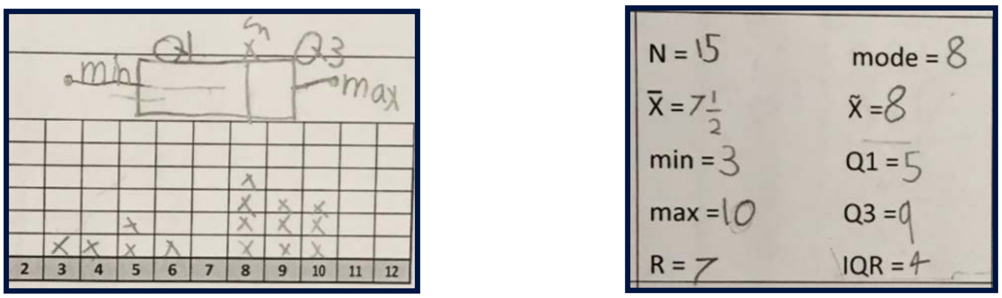
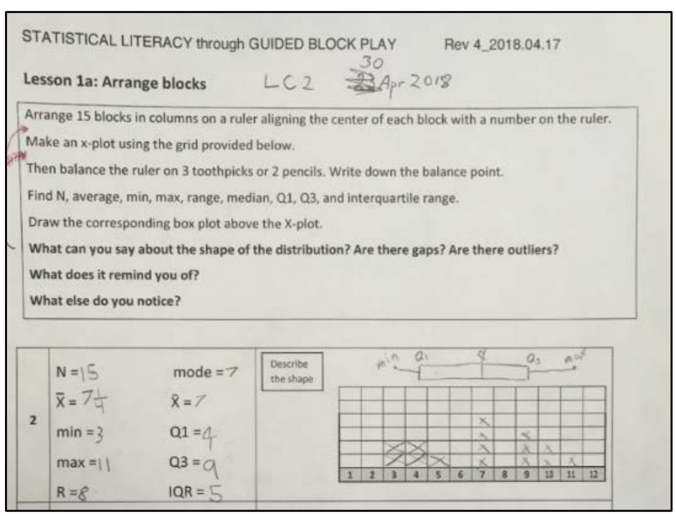


Figure 4.8. This X-plot and box plot is from Layla's third session. She found ten statistics and labeled the box plot. She was introduced to fractions in her first session.


STATISTICAL LITERACY through GUIDED BLOCK PLAY Rev 4_2018.04.17

Lesson 1a: Arrange blocks LC2 30 Apr 2018

Arrange 15 blocks in columns on a ruler aligning the center of each block with a number on the ruler.
Make an x-plot using the grid provided below.
Then balance the ruler on 3 toothpicks or 2 pencils. Write down the balance point.
Find N, average, min, max, range, median, Q1, Q3, and interquartile range.
Draw the corresponding box plot above the X-plot.
What can you say about the shape of the distribution? Are there gaps? Are there outliers?
What does it remind you of?
What else do you notice?

2	N = 15 X̄ = 7 1/4 min = 3 max = 11 R = 8	mode = 7 X̃ = 7 Q1 = 4 Q3 = 9 IQR = 5	Describe the shape

Figure 4.9. This Dice Roll worksheet from Layla's second session shows the X-plot, box plot and ten statistics. The tutor made the first few X's in the X-plot then Layla completed it. The tutor made the box of the boxplot and Layla made the whiskers.

Block Weights

In her third session, Carla chose to weigh 15 blocks and record the weights to a tenth of a gram on the data collection sheet (Figure 4.10). In the previous lesson, she learned to zero the scale and she learned the symbol for grams (g). The tutor guided her in rounding the weights to the nearest gram and he wrote the rounded weights in a third column as she dictated to him. The tutor attempts to explain variation.

Tutor: What do you want to do now?

Carla: The weighing.

Tutor: Ok. Do you know what I did?

Carla: Ah?

Tutor: I got. Uh. Well, I've got two things here. I've got. I made a data collection sheet. I also have the data you collected from last time where you weighed the blocks. [19:02]

Carla: And Mr. Big Fat Gummy Bear. [Carla had an eraser in the shape of a gummy bear. She weighed the "Big Fat Gummy Bear" in the previous session.]

Tutor: And the big fat gummy bear. Yes. How much did the big fat gummy bear weigh?

Carla: Two. Twenty-five, point two pounds … grams [She corrects herself]

Tutor: Grams. Yeah. That's it. You got it. Um. These may not be the same blocks, they may be different, so the weights may be a little bit different. So. Um. Do you wanna weigh 'em again?

Carla: Yeah.

Tutor: Ok. We'll start over. Ok, do you remember how to turn on the scale?

Carla: Here? [She points to a button on the scale.]

Tutor: Well, try it. Ok. And …

CB: Weigh a block?

Tutor: Oh, wait a minute. It's upside down again. [20:00] Is it giving you ounces or grams?

Carla: Grams.

Tutor: Grams. How do you know?

Carla: By the "g" on top.

Tutor: Yeah. It's got a little "g" for grams. Ok. So, let's zero it. Make sure it's zeroed. And now, what we want to do: make …

Carla: A list.

Tutor: Column headings before we start taking the data down. So, we put … you could put block number here.

Carla: So, like one, two, three, four. [Carla starts numbering in the first row instead of putting the column headings while the tutor's attention is on the camera.]

Tutor: Oh. No wait. No, we want to label the columns for… Oh, that's Ok. You know what we could do? You can just write it above the column here. Do you want me to do that? [See Figure 4.10.]

Carla: Yeah.

Tutor: Ok. I'll put the column heading on here. Ok. "Block Number" And then here we'll put the weight. [21:00] And we're weighing in grams. So, we'll put the little "g" there. We put that in parentheses.

*   *   *   *

Carla: And weighs [she gets ready to write the first weight on the sheet.]

Tutor: Say the weight loud enough so that the microphone hears you. How much does it weigh?

Carla: 12.8

Tutor: Are you sure that's an 8?

Carla: Zero.

Tutor: Yeah. That's a zero. 12.0 what?

Carla: Grams

Tutor: Grams. Right.

Carla: The Last block we weighed [whispers] [23:00] 11.7 grams.

Tutor: Right.

Carla: Second block. Third. 10.9 grams; 11.5 grams; 10.9 grams [24:09]

Tutor: Oh. [The tutor looks at the data collection sheet.] Let me show you something here. That's really good that you're thinking of the grams. But see when I put it up here, that means that everything underneath it is in grams. So, you just put it once. That way you don't have to write it. You don't have to erase it. You can just leave 'em there. That's fine.

Carla: Ok

Tutor: You just don't have to put it on the next ones.

Carla: Did I already do that block?

Tutor: Uh. How much does it weigh?

Carla: 10.9 grams.

Tutor: Ok. So, it looks like.

Carla: 10.3 grams; [25:00] 11.5 grams; 11.1 grams; 10.9 grams; 11.2 grams; [26:00] 11.7 grams; 12.3 grams; 10.2 grams; 11.5 grams; [27:00] 9.5 grams.

Tutor: Is that all the blocks?

Carla: Yeah. [27:30]

Tutor: Do you want more blocks to weigh? Or is that enough for now?

Carla: That's enough for now.

Tutor: Ok. Let me show you something that we can do with that now. What we want to do is we want to look at the distribution. Um. And we'll talk more about that. But, you see how they're all different? All the weights are different?

Carla: Yeah.

Tutor: Maybe there are a couple that are the same. But that's… We call that variation. There's variation in the weights of the blocks.

Carla: What's variation?

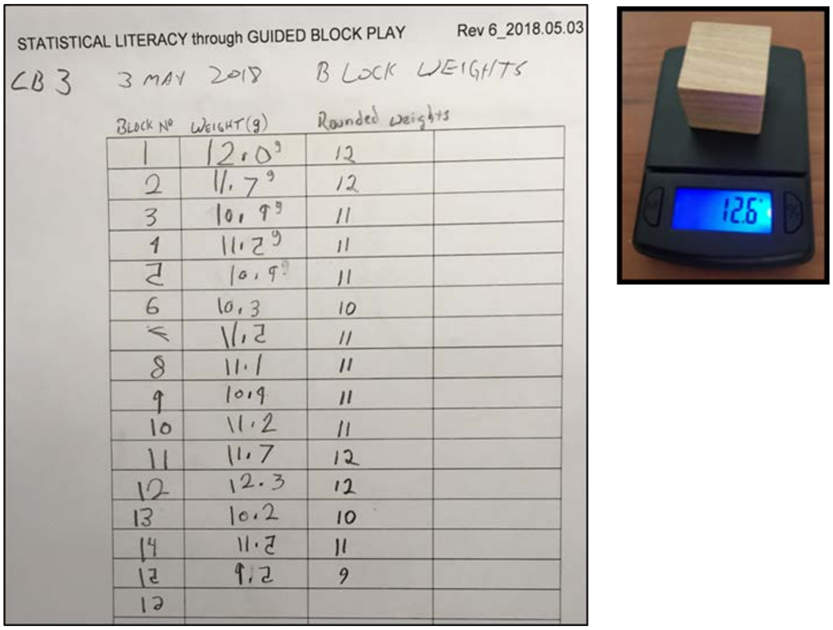
Tutor: Um. [28:00] It means they're not all the same. It means they're different. It's like when we have. Here we have 15 blocks, if we had no variation, they'd all weigh exactly the same. But since the weights vary, that means that they're different. But, what we want to know is if there's a pattern in how they change. Some are higher; some are lower. But we want to find out if there's a pattern to the variation. If we look at a whole. We call this a data set. Each weight of a block is a data point. So, we call it a data set. And we're gonna look at that all together to see if we can learn something about the weights of the blocks. I think it will become clearer when we go through it. You want to? [she looks at the data sheet.] [29:00] I know I'm not explaining it in a way that's very clear. But, if we walk through it together …

Carla: Yeah.

Tutor: I'll show you what I mean. [she starts to play with her pencil.] It's a lot easier to show you than to explain it. The first thing we want to do is round the numbers. Do you know how to round a number?

Carla: No

Tutor: No? I'll show you how to do that.



STATISTICAL LITERACY through GUIDED BLOCK PLAY   Rev 6_2018.05.03

CB 3   3 MAY 2018   BLOCK WEIGHTS

Block №	Weight (g)	Rounded weights	
1	12.0g	12	
2	11.7g	12	
3	10.9g	11	
4	11.2g	11	
5	10.9g	11	
6	10.3	10	
7	11.2	11	
8	11.1	11	
9	10.9	11	
10	11.2	11	
11	11.7	12	
12	12.3	12	
13	10.2	10	
14	11.2	11	
15	9.2	9	
16			

Figure 4.10. Carla weighed 15 blocks with a small digital scale and wrote the block weights in the data collection sheet. She then rounded the weights to the nearest gram and the tutor wrote the rounded weights as Carla called them out. The tutor also wrote the column headings.

In a similar exercise, Layla weighed 15 blocks, rounded the weights to the nearest gram, made an X-plot of the rounded block weights, arranged the blocks accordingly on the ruler, then balanced the ruler to find the average as shown in Figure 4.11.

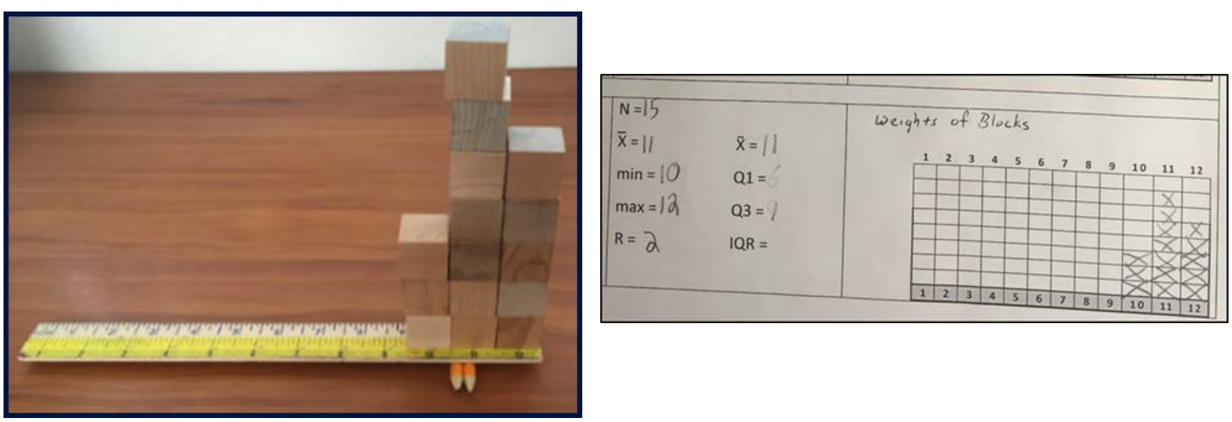


Figure 4.11. Layla made an X-plot and arranged the blocks on the ruler according to the rounded weights and balanced the ruler to find the average. She found the min, max, range, and median.

Marble Weights

In her fifth session, when Fiona had a choice of doing the Marble Grab or weighing the marbles, she chose the marbles. The tutor demonstrated zeroing the scale, explained decimals and grams as a unit of measure. The tutor made a data collection sheet including title and column headings. Fiona weighed ten marbles to a tenth of a gram and wrote them in the data collection sheet as shown in Figure 4.12. The tutor was going to remove the boulder but when Fiona saw it, she was immediately attracted to it. This turned out to be fortuitous as it introduced her to outliers. She recorded the marble weights to a tenth of a gram in the data collection sheet.

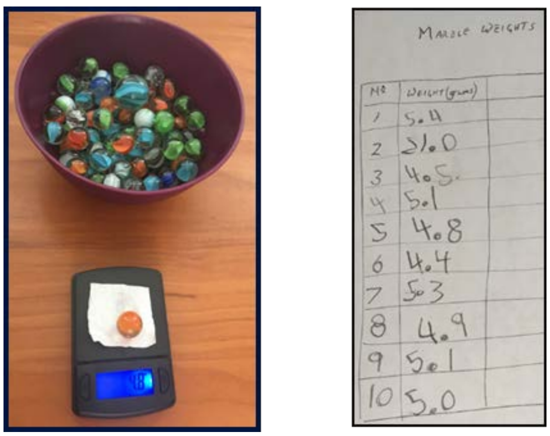


Figure 4.12. Fiona weighed ten marbles from the marble bowl, including the extra-large marble (commonly known as a “boulder”).

Fiona made the Marble Plot shown in Figure 4.13 and found the minimum, maximum, N, range, mode, and median. The tutor brought her attention to the shape of the distribution, including symmetry, gaps, and outliers. She decided to label the outlier. The tutor drew the axis and numbered the scale from 4.0 to 6.0 grams. The tutor wrote the numbers 1 to 12 underneath the axis of the marble plot and showed Fiona how to transform the marble weights to block positions on the ruler. She placed blocks on the ruler in the corresponding location and found the balance point as shown in Figure 4.14. She transformed the balance point back to the scale of the marble weights and wrote the average on the worksheet. The tutor suggested she not find Q1, Q3, and IQR since she had already been in the session for over 35 minutes and might be tired, but she decided to continue. Then she made a boxplot.

When asked at the end of the session what she liked most about it, she said she liked making the boxplot because it was challenging. She liked weighing the marbles and making the marble plot. She said it was hard for her to figure out what the tutor was drawing when he drew a break in the *x*-axis to accommodate the outlier. She was fascinated by the outlier.

In session 6, Fiona reviewed the previous lesson with the tutor and completed the marble plot. She transformed X-bar, Q1, Q3, and IQR back to grams and wrote the converted values on the Marble Plot. She added a title and labeled the boxplot as shown in Figure 4.13.

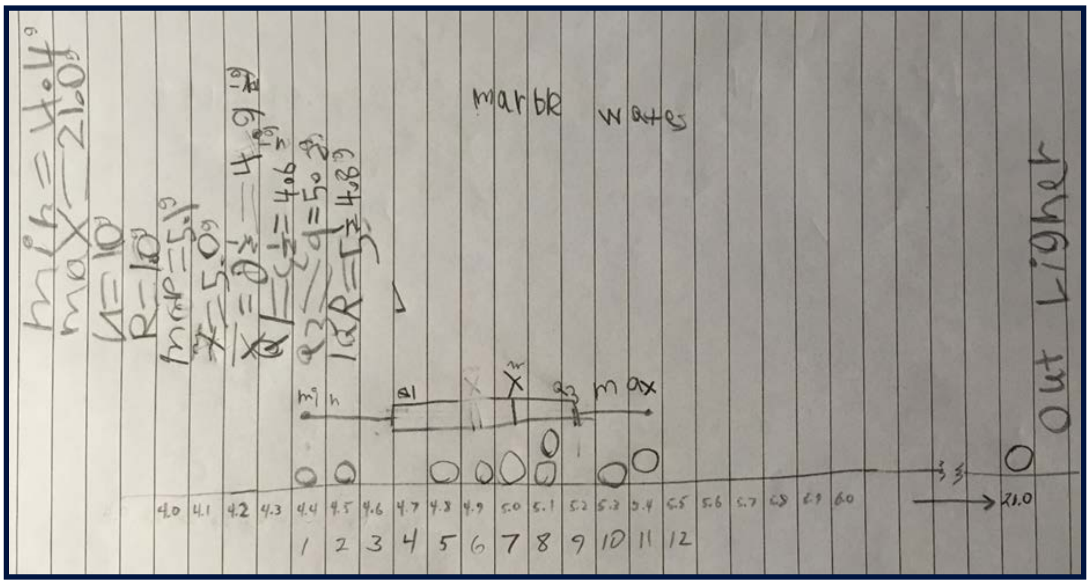


Figure 4.13. The tutor drew the first circle on the Marble Plot at 4.4 grams and Fiona drew the rest. She made a box plot and labeled the five points. She was intrigued by the outlier and the break in the axis. She decided to label the "Out Ligher."

Figure 4.14. Fiona arranged the blocks on the ruler according to the transformed scale, found the balance point, and recorded the average of the dataset.

Marble Grab

In the Marble Grab, the tutor posed the question: “Which hand is bigger, your left or your right?” In this activity, the student grabs a handful of marbles and counts them, records the number in the data collection sheet then repeats with the opposite hand. This continues for several trials. In the following example, Edgar grabs a handful of marbles, counts them, and records the number of marbles in the data collection sheet as shown in Figure 4.15. Then he switches hands and repeats for a total of five trials. They discussed the variation in marble counts between hands and from trial to trial on the same hand.

Tutor: [36:00] Do you want to play a statistics game with marbles or blocks, and a scale?

Edgar: I want to play sta-tis-tics … statistics with marbles.

Tutor: Ok. Let’s play statistics with marbles. There are two kinds of statistics. There are statistics we get from counting, [Edgar yawns] and statistics we get from measuring.

Edgar: Counting. [Edgar looks up at the Tutor.]

Tutor: You want counting? Ok. [The tutor pours the marbles into a bowl. Edgar’s face lights up and he leans forward in his chair.]

Edgar: You have a bowl of marbles. [He shows keen interest.]

Tutor: We have a bowl of marbles. [The tutor takes the big marble out of the bowl.]

Edgar: Why did you took out the big one? [Edgar gets up off his chair and walks around the table.]

Tutor: Because we're gonna measure your handful. You wanna keep the big one in there, don't you? [37:00]. [The big marble draws Edgar's attention as it did Fiona's.]

* * * * * * * *

Edgar: [39:00] Well I thought you had another big marble.

Tutor: You thought I had what?

Edgar: I thought you had another big marble. [Tutor continues drawing the table for data collection while Edgar noisily plunges his hands into the marble bowl.]

Tutor: No, I just have the one big one. [Tutor finishes drawing the lines] So, here we go. Our data collection sheet, what we're gonna do is … Trial. We're gonna start with Trial 1. And then we go to Trial 2. So, for each of these trials, you're gonna grab a handful with your left hand, and you can dump them in here [the plastic container], and then we can count them. And then we'll put 'em back. And then you can grab a handful with your right hand. Then compare the two. [Edgar continues to noisily play with the marbles.] That's why I took the big on out because it kind of messed it up. [39:50] (Transcript EM4).

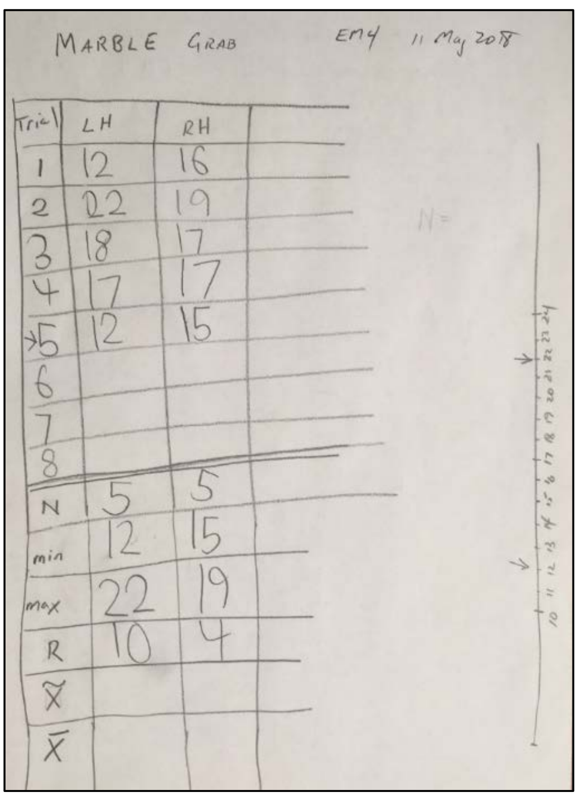


Figure 4.15. This was Edgar's fourth session. He compared two datasets from the Marble Grab. He found min, max, and range. The session ended before he could find the rest of the statistics.

The Marble Grab was a favorite activity among all the students. It provided a rich sensorimotor experience of variation both within and between datasets thus providing an embodied foundation for the *t*-test and analysis of variance (ANOVA). These topics usually encountered in university level courses are found to be difficult by many adult students. The boxplot is an ideal tool for comparing data sets. However, after conducting several Marble Grabs with different students, it seemed to the researcher that introducing the activity with just one hand might help to establish the procedure, then when the student has some procedural familiarity after say five or six rounds, extend the activity to two hands. Consequently, during the fifth and last session with Edgar, he did the one-handed Marble Grab, made a Marble Plot, found the ten statistics, and made a boxplot as shown in Figure 4.16. This activity seemed to integrate the previous lessons.

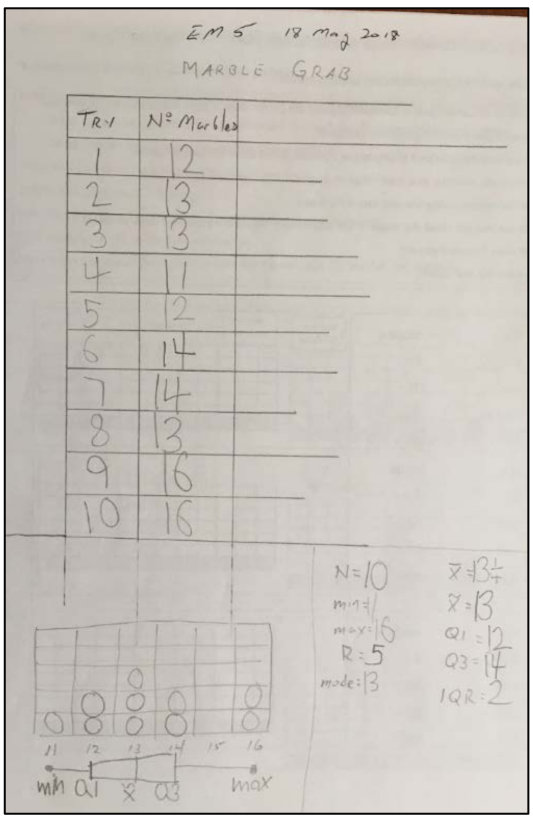


Figure 4.16. Edgar performed ten trials of the Marble Grab with one hand, recorded the results, organized the dataset into a marble plot, found ten statistics, and made and labeled a boxplot. The tutor made the data collection sheet and wrote the symbols for the statistics and numbered the axis of the marble plot.

“Doing the marbles” was a favorite activity for all the students as captured in the following conversation between the tutor and Fiona, a kindergarten student.

Tutor: Let me just ask you a couple of things before you go. Tell me what you liked best about what we did today.

Fiona: I liked [pause] doing the marbles. [52:02]

Tutor: You liked doing marbles?

Fiona: Yes.

Tutor: What was your favorite part?

Fiona: It was doing the marbles.

Tutor: Grabbing the marbles? Or counting the marbles?

Fiona: Yup, just the whole thing of the marbles.

Tutor: The whole thing of the marbles?

Fiona: Doing the marbles; that’s what I said. I said doing the marbles.

Tutor: Doing the marbles; now we know what “doing the marbles” is. Do you want to do marbles next week too?

Fiona: Yes.

Tutor: Ok, well before we finish…

FB Doing the marbles will mean the whole thing of grabbing the marbles and counting them.

Tutor: Ok, now we know what “doing the marbles” is.

Fiona: And showing you my ideas.

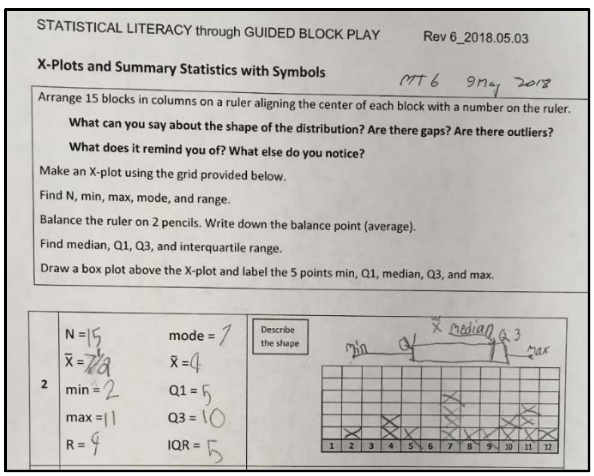
Tutor: You had some good ideas today.

Fiona: Yes. (Transcript FB3).

Student Pairs

Case 2.2 was with two second-grade students, a boy and a girl. The girl was compliant, and the boy was restless. This was the first session and having two students was unplanned. They both showed up due to a misunderstanding around scheduling and rather than turn one away, the tutor decided to accommodate them both. However, the tutor was not prepared to manage the dynamics of this pair, especially in a first session when there were still many unknowns. Case 2.2 was with two second-grade girls. This was the third session for one of them and the second session for the other. Although they were on task, there was little synergy between them. The first two cases with pairs of students did not yield much useful data.

In Case 2.3, a second grader, after having completed five sessions, tutored her kindergarten friend in the Dice Roll and the Marble Grab. This was an opportunity to assess the learning of the second grader and observe the response of the kindergartener. Vygotsky described the ZPD to include either an adult or a more advanced peer. This was an opportunity to observe a more advanced peer conducting a younger learner through the ZPD. Figure 4.17 shows the worksheet for the Dice Roll. This worksheet reflects a level of competence equal to any in the one-on-one sessions. The second-grade tutor was unsure about the symbol for the average (X-bar) and needed some guidance in making the boxplot. She had only made three boxplots previous to this session. She conducted the session competently demonstrating her procedural knowledge. One weakness of the study was the absence of assessment tools for evaluating concept development (the first research question). This opportunity to pair a more advanced peer with a new learner arose fortuitously and it suggests that placing a student in the role of tutor might be a useful strategy for assessment as well as a possible approach to scaling up the lessons to the classroom.


STATISTICAL LITERACY through GUIDED BLOCK PLAY   Rev 6_2018.05.03

X-Plots and Summary Statistics with Symbols   MT 6 9 May 2018

Arrange 15 blocks in columns on a ruler aligning the center of each block with a number on the ruler.

What can you say about the shape of the distribution? Are there gaps? Are there outliers?

What does it remind you of? What else do you notice?

Make an X-plot using the grid provided below.

Find N, min, max, mode, and range.

Balance the ruler on 2 pencils. Write down the balance point (average).

Find median, Q1, Q3, and interquartile range.

Draw a box plot above the X-plot and label the 5 points min, Q1, median, Q3, and max.

2	N = 15	mode = 7	Describe the shape
	$\bar{X}$ = 7/a	$\tilde{X}$ = 4	
	min = 2	Q1 = 5	
	max = 11	Q3 = 10	
	R = 9	IQR = 5	

Figure 4.17. In this Dice Roll exercise, a second-grade student guided her kindergarten friend in completing the X-plot, the box plot, and finding the statistics.

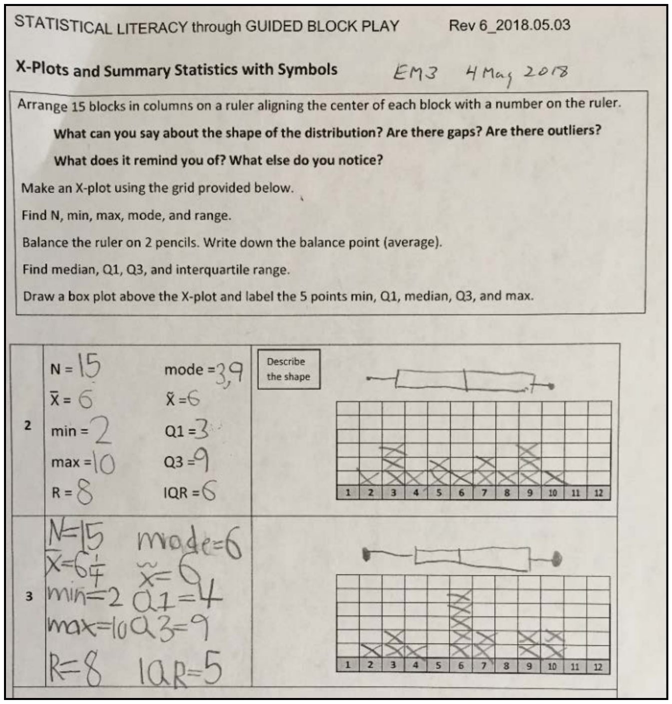
Observations of Peers in Interaction Analysis

Both interaction analysis groups saw that students were fatigued by finding quartiles and making boxplots. Whether this fatigue was from the activity itself or from the length of the lesson is not clear. Both groups concluded that the student they observed found no meaning in the task. Each group however, observed only one of many such activities finding quartiles and making boxplots. Subsequently, the researcher reviewed these interactions and saw that the guidance/autonomy dynamic was out of balance and he had become too intent on reaching a learning goal. He accordingly modified the design and sequence of the lessons.

Session 1

Three elementary Montessori teachers (2 with over 20 years' experience) and one PhD clinical psychologist observed 15 minutes of video in this 50-minute session. They were given a copy of the completed worksheet (Figure 4.18) and the transcript (Case 1.3, Session 3) for reference during the Interaction Analysis session. Participants observed a Dice Roll lesson of a 7.3-year-old first-grade boy. This was his third session. He created a symmetric Free Form block design then made the X-plot and found the statistics shown in Figure 4.18. On his own initiative, he copied the statistical symbols from the first lesson into the corresponding area for the Dice Roll lesson. He rolled the dice and made the X-plot, then found the ten statistics and made a boxplot.

Observer Comments. These observers commented that Edgar was not learning independently; the lesson did not match his developmental level of reasoning. They noted that although he could find statistics under the direction of an adult, they had no meaning for him. He lacked prior knowledge and had no point of reference or context. He had no prior knowledge of fractions. They questioned the value of this lesson at this stage of development. They agreed that it might have value in fourth or fifth grade with students working in small groups. One observer commented that the lesson is not sufficiently user-friendly to engage a younger child naturally, spontaneously, and autonomously so that he develops mastery. He needs to be two or three years older before his reasoning mind matches the level of the lesson. He can see the shape, the mode and the range, but median, mean, Q1, Q3, and boxplots are too abstract. These observers agreed that this lesson might work well with children as young as third grade. In their experience, fifth grade is a good year for conducting surveys and for introducing technical terms; sixth grade for box and whisker plots. This group felt that the lesson goes too quickly to a level of abstraction beyond the reach of the first or second grader and that the level of complexity and abstraction should increase more gradually. Drawing the blocks rather than doing an X-plot is more concrete and might have more meaning for younger learners. One observer commented that it would be interesting to do a longitudinal study to see if there is an effect on attitudes toward statistics later in life.



STATISTICAL LITERACY through GUIDED BLOCK PLAY Rev 6_2018.05.03

X-Plots and Summary Statistics with Symbols EM3 4 May 2018

Arrange 15 blocks in columns on a ruler aligning the center of each block with a number on the ruler.

What can you say about the shape of the distribution? Are there gaps? Are there outliers?

What does it remind you of? What else do you notice?

Make an X-plot using the grid provided below.

Find N, min, max, mode, and range.

Balance the ruler on 2 pencils. Write down the balance point (average).

Find median, Q1, Q3, and interquartile range.

Draw a box plot above the X-plot and label the 5 points min, Q1, median, Q3, and max.

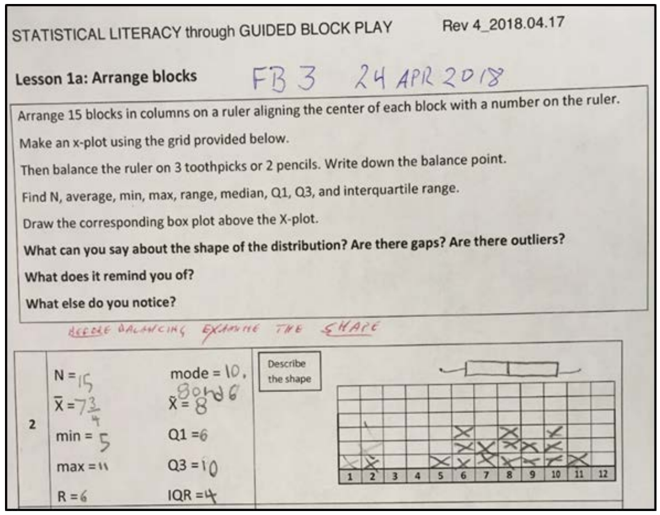
2	N = 15 $\bar{X}$ = 6 min = 2 max = 10 R = 8	mode = 3, 9 $\tilde{X}$ = 6 Q1 = 3 Q3 = 9 IQR = 6	Describe the shape
3	N=15 $\bar{X}$=6¼ min=2 max=10 R=8	mode=6 $\tilde{x}$=6 Q1=4 Q3=9 IQR=5	

Figure 4.18. In this 3rd session for a first-grade boy, he made a Free Form symmetric arrangement of blocks, made an X-plot, found the statistics, and made a boxplot. On his own initiative in the second lesson, he copied the symbols before doing the Dice Roll.

Session 2

Two retired elementary school teachers (over 30 years' experience each) and one retired early childhood specialist and trainer with over 30 years' experience observed 24 minutes of video in this 90-minute session. They were given a copy of the completed worksheet (Figure 4.19) and the transcript (Case 1.3, Session 3) for reference during the Interaction Analysis session.

Observers watched a Dice Roll lesson of a 6.3-year-old kindergarten girl. This was her third session. She created a symmetric Free Form block design then made the X-plot and found the statistics shown in Figure 4.19. The balance point was 7¾. This was a new fraction for her (she learned ½ in her first session).


STATISTICAL LITERACY through GUIDED BLOCK PLAY Rev 4_2018.04.17

Lesson 1a: Arrange blocks FB 3 24 APR 2018

Arrange 15 blocks in columns on a ruler aligning the center of each block with a number on the ruler.
Make an x-plot using the grid provided below.
Then balance the ruler on 3 toothpicks or 2 pencils. Write down the balance point.
Find N, average, min, max, range, median, Q1, Q3, and interquartile range.
Draw the corresponding box plot above the X-plot.
What can you say about the shape of the distribution? Are there gaps? Are there outliers?
What does it remind you of?
What else do you notice?

BEFORE BALANCING EXAMINE THE SHAPE

2	N = 15	mode = 10, 8 and 6	Describe the shape	
	$\bar{X}$ = 7 3/4	$\tilde{X}$ = 8		
	min = 5	Q1 = 6		
	max = 11	Q3 = 10		
	R = 6	IQR = 4		

Figure 4.19. A kindergarten girl arranged the blocks on the ruler, made the X-plot, and found the statistics. The tutor made the box of the boxplot and the student made the whiskers.

Similar concerns were raised about what is age appropriate, especially with respect to finding quartiles and making boxplots. They observed that she was getting fatigued after about 32 minutes into the session. They were impressed by her ability to stay engaged so long. Observers drew an equivalence between work and play. There was a suggestion to break the lesson down further into “baby steps” and include more repetition. There was a suggestion to limit the session to the physical interactions and leave the worksheet until second or third grade. P1: “just because they can do it doesn't mean they should be doing it or that is appropriate. But once they reach that stage where they can go from the concrete to the symbolic, that doesn't happen until the second grade or third grade.” It’s OK to engage the student with the concepts of variation and distribution but without difficult vocabulary and leaving the quartiles and boxplots for later years. There was support for the idea of doing this with adults.

A Redesigned Learning Trajectory Based on the Findings

The findings of the Interaction Analysis suggest quartiles and box plots be postponed. Extant research suggests that statistical literacy education begin with inquiry. Considering these, in the following redesigned learning module, students develop more familiarity with measurement, data collection, data representations, and embodied interaction with learning tools before making box plots and comparing data sets.

Module 1 is designed to introduce basic dispositions, knowledge and skills of data literacy and statistical inquiry through tutorial guidance. These lessons are intended to be playful and enjoyable. In these lessons, students will learn to use and care for measuring instruments. They will exercise hand-eye coordination through manipulation of materials. They will practice data organization skills using data collection sheets and visual representations of data sets. These lessons are designed to lay a foundation of conceptual understanding of a data set as an aggregate with shape, spread, and center. Subsequent Modules to be developed will include the use of additional measuring instruments, histograms, time-series data, and confidence intervals. They will similarly be non-mathematical in conveying statistical concepts but will offer opportunities to introduce mathematical ideas in an experiential way.

COURSE: Developing Knowledge and Skills of Statistical Inquiry

Module 1: Shape, Spread, and Center of a Data Set

Essential Understandings

- A data set has a shape that we can describe using ideas like gaps, clumps, outliers, and mode
- A data set can be spread out or all close together. The range, IQR, and MAD tell us how spread out the data are.
- A data set has a middle that we can find in different ways like balancing a ruler or finding the block in the middle.
- A data set can be represented by an arrangement of blocks on a ruler
- Data are generated by counting and by measuring;
- Statistics describe a data set
- We can make visual representations of a data set, some of them show all the data and some only a few key statistics like a box plot.
- We can compare data sets to help answer questions
- A ruler can help us understand numbers
- Labeling is important
- Statistics can be fun
- We can transform numbers back and forth between two different scales
- We weigh small things in grams

Lesson 1: Dice Roll – gather and organize data

1. Organize blocks on a ruler according to repeated dice rolls.
2. Describe the shape of the data set (mode, symmetry, resemblances, gaps, outliers).
3. Make an X-plot (a scaffolded dot plot).

4. Find min, max, mode through observation.
5. Find N and range by counting.
6. Find the balance point to the nearest ¼ inch [learn fractions: ½, ¼, ¾ as needed].

Lesson 2: Block Weights – use measuring instruments to build a data set
1. Prepare a data collection sheet (include title, date, and column headings).
2. Zero the scale and discuss proper care of measuring instruments.
3. Weigh at least 10 blocks and record the reading to one decimal place.
4. Round measurements to the nearest whole number in the next column.
5. Arrange blocks on the ruler according to the rounded weights.
6. Repeat Lesson 1.

Lesson 3: Marble Weights – use measuring instruments to build a data set
7. Prepare a data collection sheet (include title, date, and column headings).
8. Zero the scale and discuss proper care of measuring instruments.
9. Weigh at least 10 marbles and record the reading to one decimal place.
10. Transcribe the measurements onto the ruler.
11. Arrange blocks on the ruler according to the transcribed numbers.
12. Repeat Lesson 1.

Lesson 4: One-Handed Marble Grab – count objects to build a data set
1. Prepare a marble bowl
2. Prepare a data collection sheet (include title, date, and column headings)
3. Grab a handful of marbles and count the marbles
4. Record marble counts in the appropriate column.
5. Repeat Lesson 1.
6. Manipulate the blocks to find the middle block (median), Q1, and Q3
7. Find IQR by counting
8. Make a box plot.

Lesson 5: Two-Handed Marble Grab – compare two data sets
1. Prepare a data collection sheet with separate columns for left hand and right hand.
2. Consider the question: “Is one hand bigger than the other?”
3. Alternate the marble grab between the two hands.
4. Perform at least 5 trials.
5. Manipulate the blocks to find the middle block (median), Q1, and Q3
6. Find IQR by counting
7. Make a box plot.
8. Compare the two hands.
9. Discuss the observed variation.

Lesson 6: Find the mean absolute deviation (MAD) of a data set
10. Choose a data set from a previous exercise.
11. Make a data collection sheet of 2 columns.
12. List the position of each block in the first column.

13. List the distance of each block from the balance point in the second column.
14. Rearrange the blocks on the ruler according to the 2^{nd} column.
15. Balance the blocks and record the balance point (MAD).

Limitations of the Study

Participants in this study were not representative of the population of students. There was no statistical sampling. Although the researcher took steps to reduce bias in the analysis and interpretation of the data through interaction analysis, data triangulation, and positionality (explicating the philosophy and motivation of the researcher), some researcher bias remains. Although the interaction analysis groups provided valuable feedback to compensate for researcher bias, the process was not as rigorous as originally envisioned due to limits on the availability of participants. These participants were not available to review the background and design of the study before reviewing videos of the learning sessions. Their feedback was based on their extensive experience in primary education. The first interaction analysis session had technical difficulties reducing the available scheduled time from 60 to 45 minutes. Coding of transcripts, field notes, and video in the pilot study followed a rigorous methodology (Saldaña, 2013), however, coding was not corroborated by a second analyst and hence was not subject to inter-rater reliability assessment.

Sessions were held in an empty K-5 music classroom. Musical instruments in the room were sometimes a distraction. The noise of the air conditioner sometimes interfered with audio data collection. The teacher in whose classroom the study took place worked quietly in the background but sometimes her activity was distracting. The noise of children playing outside the classroom sometimes interfered with the session. Some sessions were cut short when a parent/guardian came in to pick up the student. Some data was lost due to technical failures, including automatic shutoff of the camera, battery failure, and filling up the memory card before the end of the session. Sometimes relevant activity occurred outside the field of view of the camera.

Lessons Learned

The site of future studies needs to be free of distractions and ambient noise. The camera needs to be more reliable and set where relevant activity is within the field of view. A second camera would help ensure more complete coverage and provide backup in case of technical failure. Automatic transcription software used toward the end of the data reduction period reduced the time needed for transcription. In future studies, preliminary data reduction after each session would help inform the following session. Delays in necessary approvals forced a compression of the schedule for the field work and pushed it out toward the end of the semester reducing flexibility in scheduling. Acquiring all the video data before beginning analysis limited the ability of the study to answer the research questions. Although this was an exploratory study and a certain amount of improvisation was expected, a more systematic and sequential approach to the lessons might enrich the data and provide additional insight into learning trajectories, similarities and differences among students. Although procedural understanding was demonstrated, conceptual understanding was not fairly assessed. Future studies should probe more deliberately the thinking of the students and their understanding of variation and distribution.

When the potential for the marbles to engage the students was discovered, more time was needed to integrate the activities with the marbles into the study design and re-formulate the research questions. Maintaining the balance between tutorial guidance and learner autonomy (agency) was an ongoing challenge. The tutor sometimes lapsed into verbal explanations and missed important cues from students. Careful review of the video after each session and ongoing tutor reflection on the learning dynamics in relation to the research questions might help to maintain this balance.

Asking students to make predictions about where the ruler will balance before they balance it might help develop their statistical intuition and reveal more about their thinking process. The researcher had no formal training or experience teaching children. Future studies would benefit from close collaboration with early childhood educators at all stages of the design and execution of the study.

Chapter 5: Discussion

Ideas of “readiness” and “developmentally appropriate” may be interfering with possibilities of learning in early childhood (National Research Council (US), 2001). Bruner (1960) recognized that schools may be postponing the teaching of some subjects on the assumption that the subject matter is too difficult. He claimed that “the foundations of any subject may be taught to anybody at any age in some form” (p. 12). He further claimed, “any idea can be represented honestly and usefully in the thought forms of children of school age, and that these first representations can later be made more powerful and precise the more easily by virtue of this early learning” (p. 33). Bruner concluded that it may be possible to discover methods to teach basic ideas of science and mathematics to children “considerably younger than the traditional age. It is at this earlier age that systematic instruction can lay a groundwork in the fundamentals that can be used later and with great profit at the secondary level” (pp. 44-45).

Statistics is an intimidating subject for many adult learners. By incorporating an element of play and using blocks and marbles as a primary mediating artifacts, this study showed that young children are capable of far more than adults generally recognize. Foundational statistical concepts can be *embodied* in a way that children find engaging and no computation is needed to find common measures of dispersion and central tendency. “It may indeed be the case that such an early science and mathematics 'pre-curriculum’ might go a long way toward building up in the child the kind of intuitive and more inductive understanding that could be given embodiment later in formal courses in mathematics and science” (Bruner, 1960, p. 46). This study showed that a pre-curriculum in embodied learning of statistics is accessible and engaging for K-2 students. Although this study did not reveal the degree of conceptual understanding achieved by the participating students, it did suggest a way of introducing statistics education to young children. Refining these activities and scaling them up to a classroom calls for collaborative research with practicing teachers. Beginning with experimental classrooms then introducing these activities to more typical classrooms: “We must operate always under the constraint that an effective intervention should be able to migrate from our experimental classroom to average classrooms operated by and for average students and teachers, supported by realistic technological and personal support" (Brown, 1992, p. 143). The challenge then is how to educate teachers. This speaks to *how these findings might inform the design of learning experiences for adults and instigate innovations in statistics education for all ages* (the third research question).

Statistics Education of Teachers

The *GAISE Report* (Franklin et al., 2005) identifies three levels of development that roughly correspond to elementary, middle, and high school. However, the report emphasizes that the levels are based on development in statistical thinking, rather than age. Teachers need to understand statistics as distinct from mathematics, particularly the focus on context, distribution, and variability. Statistical reasoning and mathematical reasoning are different ways of thinking. Teachers need conceptual, not just procedural understanding. The report emphasizes the need "to provide teachers with courses and professional development that cultivate their statistical understanding, as well as the pedagogical knowledge to develop statistical literacy in the next generation of learners" (p. 5). Traditional courses in statistics are unlikely to accomplish this. The *Statistical Education of Teachers* (*SET*) report (Franklin et al., 2015) recommends that pre-service teachers "learn statistics in ways that enable them to develop a deep conceptual understanding of the statistics they will teach" and "engage in the statistical problem-solving process – formulate statistical questions, collect data, analyze data, and interpret results – regularly in their courses" (p. 8). Many practicing teachers did not learn statistics in their teacher preparation program; this may be an advantage given the pervasive failure of traditional courses to develop conceptual understanding and their tendency to breed misconceptions. "Robust professional development opportunities need to be developed for advancing in-service teachers' understanding of statistics" (p. 8). The *SET* report advocates "using manipulatives to aide in the collection, exploration and analysis, and interpretation of data" (p.22). This study takes this a step further in using manipulatives to "play" with statistical concepts. The lessons explored in this study can be adapted to teach teachers how to teach their students while learning statistics themselves, perhaps for the first time. However, innovations in professional development that prepares elementary teachers to teach statistics are already in the field.

Professional Development

EarlyStatistics is an online professional development project designed to address the issues identified in the *SET* report targeting European elementary and middle school teachers (Meletiou-Mavrotheris, Mavrotheris, & Paparistodemou, 2011; Serradó Bayés, Meletiou-Mavrotheris, & Paparistodemou, 2014). This project is described as *support-led* rather than *package-led*. An online community of practice was established through computer-supported collaborative learning after a week-long face-to-face session. Participants "take part in authentic educational activities which give them the opportunity to reflect on the 'big ideas' of statistics and their applications, and to explore ways of improving statistics instruction through the adoption of a coherent technology-rich curriculum based on the statistical problem solving process" (p. 7). This is a promising program with a track record that merits further development and replication.

Another promising approach to developing the professional capabilities of teachers to cultivate statistical literacy is the ECHO for Education project (Giebitz & Stanton, 2018). ECHO for Education is an extension of the ECHO Institute conceived and developed as a method of providing medical training and care to remote and underserved communities in New Mexico (Arora et al., 2011). "Launched in 2003, the ECHO model™ makes specialized medical knowledge accessible wherever it is needed to save and improve people's lives." ECHO has grown into a global network providing medical services from over 220 hubs in over 30 countries on six continents. Patterned after this successful model for delivering medical services, *ECHO for Education* is a "hub and spoke" model of networked learning resources connecting experts at the hub with teachers around the state (the spokes). Equally important, it connects the teachers with each other for collaborative learning and problem solving, and with a bank of online resources. In 2018, A pilot program began with a 2-day face-to-face event followed by weekly one-hour videoconferencing sessions. Experts in topics relevant to the participants presented from the hub, then participants presented their own case studies and received feedback and recommendations from their peers located at schools around the state. They met again face-to-face at the end of the program in a statewide conference to share their experiences with a wider audience and to participate in the program evaluation. The ECHO model shows promise for developing an infrastructure of just-in-time, relevant, targeted training for professional development of educators. A proposal for funding a pilot project to prepare elementary school teachers to teach statistics is in progress.

Design Based Research

Design-Based Research (DBR) is another strategy that shows promise for cultivating statistical literacy throughout a community of practicing teachers (Bakker & Eerde, 2014; Design-Based Research Collective, 2003). DBR has roots in the design experiments of Ann Brown (1992). Brown describes her “attempts to engage in design experiments intended to transform classrooms from academic work factories to learning environments that encourage reflective practice among students, teachers, and researchers” and “the need for new and complex methodologies to capture the systemic nature of learning, teaching, and assessment” (p. 174). Brown sought to develop “a theoretical model of learning and instruction rooted in a firm empirical base” and “engineer interventions that not only work by recognizable standards but are also based on theoretical descriptions that delineate why they work, and thus render them reliable and repeatable” (p. 143). Brown describes “the intentional learning classroom” in which “students are encouraged to engage in self-reflective learning and critical inquiry” (p. 149). The learning track follows a spiral curriculum in which a few salient themes recur, “themes that students come to understand deeply and recognize at increasingly deeper levels of explanatory coherence and theoretical generality” (p. 150). The trajectory from learning statistical structures in block play to a level of competence in statistical inquiry remains to be explored.

DBR is an iterative process of investigation that seeks to develop theory while improving learning outcomes (Bakker & Eerde, 2014; P. Cobb, Confrey, DiSessa, Lehrer, & Schauble, 2003; Collins, Joseph, Bielaczyc, Collins, & Bielaczyc, 2016). It is particularly promising where variables cannot be easily isolated or controlled and where relevant variables may not even be known. DBR methodology combines collaboration in a teaching/learning context with refinement of design and theory through successive iterations. It can adjust to emergent phenomena, discoveries, and insights. Conjectures are generated and tested through successive iterations and the research design is modified according to emergent needs. The Design-Based Research Collective (2003) proposed five key features:

1) overlapping goals of designing learning environments and developing “proto-theories” of learning; 2) meeting these goals through “continuous cycles of design, enactment, analysis, and redesign”; 3) research outcomes that are useful to practitioners; 4) feasible in authentic settings and advance our understanding; and 5) methods connect and document enacted processes and outcomes (p. 5). DBR can be a powerful element in the professional development of teachers while strengthening local resources for research into what works in specific contexts with specific populations. This could be a framework for launching a block play approach and developing the research infrastructure to build toward a comprehensive, adaptable, culturally responsive curriculum in statistics.

A Lesson Analysis Framework

The *Lesson Analysis Framework* of Santagata, Yeh, & Mercado (2018) integrates theory, research-based knowledge, and mathematics teaching practice through systematic reflection and analysis. The core practices of this framework are: “eliciting and responding to student ideas, designing and sequencing instructional episodes that build conceptual understanding as the basis for procedural fluency, using multiple mathematical representations to support students’ development of conceptual understanding, and orchestrating classroom discussion” (p. 481). This framework follows four steps: (1) specify learning goals; (2) conduct empirical observations to collect evidence of student learning; (3) generate hypotheses about interactions that promoted student learning; and (4) apply implications of these hypotheses to improve the design of the learning session (p. 483). This approach might provide additional specificity to a DBR strategy.

Concept Formation and Conceptual Change in Statistical Literacy

The question of conceptual understanding of statistics is at the heart of this study. However, there is a kaleidoscope of varied opinions and perspectives on concept formation and change (diSessa, 2014). “There are no widely accepted, well-articulated, and tested theories of conceptual change” (p. 89). Such ambiguity is reflected at both the macro[xviii] and the micro level. At a micro level, neuronal group selection as a dynamic process in the formation of conceptual metaphors grounded in sensorimotor experience is an active area of research (Lakoff, 2015b, 2015a). Evidence suggests that rather than isolatable structures in the brain, concepts are better understood as a process of coordinated neuronal activity occurring within a whole situated organism (Feldman & Narayanan, 2004).

DiSessa proposes a knowledge-in-pieces perspective that is fine-grained enough to track learning so that it reveals individual differences without depending on a grand theory. He maintains that "good intuitive design can override the power of current theory to prescribe successful methods" (p. 102). Pre/post evaluations have little to say about what happens in the time between the pre and the post assessments. diSessa concludes, "almost no research on conceptual change tracks students" moment-by-moment thinking while learning … Filling in the big "before-and-after" views of change with the details of exactly what changes when may be the gold ring of conceptual change research" (p. 105). Cognitively Guided Instruction (Carpenter, Fennema, Franke, Levi, & Empson, 2000) offers tools to help teachers understand the development of mathematical thinking in young children. These tools might also help teachers understand children's statistical thinking. A framework developed for training teachers to understand the mathematical thinking of children (Ginsburg, 2018) might also help them understand their statistical thinking.

Thelen, Schöner, Scheier, and Smith (2001) developed "a formal dynamic theory and model based on cognitive embodiment" (p. 1) that integrates our understanding of the learning process throughout life. They explored the developmental origins of cognition in infancy. Their *dynamic systems* approach affirms the continuity of the learning process from the first months of life through adulthood. They studied the goal-directed actions of an infant interacting with a toy and conjectured:

> *If we can understand this particular infant task and its myriad contextual variations in terms of coupled dynamic processes, then the same kind of analysis can be applied to any task at any age. If we can show that 'knowing' cannot be separated from perceiving, acting, and remembering, then these processes are always linked. There is no time and no task when such dynamics cease and some other mode of processing kicks in. Body and world remain ceaselessly melded together (p. 2).*

The authors speculate about the usefulness of their model for "integrating multiple, time-based processes of human cognition and action" (p. 3). The division between the “conceptual” and the “perceptual-motor” is somewhat arbitrary and may be counterproductive. Thinking begins with perception and action and retains this embodiment throughout life. “The goal of development is not to rise above the mere sensorimotor but for cognition to be at home within the body” (Thelen, 2000, p. 8). “The important metric is not whether the mental activity is truly ‘conceptual’ or merely ‘sensorimotor’ but the flexibility of the coupling between thinking and acting” (p. 14). Skills are created *by* performance, not just reflected *in* performance; cognition is acquired from the outside in and depends on perceptual-motor pathways to access higher functions in the brain. This dynamic grounding of higher forms of cognition in sensorimotor pathways – much as they were formed in infancy – remains throughout life.

Lakoff and Johnson (1980) maintained that conceptual systems are built up of metaphors, and metaphors are grounded in embodied experience in the world. The primary metaphors that comprise the bedrock of statistical reasoning might be forged through sensorimotor experience in early childhood helping to ensure the adult is equipped for statistical thinking and reasoning. Building on the *neural theory of language*, Feldman and Narayanan (2004) proposed a *neural theory of metaphor* that accounts for the linking of the sensorimotor system through neural circuitry to higher cortical areas giving rise to metaphorical thought. Primary conceptual metaphors are “learned unconsciously and automatically in childhood simply by functioning in the everyday world with a human body and brain” (pp. 256-257). Neuronal maps are physically embodied in our body and brain through *neural recruitment* between clusters of neurons, or *nodes.* “This neural learning mechanism produces a stable, conventional system of primary metaphors that tend to remain in place indefinitely within the conceptual system and are independent of language” (p. 256). This underscores the importance of establishing foundational metaphors, such as a frequency distribution or a dataset as an aggregate with emergent properties of shape, spread, and center, early in the learning journey to avoid later conceptual confusion.

Statistics Education and Data Science: Two Cultures

There is a growing demand for statistical skills. However, two contrasting cultures are contending with the oceans of data that have come to characterize contemporary life. The statistics community has its roots in concerns of the State (hence the term *statistics*) such as demographics and economics. But the tremendous increase in data gathering and computational capabilities has given rise to data mining, data scraping, data science, and analytics – phenomena grounded in computer science, not statistics. The computer science community often sees statistics as a "bag of tricks" rather than as a way of thinking and reasoning about data. But without a firm grounding in statistics, the data scientists' enthusiasm for computational algorithms can become (or remain) unmoored from statistical literacy.

Breiman (2001) describes the challenge presented to the "data modeling culture" by the "algorithmic modeling culture:"

> There are two cultures in the use of statistical modeling to reach conclusions from data. One assumes that the data are generated by a given stochastic data model. The other uses algorithmic models and treats the data mechanism as unknown. The statistical community has been committed to the almost exclusive use of data models. This commitment has led to irrelevant theory, questionable conclusions, and has kept statisticians from working on a large range of interesting current problems. Algorithmic modeling, both in theory and practice, has developed rapidly in fields outside statistics. It can be used both on large complex data sets and as a more accurate and informative alternative to data modeling on smaller data sets. If our goal as a field is to use data to solve problems, then we need to move away from exclusive dependence on data models and adopt a more diverse set of tools. (Breiman, 2001, p. 199).

These two cultures need to find common ground. "Computer scientists bring useful skills and approaches to tackle the analysis of large, complex datasets. Statisticians bring important expertise in terms of the understanding of variability and bias to help ensure that conclusions are justified" (Wild et al., 2018, p. 31). Statistics education must change to ensure its relevance in a world of big data and complex algorithms.

Both cultures seek to reach conclusions from data, but their models of how to do so differ sharply. Statisticians assume data are generated by a stochastic model such as a normal or an exponential distribution. Computer scientists and data scientists, on the other hand, create algorithmic models concerned less with understanding the underlying mechanism than with predictive accuracy. One group thinks in terms of statistical inference, the other in terms of algorithms. “Data models are rarely used in this community [algorithmic community]. The approach is that nature produces data in a black box whose insides are complex, mysterious, and, at least, partly unknowable… the models that best emulate nature in terms of predictive accuracy are also the most complex and inscrutable” (Breiman, 2001, pp. 205- 209). Gould (2017) takes the position that data literacy *is* statistical literacy (see Chapter 1); the distinction is counterproductive. Gould calls on the statistics community to expand the scope of statistical literacy to embrace a new data landscape.

Conclusion

Learning is fueled by curiosity and questions. Inquiry is fundamental to being human. The use of data makes the inquiry process considerably more powerful for understanding how the world works. Statistical literacy is an invaluable constellation of knowledge, skills, and dispositions to support data-based inquiry and it can begin early in the life of a child. Students of all ages, citizens of all stripes, need the capabilities to frame questions and seek answers for themselves. By shifting the focus “from answering questions to asking them, inquiry emerges as a tool for harnessing not only the passion of students but also the stockpile of tacit knowledge that comes from a lifetime of experience doing the things that have become second nature to them” (Thomas & Seely Brown, 2011, p. 85). “Learning that is driven by passion and play is poised to significantly alter and extend our ability to think, innovate, and discover in ways that have not previously been possible” (p. 89).

References

Abrahamson, D., & Lindgren, R. (2014). Embodiment and embodied design. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (2nd ed., pp. 358–376). New York: Cambridge University Press.

Albert, L. R. (2014). The history of base-ten-blocks: Why and who made base-ten-blocks? *Mediterranean Journal of Social Sciences, 5*(9), 356–365. https://doi.org/10.5901/mjss.2014.v5n9p356

Anshen, R. N. (1986). *Biography of an Idea*. Mt. Kisko, New York: Moyer Bell.

Arnold, P., Confrey, J., Jones, R. S., Lee, H. S., & Pfannkuch, M. (2018). Statistics learning trajectories. In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International Handbook of Research in Statistics Education* (pp. 295–326). Cham, Switzerland: Springer.

Arora, S., Thornton, K., Murata, G., Deming, P., Pharm, D., Kalishman, S., … Qualls, C. (2011). Virus Infection by Primary Care Providers. *New England Journal of Medicine, 364*(23), 2199–2207.

Baker, M., & Penny, D. (2016). Is there a reproducibility crisis? *Nature*, *533*(7604), 452–454. https://doi.org/10.1038/533452A

Bakker, A., & Eerde, D. van. (2014). An introduction to design-based research with an example from statistics decucation. In A. Bikner-Ahsbahs, C. Knipping, & N. Presmeg (Eds.), *Approaches to qualitative research: Methodology and methods in mathematics education* (pp. 429–466). New York: Springer.

Bakker, A., & Gravemeijer, K. P. E. (2004). Learning to reason about distribution. In D. Ben-Zvi & J. Garfield (Eds.), *The Challenge of Developing Statistical Literacy, Reasoning and Thinking* (pp. 147–168). Dordrecht: Kluwer Academic Publishers.

Bang, M., & Medin, D. L. (2010). Cultural processes in science education: Supporting the navigation of multiple epistemologies. *Science Education*, *94*(6), 1008–1026.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*(1), 617–645. https://doi.org/10.1146/annurev.psych.59.103006.093639

Barsalou, L. W., Kyle Simmons, W., Barbey, A. K., & Wilson, C. D. (2003). Grounding conceptual knowledge in modality- specific systems. *Trends in Cognitive Sciences*, *7*(2), 84–91. https://doi.org/10.1016/S1364-6613(02)00029-3

Beaudry, J. S., & Miller, L. (2016). *Research literacy: A primer for understanding and using research*. New York: Guilford Press.

Ben-Zvi, D., & Garfield, J. (2004a). Statisical literacy, reasoning and thinking : Goals, definitions, and challenges. In D. Ben-Zvi & J. Garfield (Eds.), *The Challenge of*

Developing Statistical Literacy, Reasoning and Thinking (pp. 3–16). Dordrecht: Kluwer Academic Publishers.

Ben-Zvi, D., & Garfield, J. (2004b). *The challenge of developing statistical literacy, reasoning and thinking*. Dordrecht: Kluwer Academic Publishers.

Bertalanffy, L. von. (1968). *General system theory: Foundations, development, applications* (Revised). New York: George Braziller.

Best, J. (2004). *More damned lies and statistics: How numbers confuse public issues*. Berkeley, CA: University of California Press.

Biehler, R., Frischemeier, D., Reading, C., & Shaughnessy, J. M. (2018). Reasoning about data. In D. Ben Zvi (Ed.), *International Handbook of Research in Statistics Education* (pp. 139–192). Cham, Switzerland: Springer.

Biggs, J. B., & Collis, K. F. (1991). Multimodal learning and the quality of intelligent behavior. In H. A. H. Rowe (Ed.), *Intelligence, reconceptualization and measurement* (pp. 57–76). Hillsdale, NJ: Laurence Elbraum Associates.

Bohm, D. (2004). *On dialogue*. (L. Nichol, Ed.). London: Routledge.

Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior*, *11*, 717–726.

Brayboy, B. M. J., & McCarty, T. L. (2010). Indigenous knowledges and social justice pedagogy. New York: Routledge.

Breiman, L. (2001). Statistical Modeling: The Two Cultures. *Statistical Science*, *16*(3), 199–231. https://doi.org/10.2307/2676681

Breyfogle, F. W. (2003). *Implementing Six Sigma: Smarter Solutions using statistical methods* (2nd ed.). Hoboken, NY: John Wiley & Sons.

Brookhart, S. M. (2013). *How to create and use rubrics for formataive assessment and grading*. Alexandria, VA: ASCD.

Broström, S. (1999). Drama games with 6-year-old children: Possibilities and limitations. In Y. Engeström, R. Miettinen, & R.-L. Punamaki (Eds.), *Perspectives on activity theory*. Cambridge, UK: Cambridge University Press.

Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *The Journal of the Learning Sciences*, *2*(2), 141–178.

Brown, S. (2009). *Play: How it shapes the brain, opens the imagination, and invigorates the soul*. New York: Penguin Group.

Bruner, J. S. (1960). The process of education. *Phi Delta Kappan*, *53*(1), 97. https://doi.org/10.1017/CBO9781107415324.004

Bruner, J. S. (1961). The act of discovery. *Harvard Educational Review*. Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:The+Act+of+Discovery#0

Bruner, J. S. (1964). The course of cognitive growth. *American Psychologist*, *19*(1), 1–15. https://doi.org/10.1037/h0044160

Bruner, J. S. (1971). "The process of education" revisited. *Phi Delta Kappan*, *53*(1), 97. https://doi.org/10.1017/CBO9781107415324.004

Cajete, G. (2000). *Native science: Natural laws of interdependence*. Santa Fe, NM: Clear Light Publishers.

Carpenter, T. P., Fennema, E., Franke, M. L., Levi, L., & Empson, S. B. (2000). *Cognitively guided instruction: A research-based teacher professional development program for elementary school mathematics*. Madison, WI. https://doi.org/No. 003

Cartmill, E. A., Beilock, S., & Goldin-Meadow, S. (2012). A word in the hand: Action, gesture and mental representation in humans and non-human primates. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1585), 129–143. https://doi.org/10.1098/rstb.2011.0162

Chan, W.-T. (1963). *A sourcebook in Chinese philosophy*. Princeton, NJ: Princeton University Press.

Chew, P. K. H., & Dillon, D. B. (2014). Statistics Anxiety Update: Refining the Construct and Recommendations for a New Research Agenda. *Perspectives on Psychological Science*. https://doi.org/10.1177/1745691613518077

Chick, H., & Watson, J. (2001). Data representation and interpretation by primary school students working in groups. *Mathematics Education Research Journal*, *13*(2), 91–111.

Chinn, C. A., & Sherin, B. L. (2014). Microgenetic methods. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (2nd ed., pp. 171–190). New York: Cambridge University Press.

Clark, A. (1997). *Being there: Putting brain, body, and world together again*. Cambridge, MA: MIT Press.

Clements, D. H., & Sarama, J. (2004). Hypthetical learning trajectories. *Mathematical Thinking and Learning*, *6*(2).

Cobb, G. W., & Moore, D. S. (1997). Mathematics, Statistics, and Teaching. *American Mathematical Monthly*, *88*(9), 686–695.

Cobb, P., Confrey, J., DiSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher*, *32*(1), 9–13. https://doi.org/10.3102/0013189X032001009

Cobb, P. M. (2004). Principles of instructional design for supporting the development of students' statistical reasoning. In D. Ben-Zvi & J. Garfield (Eds.), *The Challenge of Developing Statistical Literacy, Reasoning and Thinking* (pp. 375–395). Dordrecht: Kluwer Academic Publishers.

Cohen, J. (1994). The earth is round (p < .05): Rejoinder. *American Psychologist*, *50*(12), 1103–1103. https://doi.org/10.1037/0003-066X.50.12.1103

Collins, A., Joseph, D., Bielaczyc, K., Collins, A., & Bielaczyc, K. (2016). Design Research : Theoretical and Methodological Issues Linked references are available on JSTOR for this article : Design Research : Theoretical and Methodological Issues, *13*(1), 15–42.

Crivellato, E., & Ribatti, D. (2007). Soul, mind, brain: Greek philosophy and the birth of neuroscience. *Brain Research Bulletin*, *71*(4), 327–336. https://doi.org/10.1016/j.brainresbull.2006.09.020

Csikszentmihalyi, M. (2014). Flow and education. In M. Csikszentmihalyi (Ed.), *Applications of Flow in Human Development and Education* (pp. 129–151). Dordrecht: Springer.

Cuisenaire, G., & Gattegno, C. (1962). *Initiation a la méthode-les nombres en couleurs*. Lonay, Switzerland: Delachaux et Niestlé.

Davydov, V. V. (2008). *Problems of developmental instruction*. New York: Nova Science Publishers.

Delmas, R. C., Garfield, J. B., Ooms, A., & Chance, B. L. (2007). Assessing students' conceptual understanding after a first course in statistics. *Statistics Education Research Journal*, *6*(2), 28–58. Retrieved from http://eprints.kingston.ac.uk/21003/

Deming, W. E. (1986). *Out of the crisis*. Cambridge, MA: MIT Press.

Deming, W. E. (1993). *The New economics: For industry, government, education* (2nd ed.). MIT Press.

Derry, S. J. (2007). *Guidelines for video research in education: Recommendations from an expert panel*. Chicago.

Derry, S. J., Pea, R. D., Barron, B., Engle, R. A., Erickson, F., Goldman, R., … Sherin, B. L. (2010). Conducting video research in the learning sciences: Guidance on selection, analysis, technology, and ethics. *Journal of the Learning Sciences*, *19*(1), 3–53. https://doi.org/10.1080/10508400903452884

Design-Based_Research_Collective. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher*, *32*(1), 5–8, 35–37.

Dewey, J. (1910). *How we think*. Boston: D. C. Heath and Company.

Dewey, J. (1938a). *Experience and education*. New York: Macmillan.

Dewey, J. (1938b). *Logic: The theory of inquiry*. New York: Henry Holt and Co.

Dienes, Z. P. (1964). *The power of mathematics: A study of the transition from the constructive to the analytical phase of mathematical thinking in children*. London: Hutchinson Educational.

diSessa, A. A. (2014). A history of conceptual change research: Threads and fault lines. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (2nd ed., pp. 88–108). New York: Cambridge University Press.

Donald, M. (1991). *Origins of the modern mind: Three stages in the evolution of culture and cognition*. Cambridge, MA: Harvard University Press.

Driscoll, M. P. (2000). *Psychology of learning for instruction* (2nd ed.). Boston: Allyn and Bacon.

Duschl, R. (2008). Science education in three-part harmony: Balancing conceptual, epistemic, and social learning goals. *Review of Research in Education*, *32*(1), 268–291. https://doi.org/10.3102/0091732X07309371

Edelman, G. M. (1989). *The remembered present: A biological theory of consciousness*. New York: Basic Books.

Editor, (2017). Towards greater reproducibility. *Nature*, *546*(7656), 8.

Einstein, A. (1905). On the electrodynamics of moving bodies (Zur Elektrodynamik bewegter Körper). *Annalen Der Physik*, *17*, 891. https://doi.org/10.1002/andp.19053221004

Einstein, A., & Infeld, L. (1938). *The evolution of physics: The growth of ideas from early concepts to relativity and quanta*. Cambridge, UK: Cambridge University Press.

Eisley, L. C. (1962). *Francis Bacon and the modern dilemma*. Lincoln, NE: University of Nebraska Press.

Engeström, Y. (1999). Activity theory and individual and social transformation. In Y. Engeström, R. Miettinen, & R.-L. Punamaki (Eds.), *Perspectives on activity theory* (pp. 1–16). Cambridge, UK: Cambridge University Press.

Feldman, J., & Narayanan, S. (2004). Embodied meaning in a neural theory of language. *Brain and Language*, *89*(2), 385–392. https://doi.org/10.1016/S0093-934X(03)00355-9

Feyerabend, P. (1975). *Against method* (4th ed.). London: Verso.

Fielding-Wells, J. (2018). Scaffolding statistical inquiries for young children. In A. Leavy, M. Meletiou-Mavrotheris, & E. Paparistodemou (Eds.), *Statistics in early childhood and primary education* (pp. 109–127). Singapore: Springer.

Fleer, M. (2009). *Play and learning in early childhood settings: International perspectives*. Springer.com.

Fleer, M. (2011). "Conceptual play": Foregrounding imagination and cognition during concept formation in early years education. *Contemporary Issues in Early Childhood, 12*(3), 224–240. https://doi.org/10.2304/ciec.2011.12.3.224

Franklin, C., Bargagliotti, A. E., Case, C. A., Kader, G., Scheaffer, R., & Spangler, D. A. (2015). *The statistical education of teachers*. Retrieved from http://www.amstat.org/education/SET/SET.pdf

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2005). *Guidelines for assessment and instruction in statistics education (GAISE) report: A pre-K-12 curriculum framework*. *Report*. Alexandria, VA: American Statistical Association. https://doi.org/10.3928/01484834-20140325-01

Freire, P. (1970). *Pedagogy of the oppressed*. New York: Continuum Publishing Company. https://doi.org/10.1215/00382876-1472612

Fröbel, F. (n.d.). *The education of man*. (W. N. Hailmann, Ed.). Mineola, NY: Dover.

Fröbel, F. (1887). *The education of man*. (W. N. Hailmann, Ed.). New York: D. Appleton.

Furtak, E. M., Seidel, T., Iverson, H., & Briggs, D. C. (2012). Experimental and quasi-experimental studies of inquiry-based science teaching: A meta-analysis. *Review of Educational Research*, *82*(3), 300–329. https://doi.org/10.3102/0034654312457206

Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International Statistical Review*, *70*(1), 1–51. https://doi.org/10.1111/j.1751-5823.2002.tb00336.x

Gal, I., Ginsburg, L., & Schau, C. (1997). Monitoring attitudes and beliefs in statistics education. In I. Gal & J. B. Garfield (Eds.), *The Assessment Challenge in Statistics*

Education (Vol. 1997, pp. 37–51). Amsterdam: IOS Press. Retrieved from http://www.stat.auckland.ac.nz/~iase/publications/assessbkref.

Galton, F. (1894). Natural inheritance. London: Macmillan & Co. Retrieved from file:///home/joel/.local/share/data/Mendeley Ltd./Mendeley Desktop/Downloaded/Unknown - Unknown - Natural inheritance. Galton Francis.pdf.pdf

Garfield, J., & Ahlgren, A. (1988). Difficulties in learning basic concepts in probability and statistics: Implications for research. *Journal for Research in Mathematics Education, 19*(1), 44–63.

Garfield, J. B., & Ben-Zvi, D. (2008). *Developing students' statistical reasoning: Connecting research and teaching practice*. Dortrecht: Springer. https://doi.org/10.1007/978-1-4020-8383-9

Garfield, J., & Ben Zvi, D. (2007). How students learn statistics revisited: A current review of research on teaching and learning statistics. *International Statistical Review*, (December 2007), 1–43. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1751-5823.2007.00029.x/full%5Cnfile:///Users/Njaal/Dropbox/Library.papers3/Files/2E/2E5A84DF-1C76-4A45-8EFA-940995A195B9.pdf%5Cnpapers3://publication/doi/10.1111/j.l751-5823.2007.00029.x

Gattegno, C. (1961). *Mathematics at your fingertips*. Canada: National Film Board of Canada. Retrieved from http://www.cuisenaire.co.uk/index.php/home/videos

Ghoshal, S. (2005). Bad management theories are destroying good management practices. *Academy of Management Learning & Education*, *4*(1), 75–91. https://doi.org/10.1109/EMR.2005.26768

Giebitz, R. (2015). Learning statistics through playing with blocks: An exploratory study in embodied interaction with a frequency distribution. In *Southwest Consortium for Innovative Psychology in Education (SCIPIE)*. Albuquerque, NM.

Giebitz, R. P., & Stanton, M. (2018). *Evaluation report for the New Mexico ECHO for Education inaugural network*. Albuquerque, NM.

Ginsburg, H. P. (2018). Helping teacher educators in institutions of higher learning to prepare prospective and practicing teachers to teach mathematics to young children. In G. Kaiser, H. Forgasz, M. Graven, A. Kuzniak, E. Simmt, & B. Xu (Eds.), *Invited Lectures from the 13th International Congress on Mathematical Education* (pp. 135–154). New York.

Goldman, R., Zahn, C., & Derry, S. J. (2014). Frontiers of digital video research in the learning sciences: Mapping the terrain. In R. K. Sawyer (Ed.), *The Cambridge* *handbook of the learning sciences* (2nd ed., pp. 213–232). Cambridge, UK: Cambridge University Press.

Goodman, Y. M., & Goodman, K. S. (1990). Vygotsky in a whole-language perspective. In L. C. Moll (Ed.), *Vygotsky and Education: Instructional Implications and Applications of Sociohistorical Psychology* (pp. 223–250). Cambridge, UK: Cambridge University Press.

Gould, R. (2017). Data literacy is statistical literacy. *Statistics Education Research Journal*, *16*(1), 22–25.

Goutard, M. (1964). *Mathematics and children*. Reading, UK: Educational Explorers.

Grandy, R., & Duschl, R. A. (2007). Reconsidering the character and role of inquiry in school science: Analysis of a conference. *Science and Education*, *16*(2), 141–166. https://doi.org/10.1007/s11191-005-2865-z

Green, R. (1980). Culturally-based science: The potential for traditional people, science and folklore. *American Association for the Advancement of Science*, 204–212.

Grow, G. O. (1991). Teaching learners to be self-directed. *Adult Education Quarterly, 41*(3), 125–149.

Hall, R., & Jurow, A. S. (2015). Changing concepts in activity: Descriptive and design studies of consequential learning in conceptual practices. *Educational Psychologist*, *50*(3), 173–189. https://doi.org/10.1080/00461520.2015.1075403

Hannigan, A., Gill, O., & Leavy, A. M. (2013). An investigation of prospective secondary mathematics teachers' conceptual knowledge of and attitudes towards statistics.

Hattie, J., & Yates, G. C. R. (2014). *Visible learning and the science of how we learn*. London: Routledge.

Hewitt, K. (2001). Blocks as a tool for Learning: Historical and contemporary perspectives. *Young Children*, (January), 6–13.

Huizinga, J. (1950). Homo ludens: A study of the play-element in culture, (1956), 226.

Hutchins, E. (1995). *Cognition in the wild*. Cambridge, MA: MIT Press.

Hutchins, E. (2014). The cultural ecosystem of human cognition. *Philosophical Psychology*, *27*(1), 34–49. https://doi.org/10.1080/09515089.2013.830548

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, *2*(8), 0696–0701. https://doi.org/10.1371/journal.pmed.0020124

Isaacs, W. (1999). *Dialogue and the art of thinking together*. New York: Doubleday.

Jones, G. A., Langrall, C. W., Mooney, E. S., & Thornton, C. A. (2004). Models of development in statistical reasoning. In D. Ben-Zvi & J. Garfield (Eds.), *The Challenge of Developing Statistical Literacy, Reasoning and Thinking* (pp. 97–118). Dordrecht: Kluwer Academic Publishers.

Jones, G. A., Thornton, C. A., Langrall, C. W., Mooney, E. S., Perry, B., & Putt, I. J. (2000). A framework for characterizing children's statistical thinking. *Mathematical Thinking and Learning*, *2*(4), 269–307. https://doi.org/10.1207/S15327833MTL0204

Jordan, B., & Henderson, A. (1995). Interaction Analysis: Foundations and practice. *The Journal of the Learning Sciences*, *4*(1), 39–103. https://doi.org/10.1207/s15327809jls0401

Kahneman, D. (2011). *Thinking fast and slow*. New York: Farrar, Straus and Giroux.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: heuristics and biases*. Cambridge, MA: Cambridge University Press.

Kant, I. (1998). *Critique of Pure Reason*. Cambridge, UK: Cambridge University Press.

Kaplan, R. S., & Norton, D. P. (1996). *The Balanced Scorecard: Translating strategy into action*. Boston: Harvard Business School Press.

Kidman, J., Abrams, E., & McRae, H. (2011). Imaginary subjects: School science, indigenous students, and knowledge - power relations. *British Journal of Sociology of Education*, *32*(2), 201–220.

Krathwohl, D. R. (2002). A Revision of Bloom's Taxonomy : An overview. *Theory Into Practice*, *41*(4), 212–218. https://doi.org/10.1207/s15430421tip4104

Kuhn, T. S. (1962). *The Structure of Scientific Revolutions* (2nd ed.). Chicago: University of Chicago Press.

Kunert, J., Montag, A., & Pöhlmann, S. (2001). The quincunx: History and mathematics. *Statistical Papers*. Springer-Verlag.

Lakoff, G. (2015a). How Brains Think: The Embodiment Hypothesis. In *International Convention of Psychological Science ICPS 2015*. Amsterdam: Association for Psychological Science (APS). Retrieved from https://www.youtube.com/user/PsychologicalScience/search?query=lakoff

Lakoff, G. (2015b). The Brain's Mathematics: The Cognitive and Neural Foundations of Mathematics. In *2016 Breakthrough Prize Symposium*. Berkeley, CA: University of California Berkeley. Retrieved from https://www.youtube.com/watch?v=S-TZBvjc3s8

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.

Lakoff, G., & Nuñez, R. E. (2000). *Where mathematics comes from: How the embodied mind brings mathematics into being*. New York: Basic Books.

Lange, E. A. (2012). Transforming transformative learning through sustainability and the new science. In E. W. Taylor & P. Cranton (Eds.), *The handbook of tranformative learning: Theory, research and practice* (pp. 195–211). San Francisco, CA: Jossey-Bass.

Lau, D. C., & Ames, R. T. (1998). *Yuan Dao: Tracing Dao to Its Source*. New York: Ballantine Books.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, UK: Cambridge University Press.

Leavy, A. M., & Hourigan, M. (2018). Inscriptional capacities and representations of young children engaged in data collection during a statistical investigation. In A. Leavy, M. Meletiou-Mavrotheris, & E. Paparistodemou (Eds.), *Statistics in early childhood and primary education* (pp. 89–107). Singapore: Springer.

Lighthall, F. F. (1991). Launching the space shuttle Challenger: Disciplinary deficiencies in the analysis of engineering data. *IEEE Transactions on Engineering Management*, *38*(1), 63–74. https://doi.org/10.1109/17.65761

Liker, J. K., & Meier, D. P. (2007). *Toyota talent: Developing your people the Toyota way*. New York: McGraw Hill.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Los Angeles: SAGE.

Locke, J. (1693). *Some thoughts concerning education*. *Philosophy of education: The essential texts*. London: Black Swan. https://doi.org/10.2307/3118337

Loewit-Phillips, P. M., & Goldbas, A. (2013). History and Impact. *International Journal of Childbirth Education*, *28*(4), 62–68.

Makar, K. (2016). Developing young children’s emergent inferential practices in statistics. *Mathematical Thinking and Learning*, *18*(1), 1–24. https://doi.org/10.1080/10986065.2016.1107820

Makar, K. (2018). Theorising links between context and structure to introduce powerful statistical ideas in the early years. In A. Leavy, M. Meletiou-Mavrotheris, E. Paparistodemou, & efi (Eds.), *Statistics in early childhood and primary education* (pp. 3–20). Singapore: Springer.

Malouf, D. B., & Taymans, J. M. (2016). Anatomy of an evidence base. *Educational Researcher*, *45*(8), 454–459. https://doi.org/10.3102/0013189X16678417

Meletiou-Mavrotheris, M. ., Mavrotheris, E. ., & Paparistodemou, E. . b. (2011). Distance learning for teacher professional development in statistics education. *Teaching Statistics*, *33*(1), 2–8. https://doi.org/10.1111/j.1467-9639.2010.00425.x

Merriam, S. B. (1998). *Qualitative research and case study applications in education*. San Francisco, CA: Jossey-Bass.

Miller, J. (2008). Toyota's top engineer on how to develop thinking people. Retrieved October 30, 2017, from http://blog.gembaacademy.com/2008/08/04/toyotas_top_engineer_on_how_to_develop_thinking_pe/

Montessori, M. (1912). *The Montessori Method*. New York: Schoken Books.

Nagel, T. (1974). What is it like to be a bat ? *The Philosophical Review*, *83*(4), 435–450.

Nakane, J., & Hall, R. W. (2002). Ohno' s Method: Creating a survival work culture. *Target*, *18*(1), 6–15. Retrieved from www.ame.org

National Research Council (U.S.). (1996). *National Science Education Standards*. Washington D.C.: National Academy Press. https://doi.org/10.17226/4962

National Research Council (U.S.). (1999). *How people learn: Brain, mind, experience, and school*. (J. D. Bransford, A. L. Brown, & R. R. Cocking, Eds.), *Committee on learning research and educational practice* (Expanded). Washington, D.C.: National Academy Press. https://doi.org/10.1016/0885-2014(91)90049-J

National Research Council (U.S.). (2001). *Eager to learn: Educating or preschoolers*. (B. T. Bowman, M. S. Donovan, & M. S. Burns, Eds.). Washington, DC: National Acadmey Press.

National Research Council (U.S.). (2005). *How students learn: Science in the classroom*. (M. S. Donovan & J. d. Bransford, Eds.). Washington D.C.: The National Academies Press. https://doi.org/10.17226/661

National Research Council (U.S.). (2012). *Education for life and work: Developing transferable knowledge and skills in the 21st century*. (J. W. Pellegrino, M. L. Hilton, & D. D. Learning, Eds.). Washington D.C.: The National Academies Press. Retrieved from https://www.hewlett.org/wp-content/uploads/2016/08/Education_for_Life_and_Work.pdf

Neave, H. R. (1990). *The Deming dimension*. Knowxville, TN: SPC Press.

O'Sullivan, E. (2012). Deep transformation: Forging a planetary worldview. In E. W. Taylor & P. Cranton (Eds.), *The handbook of tranformative learning: Theory, research and practice* (pp. 162–177). San Francisco, CA: Jossey-Bass.

Oceans of Data Institute. (2015). *Building global interest in data literacy : A dialogue*. Waltham, MA. Retrieved from http://oceansofdata.org/

OECD. (2012). *Preparing teachers and developing school leaders for the 21st century: Lessons from around the world*. (A. Schleicher, Ed.). Paris: OECD Publishing. https://doi.org/10.1787/9789264174559-en

Ohno, T. (1978). *The Toyota Production System: Aiming to manage free from economies of scale*. Tokyo: Diamond.

Onwuegbuzie, A. J., & Wilson, V. A. (2003). Statistics anxiety: Nature, etiology, antecedents, effects, and treatments: A comprehensive review of the literature. *Teaching in Higher Education*, *8*(2), 195–209. https://doi.org/10.1080/1356251032000052447

Pea, R. D. (1993). Practices of distributed intelligence and designs for education. In G. Salomon (Ed.), *Distributed cognitions: Psychological and educational considerations* (pp. 47–87). Cambridge, UK: Cambridge University Press.

Pfannkuch, M., & Wild, C. (2004). Towards an understanding of statistical thinking. In D. Ben-Zvi & Joan Garfield (Eds.), *The Challenge of Developing Statistical Literacy, Reasoning and Thinking* (pp. 17–46).

Piaget, J. (1948). *To understand is to invent: The future of education*. New York: Grossman Publishers.

Piaget, J. (1970). *Genetic Epistemology*. New York: Norton & Co.

Polanyi, M. (1958). *Personal knowledge: Towards a post-critical philosophy*. Chicago: University of Chicago Press.

Powell, A. B., Francisco, J. M., & Maher, C. A. (2003). An analytical model for studying the development of learners' mathematical ideas and reasoning using videotape data. *Journal of Mathematical Behavior*, *22*(4), 405–435. https://doi.org/10.1016/j.jmathb.2003.09.002

Pullinger, J. (2013). Statistics making an impact. *Journal of the Royal Statistical Society. Series A: Statistics in Society*, *176*(Part 4), 819–839.

Ramirez, C., Schau, C., & Emmioğlu, E. (2012). The importance of attitudes in statistics education. *Statistics Education Research Journal*, *11*(2), 57–71.

Read, J. (1992). A short history of children's building blocks. In P. Gura (Ed.), *Exploring learning: Young children and blockplay* (pp. 1–13). London: The Froebel Blockplay Research Group.

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

Saldaña, J. (2013). *The coding manual for qualitative researchers* (2nd ed.). Los Angeles: SAGE.

Santagata, R., Yeh, C., & Mercado, J. (2018). Preparing Elementary School Teachers to Learn From Teaching: A Comparison of Two Approaches to Mathematics Methods Instruction. *Journal of the Learning Sciences*, *27*(3), 1–43. https://doi.org/10.1080/10508406.2018.1441030

Schau, C. (2008). Common issues in SATS © research. In *Joint Statistical Meetings Proceedings* (pp. 1–4). Denver, CO: American Statistical Association.

Schau, C., & Stevens, J. (1995). The development and validation of the Survey of Attitudes Toward Statistics. *Educational and Psychological Measurement*, *55*(5), 868–875.

Schonberger, R. J. (1986). *World Class Manufacturing: The lessons of simplicity applied*. New York: The Free Press.

Schuyler W. Huck. (2009). *Statistical misconceptions*. New York: Routledge.

Senge, P. M. (1990). *The fifth discipline: The art and practice of the learning organization*. New York: Doubleday.

Serradó Bayés, A., Meletiou-Mavrotheris, M., & Paparistodemou, E. (2014). EarlyStatistics: A course for developing teachers' statistics technological and pedagogical content. *Statistique et Enseignement*, *5*(1), 5–29. Retrieved from http://publications-sfds.math.cnrs.fr/index.php/StatEns/article/view/298/279

Sfard, A., & McClain, K. (2002). Analyzing tools: Perspective on the role of designed artifacts in mathematics learning. *The Journal of the Learning Sciences*, *11*(2&3), 153–161. https://doi.org/10.1207/S15327809JLS11,2-3n_1

Sheets-Johnstone, M. (2011). *The primacy of movement* (2nd ed.). Amsterdam: John Benjamins Publishing Company.

Shewhart, W. (1931). *Economic control of quality of manufactured product*. New York: Van Nostrand Company.

Simon, M. a. (1995). Reconstructing mathematics pedagogy from a constructivist perspective. *Journal for Research in Mathematics Education*, *26*(2), 114–145. https://doi.org/10.2307/749205

Sohn, D. (1991). The interaction concept in scientific discourse and the analysis of variance. *Journal of Psychology*, *125*(6), 621–629.

Soledad, E. (2018). Data representations in early statistics: Data sense, meta-representational competence and transnumeration. In *Statistics in early childhood and primary education* (pp. 239–256). Singapore.

Stake, R. E. (2006). *Multiple case study analysis*. New York: Guilford Press.

The Deming Institute. (1980). Red Bead Experiment. Retrieved January 22, 2018, from https://deming.org/explore/red-bead-experiment

Thelen, E. (2000). Grounded in the world: Developmental origins of the embodied mind. *Infancy*, *1*(1), 3–28. https://doi.org/10.1207/S15327078IN0101

Thelen, E., Schöner, G., Scheier, C., & Smith, L. B. (2001). The dynamic of embodiment: a field theory of infant preservative reaching. *Behavioral and Brain Sciences*, *24*, 1–86.

Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press.

Thomas, D., & Seely_Brown, J. (2011). *A new culture of learning: Cultivating the imagination for a world of constant change*. Lexington, KY: CreateSpace.

Tufte, E. R. (1997). Visual and statistical thinking: Displays of evidence for making decisions. In *Visual explanations: Images and quantities, evidence and narrative*. Cheshire, CT: Graphics Press.

Tukey, J. W. . (1962). The future of data analysis. *Annals of Mathematical Statistics*, *33*(1), 1–67.

Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin*, *2*, 105–110.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124

UNESCO. (2006). *Millennium Development Goals and statistical literacy*.

United Nations Economic Commission for Europe. (2012). *Making data meaningful Part 4: A guide to improving statistical literacy*. Geneva. Retrieved from http://www.unece.org/fileadmin/DAM/stats/documents/writing/Making_Data_Meaningful_Part_4_for_Web.pdf

United Nations General Assembly. Transforming our world: The 2030 Agenda for Sustainable Development, 16301 Resolution adopted by the General Assembly on 25

September 2015 § (2015). New York: United Nations General Assembly. https://doi.org/10.1007/s13398-014-0173-7.2

University of Minnesota. (2006). Assessment Resource Tools for Improving Statistical Thinking. Retrieved December 9, 2016, from https://apps3.cehd.umn.edu/artist/index.html

Uttal, D. H., Marzolf, D. P., Pierroutsakos, S. L., Smith, C. M., Troseth, G. L., Scudder, K. V., & DeLoache, J. S. (1998). Seeing through symbols: The development of children's understanding of symbolic relations. In O. N. Saracho & B. Spodek (Eds.), *Multiple perspectives on play in early childhood education* (pp. 59–79). Albany, NY: State University of New York.

Varela, F. J., Thompson, E., & Rosch, E. (2016). *The embodied mind: Cognitive science and human experience* (Revised). Cambridge, MA: MIT Press.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.). Cambridge, MA: Harvard University Press.

Watson, J., Fitzallen, N., Fielding-Wells, J., & Madden, S. (2018). The practice of statistics. In D. Ben Zvi, K. Makar, & J. Garfield (Eds.), *International Handbook of Research in Statistics Education* (pp. 105–138). Cham, Switzerland: Springer.

Watson, J. M. (2006). *Statistical literacy at school: Growth and goals*. Mahwah, NJ: Lawrence Elbraum Associates.

Watson, J. M. (2011). Foundations for improving statistical literacy. *Statistical Journal of the IAOS*, *27*(3–4), 197–204. https://doi.org/10.3233/SJI-2011-0728

Whitaker, D., Foti, S., & Jacobbe, T. (2015). The Levels of Conceptual Understanding in Statistics (LOCUS) Project: Results of the pilot study. *Numeracy*, *8*(2), 1–16. https://doi.org/10.5038/1936-4660.8.2.3

Wilbur, K. (1984). *Quantum questions: Mystical writings of the world's great physicists*. Boulder, CO: Shambhala.

Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, *67*(3), 223–248.

Wild, C. J., Pfannkuch, M., Regan, M., & Horton, N. J. (2011). Towards more accessible conceptions of statistical inference. *Journal of the Royal Statistical Society. Series A: Statistics in Society*, *174*(2), 247–295. https://doi.org/10.1111/j.1467-985X.2010.00678.x

Wild, C. J., Utts, J. M., & Horton, N. J. (2018). What is statistics? In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International Handbook of Research in Statistics Education* (pp. 5–36). Cham, Switzerland: Springer.

Wilkinson, L. (1999). Statistical methods in psychology journals. *American Psychologist*, *54*(8), 594–604.

Wilson, F. R. (1998). *The hand: How its use shapes the brain, language, and human culture.* New York: Random House.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, *17*, 89–100.

Yeager, D. M. (2004). Confronting the minotaur : Moral inversion and Polanyi 's moral philosophy, (M 23), 22–48.

Yin, R. K. (2014). *Case study research: Design and methods* (5th ed.). Los Angeles: SAGE.

Appendices

Appendix A: Case Summaries

Case 1.1: Fiona (pseudonym)

This 6.3-year-old kindergarten girl likes recess and reading. She likes challenges, is outgoing and talkative. She began the session by stacking the blocks. The stack fell twice but she eventually made a free-standing stack of 14 blocks. Then she made a square and then a castle. She told a story about the castle and a princess who ripped her dress. Then she made a person. I joined her in the block play. Eleven minutes into the session, I transitioned to Lesson 1 showing her how to align blocks with the numbers on the ruler then balance them on a pair of pencils. She said confidently, "I could do that." She arranged the blocks in a symmetric pattern and they fell off the ruler when she tried to balance it. She put them back on and balanced the ruler. She read the balance point on the ruler and wrote it on the worksheet. I demonstrated how to make an X-plot. She completed the X-plot and checked it against the blocks. I demonstrated a method for finding "the block in the middle" (median) by taking blocks off the end, two at a time, one with each hand, until there was one left. I wrote the symbol $\tilde{X}$ on the worksheet. We called this symbol "X-squiggly." In the fourth learning interaction, she found Q1 and Q3; she seemed to get a little fatigued. During this and many following sessions, she talked aloud about what she was doing and what she was thinking. She also talked to herself, often inaudibly.

In the fifth learning interaction, she found the minimum (min) and the maximum (max) and recorded the number of blocks (N). This took about a minute and half. This might have been a good time to take a break (27 minutes into the session) but I asked her to continue by making another arrangement of blocks on the ruler. About a minute into this interaction, she became more playful. I asked her to write down max=10 which prompted the following exchange (numbers in brackets are time stamps]:

Fiona: I can't show you that with my fingers because I only have 8 fingers and 2 thumbs.

Tutor: Tell me more.

Fiona: Uh.

Tutor: If you use your thumb, do you think you could count to 10? [she nods yes]. On your fingers? [she holds up her fingers.]

Fiona: 1,2,3,4,5,6,7,8,9,10 [she folds down her fingers as she counts]

Tutor: Very good.

Fiona: 11,12,13,14,15,16,17,18,19,20. If you count your toes.

Tutor: You'd have to take your shoes off [30:00] for that.

Fiona: But you still could count your toes [she chuckles].

Tutor: Yeah. Can you count them without seeing them?

Fiona: 1,2,3,4,5 [she counts her toes through her shoes.]

Tutor: Can you count your toes without seeing them?

Fiona: Six, seven, eight, nine, ten [she grabs her other foot.]

Tutor: Oh. You can feel them through your shoes. Ok.

Fiona: My shoes stop right there, and my toes stop right there. [She squeezes her shoe to show the tips of her toes.] There.

I understood this playful exchange as a signal to shift to a different activity. She assented, and we began the *Dice Roll* lesson. As we got into the lesson, her energy level rose. Early in this lesson, we had the following exchange after she rolled a five and a six:

Fiona: Ok. [FB grabs the dice, shakes them, and throws them across the table.] Five plus six? I don't have enough fingers for that. [34:00]

Tutor: Oh no! Do you want to borrow one of mine?

Fiona: Ok. [Smiling]

Tutor: Here. You can borrow one of mine. Which one do you want to borrow?

Fiona: Uhh. I don't know.

Tutor: This one? [Tutor holds up a finger.]

Fiona: I guess.

Tutor: Ok.

Fiona: 1,2,3,4,5,6,7. Ok. Ten. [she extends a finger with each count.] Ok. Eleven. [She counts the tutor's finger.] I'll put a block on eleven. [she places a block on the ruler at 11].

Fiona completes the lesson making an X-plot and finding min, max, N, $\bar{X}$, $\tilde{X}$, Q1, and Q3. This lesson took about 20 minutes; however, some video was lost due to a camera malfunction. During this lost segment of video, Fiona made her first fraction after reading the balance point on the ruler half way between 6 and 7. After finding Q1, we had the following exchange:

Fiona: Well that took forever.

Tutor: Yeah. That took forever. It's getting late. It's time that we stop.

Fiona: Ok.

Tutor: Ok. Let's stop. Do you want to do this again next week?

Fiona: There's this one more answer to do [She points to "Q3" on the worksheet.]

Tutor: Do you want to get that one more answer? [She nods.] You do?

Fiona: Yeah.

Fiona found Q3, then we ended the session. I asked her what was fun, what was hard, and if she wanted to meet again. She liked stacking the blocks as high as she could. Finding Q1 and Q3 was difficult for her. She wanted to meet again the following week.

Case 1.1 Session 2

In our initial rapport building segment, Fiona did almost all the talking. I asked her what she remembered from last session. She remembered balancing the blocks and how they were stacked on the ruler.

Fiona completed one Free Form block arrangement, made an X-plot and found the statistics. We started a Dice Roll lesson but after a couple of dice rolls, she said, “I know what two times two is.” Then, following her prompt, I showed her how to multiply using the blocks illustrating 2x3, 3x3, and 3x4. Then I asked her to do 4x4 which she did easily. We started over with the Dice Roll but she seemed to lose interest, so I introduced the Marble Grab. I posed the problem: “Let’s figure out how big your hand is.” She suggested tracing it. [connecting to prior knowledge] I suggested finding out “how much you can hold in your hand.” She showed how she can hold two blocks, then three blocks, then four blocks. Then I ask, “How many marbles can you hold?” As I made a data collection sheet, she continued to play with the blocks. Then I gave her a bowl of marbles and showed her how to grab the marbles and count how many in a handful. We worked together to count them. We did 5 trials alternating left and right hands. The Marble Grab took about 20 minutes. We did not have time to analyze the data before the end of the session. At the end of the lesson, I began a conversation about variation, but we ran out of time.

Fiona liked stacking the blocks on the ruler. She liked the marbles more than the blocks. It was hard for her to balance the blocks on the pencils.

Case 1.1 Session 3

I reviewed the lesson plan with Fiona. She used a pencil as a unit of measure to show how much she liked the worksheet and another worksheet in one of her classes. She made a castle on the ruler like the one she made in the previous session. We talked about the shape. She found all 10 statistics. She made an X-plot and balanced the ruler on the pencils. The balance point was a fraction: 7¾. After finding the balance point, we had the following exchange:

Fiona: Oh… Why do we only do this once a week? Oh yeah, because it’s a Tuesday kind of thing.

Tutor: It’s a Tuesday kind of thing for now. Do you want to do it more often?

Fiona: Yes.

Tutor: It would be more fun for me too, to do it more often.

Fiona: I like it a whole lot.

Such enjoyment of a first introduction to statistics might prevent statistics anxiety and negative attitudes toward statistics in the future.

Boxplot

Fiona began to show signs of fatigue when we got to IQR and the boxplot (box & whiskers) after about 31 minutes into the session. She rocked back and forth for about 23 seconds then yawned. But when it came time to add the whiskers to the boxplot, she perked up:

Fiona: We need whiskers!

Tutor: Yes!

Fiona: I need whiskers; I want to make the whiskers!

Then she got more playful:

Fiona: I'm going to make three whiskers.

Tutor: Three whiskers.

Fiona: Like in the cartoon; making three whiskers on the cats.

At 34 minutes into the session, I gave her a choice between the Dice Throw and the Marble Grab. She chose the Marble Grab. Although she enjoyed the session, she was getting tired and interjected, "You know what?" Then she told a story about when she was a baby and had casts on her legs but now she takes ballet and can do the splits. She showed me her splits and I joined her in stretching out my legs. She had one more "Do you know what?" before she agreed to begin the Marble Grab 41 minutes into the session. Within 5 minutes, she had three more Do-you-know-what?'s, an indication that she was getting tired. I did not realize at the time that we should stop and reflect on what we had accomplished, or maybe just play. We continued, and she made an X-plot of the marble counts. I extended the template on the worksheet to accommodate numbers up to 14.

Case 1.1 Session 4

Fiona completed one Dice Roll, including the X-plot, 10 statistics, and a boxplot. She performed four trials of the Marble Grab and completed the Marble Grab lesson from the previous session by transforming the average, Q1, Q3, and IQR back to grams, labeling the boxplot, and giving a title to the document: "marble wates."

During the Dice Roll lesson, she referred to the ruler as being like a number line: "The ruler is practically a number line." [connecting to prior knowledge] She expressed disapproval of the 5 after rolling her fourth 5: "Ah, the 5 is getting too high!" Then, "I wanna take a break from the 5's" and "Five is too tall." Then, "You evil 5." And "Five. The evil 5. I wish 6 got one." [10:45] And finally, "The evil 5 took all the blocks." Later in the lesson she anthropomorphized her pencil: "My pencil was cheating."

After we found the min, max, mode, and N, she offered some insightful logic into language:

Fiona: Do you know what? [Interrupting] You say fifteen the backwards way you spell it. You say it the backwards way you spell it with numbers.

Tutor: You say it backwards?

Fiona: You say it with fifteen. And the teen is ten and then the five. But really, if you said it the same way it's spelled with numbers, it would be "teen fif". [She takes pleasure in her logic.]

Tutor: That's an astute observation. [Tutor repeats to himself "teenfif"]

Fiona: What does that mean?

Tutor: Astute means smart. It means clever. It means that you see things that other people don't see. It means you have insight. It means you see with your inner eye.

This was Fiona's fifth time finding Q1, Q3, and IQR. This learning interaction went along smoothly at a good pace. She invented a new term for the boxplot: "my whiskery box". After completing the boxplot, we had the following exchange:

Tutor: Perfect. You're done. You did a lot. A lot of hard playing today. Are you getting tired?

Fiona: No.

Tutor: You're not. Do you want to play with marbles now?

Fiona: Yes

Tutor: Ok. We're going to do something a little bit different than we did last time with the marbles. Ok?

After four trials grabbing handfuls of marbles and counting them we had a brief conversation about variation.

Case 1.1 Session 5

Fiona completed a series of complex tasks with a high level of engagement for 47½ minutes. She had a choice of doing the Marble Grab or weighing the marbles. She chose to weigh the marbles. I demonstrated zeroing the scale, units of measure (grams), and decimals. I demonstrated making a data collection sheet including title and column headings. She created a dataset by weighing 10 marbles to a tenth of a gram. She learned about outliers. She made a Marble Plot and found the minimum, maximum, N, range, mode, and median by inspecting the marble plot. She labeled the outlier. I wrote the numbers 1 to 12 underneath the axis of the marble plot and showed her how to transform the marble weights to block positions on the ruler. She placed blocks on the ruler in the corresponding location and found the balance point. She transformed the balance point back to the scale of the marble weights and wrote the average on the worksheet. I suggested we not find Q1, Q3, and IQR since she had already been in the session for over 35 minutes, but she decided to continue. Then she made a boxplot.

Fiona liked making the boxplot because it was challenging. She liked weighing the marbles and making the Marble Plot. She said it was hard for her to figure out what I was drawing when I drew a break in the *x*-axis to accommodate the outlier.

Case 1.1 Session 6

We played with measuring instruments. Fiona weighed the blocks piling them up on the ruler and reading the weight with each added block. I introduced the summation sign, upper case sigma. Then she added the dice and the pencils and the ruler. I showed her the dial caliper and she measured a block, the dice, and the pencil fulcrum. I showed her the Vernier caliper and she quickly put it down in favor of the dial caliper. She remembered how to zero the caliper.

Fiona chose to do the marbles. I made a data collection sheet and she did eight trials with the right hand and recorded the trial number and the marble count. When I spoke of *column headings* on the data collection sheet, she heard “colon” and started placing colons after the trial numbers. I didn’t understand this confusion until I reviewed the video. I invited her to put the blocks on the ruler to find the average of the marble counts:

Tutor: Do you think we could put blocks on the ruler and find where it balances? Maybe that would be your best guess. What do you think? Do you want to try that?

Fiona: Trying to get dirt from out my fingernails. [She is picking at her fingernails] [16:00] Oh that's clay. I got clay in my fingernails. I don't like it. It’s so uncomfortable.

Tutor: What would you think?

Fiona: I think that I don't want to do that. I'm tired. [at +53 minutes]

Tutor: Yeah. You're tired. How about if I do it; and then we'll have a look together. Ok?

I arranged the blocks on the ruler and found the balance point while she watched. After 2 more minutes she said, “Ok. I want to keep doing this. It’s fun.” But it was time to stop.

She liked playing with the marbles, stacking the blocks on the scale, and learning how to use the dial caliper. I asked her if anything was hard for her. She said, “Mmm. It was figuring out why I might have different numbers of here [sweeping the handle of the caliper across the *x*-axis of the Marble Plot] when it's the same hand.” This was especially significant since she is now wondering about variation. Unfortunately, these were the last moments of the last session and we could not pursue further discussion of variation.

Case 1.2: Layla (pseudonym)

Layla is a 6.8-year-old first grade girl. She started the session sitting up straight in her chair, hands folded in her lap, and kicking her right foot. She was more reticent than Fiona or Edgar. We spent a few minutes conversing. For the next 5 minutes, she made shapes with the blocks: a rectangle with a triangle on top, the sun in the shape of an octagon, and a teardrop. We cooperated in putting the blocks on the ruler and finding the statistics. She learned to make the X-plot in about 1.4 minutes. She found the average: 5½. I talked her through the procedure for writing 5½. She seems to have had some previous exposure to fractions. The first Free Form lesson was 21 minutes. The camera angle did not capture her face. Next, she did the Dice roll (22 minutes). She stacked the blocks in two stacks when she removed the blocks to find the median. She found 8 statistics (all except the mode and IQR). I decided to stop after 43 minutes, before finding IQR, because she was getting tired. I wrote the symbols; she wrote the values. The camera angle was poor and showed only the blocks and her hands for most of this lesson. In the wrap-up she said she liked writing, putting the X's in the boxes, balancing the blocks, and rolling the dice. Her favorite part was putting the X's in the boxes. What she found hard was "reorganizing" them [the blocks], lifting the ruler and balancing it on the pencils. She wants to do it again next week. At the end of the session, she was smiling and energetic.

Finding Q1, Q3, and IQR became tedious; the tutor talked too much. Layla did not understand the tutor's wordy explanations.

Case 1.2 Session 2

Layla remembers playing with blocks and balancing them on pencils in the previous session. She chose the Dice Roll over Free Form. She at first thought that rolling a 6 meant to put 6 blocks on the ruler. She did not remember how to do the X-plot. The blocks balanced at 7¼. She learned *mode* easily. She remembered the procedure for finding the median but took the median block off and put it on one of the two stacks. We looked at the shape at the end of the lesson since I had overlooked talking about the shape at the beginning of lesson. She saw a mountain, some stairs, and a face. After 24 minutes, she was getting tired. Then we made a boxplot; this was fatiguing for her. The boxplot could have waited for a future session. But she perked up for the next lesson: Block Weights.

Layla perked up when I showed her how to use the scale. I showed her grams and decimals. I made a data collection sheet and she weighed 15 blocks. Too much time was taken up making the data collection sheet. The transcript shows the tutor talking for as long as 3 minutes with no words from Layla. However, she then got well engaged with weighing the blocks. She wrote the block number and weight for each block in the data collection sheet. Then in the third column, she rounded the weights to the nearest whole number. She arranged the blocks on the ruler according to the rounded weights and found the mode, min, max, R, median, and X-bar. She predicted it would balance on the 6. This would have been a time to have her talk more about her reasoning with respect to the predicted balance point. She needed to be reminded how to find the median, but then found it easily.

Layla liked playing with the blocks, writing the statistics, making the X-plot and the boxplot, taking the blocks on and off the ruler, and stacking the blocks. She said finding the statistics was hard.

Case 1.2 Session 3

Layla remembered balancing the blocks, doing the X-plot, the X with the squiggly line, max, and mode. She chose the Dice Roll to start. She smiled broadly when she got the dice in her hand. She remembered how to make the X-plot and how to find the min and the max. I was hovering too close as she tried to balance the blocks. With a little coaching, she remembered how to write ½. She made a stack 14 blocks high as she removed blocks to find Q3.

I gave her a choice between weighing blocks or marbles. “Marbles!” she said as she smiled and nodded eagerly. I asked, “How many marbles do you want to weigh?” Smiling, she said, “All of em!”. She took 6½ minutes to weigh 19 marbles. While she weighed then, I labeled the *x*-axis of a Marble Plot. She completed the Marble Plot. There was one low outlier from misreading the scale and one high outlier – a 21-gram boulder. The rest of the marbles weighed between 4.4 grams and 5.5 grams. She found the statistics except for the quartiles.

Case 1.2 Session 4

Layla remembered the X-plot and the statistics. We worked on saying the word “statistics.” She rolled the dice and made an X-plot. The shape was like “the music goes like up and down up and down.” She did not remember outliers or mode, min or max. She remembered that R stands for range. The Dice Roll was a review of the last session, but she didn’t remember much.

I made a data collection sheet for the mean absolute deviation (MAD). MAD was equal to 2¾. My explanation of the concept was not clear. What was needed was *demonstration*, not *explanation*.

Layla is not enthusiastic about boxplots:

Tutor: Um. Ok. What do you want to do next? Do you want to do a boxplot? [She nods, then hesitates] Yeah?

Layla: Umm. I don’t know.

Tutor: Do you want to do a boxplot? Or do you want to weigh marbles?

Layla: Weigh marbles [She answers quickly smiling and clapping her pencil against her hand]

She wanted to weigh all of them.

Layla liked most playing with and weighing the marbles and making the Marble Plot. She wanted to do it again.

Case 1.3: Edgar (pseudonym)

Case 1.3 Session 1

Edgar is 7.3-year-old first grade boy. He was guarded in his responses to my initial questions, but he became more animated when he got his hands on the blocks. He started out stacking the blocks. They fell when he added the 12th block. Then he made a triangle. After 4 minutes, I showed him the game of stacking the blocks on the ruler. After his first block design I introduced the rule of aligning the blocks with the numbers on the ruler. He made his first

Free Form arrangement of blocks on the ruler, quickly understood and made his first X-plot, and found the statistics except for R, IQR, and mode. I showed him how to read the ruler between the numbers and how to write ½. This was his first introduction to fractions. I demonstrated how to make a boxplot. He made a second design and found the same 7 statistics. I tried to introduce the dice but he wanted to make another Free Form design. He arranged and re-arranged the blocks multiple times before settling on a design and found the 7 statistics. The video ended after 36 minutes and there was no backup. The worksheet shows that he completed a Dice Roll, made the X-plot, and found the balance point but not the other statistics. He completed several creative structures in which he used the dice and marbles along with the blocks as structural elements (Figure 4.3).

Case 1.3 Session 2

Edgar quickly began arranging the blocks on the ruler. He energetically went through four lessons in 45 minutes. He made three symmetric arrangements one after the other, made the X-plots and found the statistics. Then he made a fourth by rolling the dice. He was exposed to fractions a second time when the balance point on the Dice Roll was 6½. He made the boxplot with only one whisker, but the ends of the box were in the wrong place. I erased the boxplot and demonstrated the correct way to make it. I made a data collection sheet for the Marble Grab, but he decided to make marble races for the rest of the session. He liked stacking the blocks and playing with the dice.

Case 1.3 Session 3

Edgar was on task for about 32 minutes before losing interest. I was slow to recognize his lack of interest and kept on for another 12 minutes trying to engage him in the lessons.

He completed a Free Form block design. Looking at the shape, I introduced him to the word “symmetrical.” He practiced until he could say it fluently. He fidgeted with the dice through some of this first lesson. He completed a Dice Roll then built a 5-tier triangle tower. Then he stacked blocks on the scale. He played the xylophone while I made a data collection sheet (we were in a music room full of instruments). While my attention was on the data collection sheet, he pressed with his full weight on the scale. I chided him for this and explained about the care of measuring instruments.

We then started weighing the marbles, but he soon got off task and played with the marbles making races with the blocks and rolling the marbles down a ramp. His mind seems to switch back and forth between the task and his own play, yet he gets the task done and enjoys doing it. When asked about what he liked about today’s session, he said the marble race.

Case 1.3 Session 4

Edgar started out making a dinosaur with the blocks. I gave him an outline of the lesson telling him that he could choose among options which activities to pursue. I introduced the session by talking about variation and datasets. He chose the Dice Roll. We talked about the shape. I introduced the word *symmetrical*. He chose not to follow my suggestion to make an

X-plot before balancing the blocks so in case they fell off he would be able to put them back in the same positions. The blocks fell off. He remembered how to make the X-plot but did not remember what it was called. The balance point was between 7 and 8 and he remembered how to write 7½. I forgot to ask him about the mode when we considered the shape of the distribution. Associating the mode with the shape rather than with measures of central tendency accomplishes a couple of things. It provides a quantitative measure of shape, and it is simple and easily determined by inspection giving the student an early success in finding the statistics of the dataset. At the end of the lesson, he discovered that the boxplot had only one whisker. We both laughed. He fidgeted with the dice throughout the lesson.

He then made a Free Form distribution with the blocks. I asked if it was symmetrical and we had a lesson in saying the word *symmetrical.* Then we had a similar lesson in learning to say *statistics.* He was determined to say these words correctly. After learning the *statistics,* we had the following exchange:

Edgar: Statistics.

Tutor: Perfect [He giggles] We're playing statistics.

Edgar: We're playing statistics.

Tutor: Yes! Do you like to play statistics? [He examines the dice]. How would you like to play statistics?

Edgar: I like to play statistics because … hmmm. [35:00]

Tutor: Let me come over here. [Tutor moves the camera] Ok. Tell me why. Tell me why you like to play statistics.

Edgar: Statistics. Hm, I don't know.

Tutor: Do you like to play statistics?

Edgar: Yes.

Tutor: Ok. What do you like about it?

Edgar: Hm. I don't know.

Tutor: You don't know. Ok. [36:00] Do you want to play a statistics game with marbles or blocks, and a scale?

Edgar: I want to play sta-tis-tics … statistics with marbles.

Tutor: Ok. Let's play statistics with marbles. There are two kinds of statistics. There are statistics we get from counting, [He yawns] and statistics we get from measuring.

Edgar: Counting. [as he looks up at the tutor.]

Tutor: You want counting? Ok. [The tutor pours the marbles into a bowl. He lights up and leans forward in his chair.]

He fidgeted with the dice and a pencil throughout the lesson.

About 36 minutes into the session we started the Marble Grab. I made a data collection sheet and Edgar performed five trials alternating left and right hands. Twice, he separated out the green marbles then mixed them back in again. He found the min, max, and range and then realized he was tired. About 55 minutes after the start of the session, he was looking and acting tired:

Tutor: Ok. Are you getting tired? [He looks at the tutor] Do you want to stop?

Edgar: I'm not getting tired yet.

Tutor: You're not? Do you want to finish?

Edgar: Yes.

Tutor: You don't have to.

Edgar: I think, I want to stop.

Tutor: You think you want to stop? Ok. It's Ok. We can just stop and you can just play with the marbles if you want.

Edgar: What about the music? [There are musical instruments in the classroom.]

Tutor: Oh. Oh, I wanted to show you just one thing though. We could balance the blocks and find the average.

FIELD NOTE: The tutor is pursuing his learning agenda while Edgar wants to play.

Tutor: Do you want to do that? [18:00]

Edgar: No.

Tutor: No? Ok.

Edgar: I want to play the music right over there.

Tutor: You can play the music if you want.

Edgar: Ok.

We ended the session on a musical note.

Case 1.3 Session 5

In the previous session, we did not have time to find the average, median, Q1, Q3, IQR, and make a boxplot. We began this session with a Marble Grab at the very beginning. I made the data collection sheet beforehand and Edgar performed ten trials with just one hand. I made a grid for the marble plot and he entered the data points. I wrote the symbols for the statistics and he found and wrote the values. I transformed the scale on the ruler writing in pencil and he used the transformed scale to find the average. On finding the average, he threw his hands up in the air in delight. With my help, he read the average on the ruler where I had written the new scale: 13¼. He is getting more familiar with fractions. He made a boxplot and labeled it. During this session, as in the previous session, he separated out the green marbles. After 27 minutes, he decided it was time to play music and, on that note, the session ended.

Case 1.4: Andra (pseudonym)

Case 1.4 Session 1

Andra was a 7.7-year-old second grade girl. She is reticent. Her answers to my questions are mostly one word. She made a symmetrical arrangement of blocks on the ruler, made an X-plot, and found statistics. I showed her how to make a boxplot. She made a second arrangement of blocks similar to the first one. I asked her to make one that was more different. She made the X-plot. When she tried to balance the blocks on the ruler the pencils spread apart. I reached over pinched them together; then the blocks fell off the ruler. My involvement may have hindered more than helped. She found the statistics. I asked her to make a boxplot in the space above the X-plot. My attention was elsewhere as she made another X-plot with a rough, handmade grid. She liked playing with the blocks, writing, and making the grid. It was hard for her to balance the ruler on the pencils.

Case 1.5: Mary (pseudonym)

Case 1.5 Session 1

Mary was a 7.7-year-old second grade girl. The session started late because she was in the office waiting while I was in the classroom waiting. At the beginning of the session, she fidgeted with her fingers. She made the blocks into the shape of a person then went back to fidgeting with her fingers. This continued until I began to explain the block game; she became attentive and stopped fidgeting. When she tried to balance the ruler, the toothpicks separated. She could not read the yellow tape on the ruler because the black numbers underneath showed through. I decided to switch to another activity and deal later with the problems of the toothpicks and the ruler. This was the first one-on-one tutorial session, before switching we started using pencils instead of toothpicks for the fulcrum. Rather than blacken the numbers under the tape, I could have just put the tape on the blank side of the ruler (lessons learned).

I showed Mary how to use the scale and explained grams by showing her labels on food packages (I was improvising). She weighed the 15 blocks and I recorded the weights in the notebook. She found N, min, and max and made a dot plot. I made a number line for a dot plot and she placed the dots, but it proved to be an unsatisfactory visual representation of the distribution. This led to discovering the X-plot as a more effective visual. I introduced her to the Vernier caliper. She said the session was creative and fun. She liked weighing the blocks. It has hard to get the blocks to balance.

Case 1.5 Session 2

Mary remembered from last time weighing the blocks and measuring them with the caliper. She remembered some blocks weighed more than others (this might have been a good time to have a conversation about variation). She put the blocks on the ruler in the shape of a "sideways building." She did her first Dice Roll. Her first roll was a 5 and a 5. She read this as 55. Rather than correcting her, I just showed her another way to read the outcome of the dice roll. The blocks balanced at 7½. This was her first introduction to fractions. In addition to the average, she found min, max, range, median, Q1, and Q3. I showed her how to make a boxplot.

Mary measured the blocks while I wrote the dimensions in a data collection sheet in the field notebook, first in millimeters, then in inches. I made an X-plot grid in the notebook. We worked together to complete the X-plot.

Mary liked rolling the dice and putting the blocks on the ruler. She liked learning about X-plots. Measuring with the Vernier caliper was difficult. Video shut off at 37 minutes after she had measured a few blocks.

Case 1.5 Session 3

Mary remembered balancing the blocks on pencils and she liked using the dial caliper. She fidgeted with her fingers until she started rolling the dice. She made an X-plot and found nine statistics. I again showed her how to make a boxplot as I had the previous week.

She continued the exercise from the first session where she weighed the blocks. I explained how to round numbers and she rounded the weights to whole numbers in the next column. Then she arranged the blocks on the ruler according to the rounded weights, found the balance point (X-bar), N, median, min, max, and R. For her second Dice Roll, she found 10 statistics and made her first boxplot.

She liked finding the mode, rounding the numbers and making the boxplot. Finding the IQR was hard.

Case 1.5 Session 4

Mary remembered from the previous session how the pencils slid apart when she tried to balance the blocks. I gave her the new fulcrum made of two short pencils glued together. She rolled the dice and described the shape of the distribution. She made an X-plot. The average was 8½. I wrote 8½ in the filed notebook and she copied it. This was her third use of fractions. I asked her to talk-aloud as she found the statistics. She found the 10 statistics, made and labeled a boxplot.

I presented her with a choice of what to do next. She chose measuring with the dial caliper. She remembered to zero it.

Tutor: Alright. What do you want to do next?
Mary: Use the caliper?
Tutor: You like the caliper, don't you?
Mary: Um hmm

Tutor: Ok. Let’s think of a way to use the caliper that’s maybe more interesting than just measuring blocks. Which caliper do you like to use? [She reaches for the dial caliper.] That one. The dial caliper?

Mary: Um hm. [She smiles]

Tutor: Ok. Let’s take it out of its package here. [Tutor hands her the caliper.] What’s the first thing you want to do with a measuring instrument?

Mary: Make sure it’s at the zero?

I made a data collection sheet and she measured the diameter of 10 marbles with the dial caliper. She made an X-plot. She sees the shape as the hand sign for a local sports team (a “lobo”). I transformed the marble diameter to fit the scale of the ruler and she found X-bar on the scale of the ruler (6¾) but did not transform it back to the scale of the marble diameters. She found the other statistics except for the quartiles.

Mary liked doing the mode, Q1, Q3, and IQR. She did not find anything difficult.

Case 1.5 Session 5

I asked Mary to present to the camera as she would to the kindergarten student she will be tutoring in her next and final session. She went through the Dice Roll in 22 minutes. She presented her understanding of outliers:

Tutor: Ok? And we looked to see if there are any gaps or any outliers.

Mary: Um hmm

Tutor: Remember the outliers? [24:00]

Mary: Mm. Those were the ones farthest away.

Tutor: Yeah.

Mary: I don't know if 4 was an outlier cause it was only one space from six, seven, eight, nine, ten, eleven, and twelve.

FIELD NOTE: Mary explains her understanding of outliers.

Tutor: Uh huh. Yeah.

Mary: So

Tutor: I think, I think you're right. I don’t think 4 would be an outlier. But if it were way out there at the one, maybe it would be.

Mary: Um hm.

Tutor: Yeah. Or maybe not. [the tutor chuckles] Actually outliers really depend on a lot of other things that we won't talk about today, but you'll learn about later.

Mary: If the ruler was like that big [gesturing with hands]

Tutor: Uh huh.

Mary: An outlier would be like this [gesturing over the ruler] half. But one block would be all the way at the other side of the room.

I show her a marble plot I made of 52 marbles. We look at the shape and find the min, max, range, and median. This activity prompts the following exchange:

Tutor: So, we're doing the same thing that we do with the blocks only with a picture and a pencil and our fingers. Does that make sense?

Mary: And our brains.

Tutor: Yes. And our brains. We couldn't do any of this without our brains. But we also need our hands [38:00]

Mary: Yes.

Tutor: Our hands and our brains and our heads.

Mary: And our bones. Cause if we didn't have our bones, we'd be a squishy piece of skin and then we couldn't pick up a pencil cause our fingers would be all floppy.

Mary does 6 trials with the Marble Grab – left hand and right hand. The activity began with the tutor asking her "Which hand is bigger?" We had the following exchange:

Tutor: Fourteen? So then which hand is bigger?

Mary: This one?

Tutor: The right one? But you said before the left hand was bigger.

Mary: It looks kind of bigger [looking at both hands].

Tutor: Well let's try again. Grab again with the left hand again. We'll try to figure this out.

Mary: 1,2 [6:00] 3,4,5,6,7,8,9,10,11.

Tutor: Eleven? Did your hand shrink?

Mary: Mmm. Mm mm mm [I don't know]

Tutor: Did it? Well grab again with your right hand. We'll see if we can figure out which one is bigger.

Mary: Maybe it's because my right hand keeps grabbing it between the fingers instead of just the hand itself.

FIELD NOTE: Mary gives an explanation for common cause variation.

Tutor: Ok

Mary: 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17.

Tutor: Seventeen! Now can you explain this? It looks like your right hand is growing. 10, 14, 17. [7:00] Do you think it's just gonna keep growing?

Mary: I don't know.

Tutor: What's going on?

Mary: That's what I wanna know [both laugh]

Mary learned to be more consistent by the 5th and 6th trials.

Mary liked the Marble grab, making pictures with the blocks, balancing the blocks, and making the X-plot and the … [pause] "I think it starts with an "s" … statistics

Case 1.6: Jacob (pseudonym)

Case 1.6 Session 1

Jacob was a 7.9-year-old second grade boy. After a few minutes building rapport, he played with the blocks. He made a rocket. He made an ant's face, then an alien's face (a reference to Minecraft). Then he made Kool Aid Man and then a human, then a zombie, then Mickey Mouse. He was imagining being in a Minecraft game. Then he made Hairbrain, a character from a game called *Crossing the Road*. The tutor transitioned from Jacob's imaginary play to the blocks on the ruler:

Jacob This is a Minecraft character he has ten teeth and he's gonna walk he's gonna put it here and then he's gonna make a line with Redstone. Very far away. And then he's gonna make this wall.

Tutor: Uh huh.

Jacob [inaudible] And then, the lever, when it turns like this, it explodes. Dit dit dit dit [He lifts blocks over his head as if in an explosion.]

Tutor: You're getting a lot of uh, a lot of mileage out of these blocks? Can I show you another way to play with them? [He nods] Ok. This is uh … [16:00] Let's see. Let me get the right ruler here. Yeah, this one. Okay, I'm gonna try something a little different today. Can I see the blocks for a minute? [Jacob slides the blocks over to me.]

Jacob: I made some stairs.

Tutor: Stairs. Yeah. Now what I'm gonna ask you to do is to build something on the ruler. And here are the rules. The block has to have a number right in the middle. See that five is kind of in the middle of the block?

Jacob and I work together to put the blocks on the ruler. I showed him how to balance it. He read the balance point of the ruler. Then he made up his own rules and balanced the ruler on a block and put a pencil on either end of the ruler. We made a block design together and he shaded in the boxes on the worksheet. He went over the lines and this led to the idea of the X-plot. At 28 minutes, he seemed to be getting tired. He found the balance point, min, max, median, Q1, and Q3. I showed him how to make a boxplot.

Jacob rolled the dice for his first Dice Roll lesson. The camera shut off before he finished rolling the dice. There was no back up. The worksheet shows that he shaded the boxes but did not find any statistics.

The debriefing questions showed that he liked the blocks, "learning about grams and stuff". I asked, "What was your favorite part?" he said, "Meeting you." It was hard to get the blocks to balance.

Case 1.7: Carla (pseudonym)

Case 1.7 Session 1

Carla is an 8.1-year-old second grade girl. After making a few block structures we got right into the first Dice Roll. I demonstrated rolling the dice and putting the blocks on the ruler. I demonstrated the X-plot and she did it easily. I coached her through finding all 10 statistics then demonstrated the boxplot. The blocks fell off twice before she found the balance point.

Carla did a second Dice Roll. She made the X-plot and found N. Finding the balance point, the blocks fell off once. The camera shut off soon after she found the balance point. She found the 10 statistics. There was not time for the boxplot. There was no time at the end of the session for a debrief.

Case 1.7 Session2

Carla did the Dice Roll. She described the shape as a car, a whale, and a “dolphin shooting out water from its head.” She made the X-plot, found 10 statistics, and made a boxplot. The blocks fell off twice as she was finding the balance point. I coached her through the boxplot. After finishing the Dice Roll, about 26 minutes into the session, we had a light conversation about snacks, erasers, gummy bears, and books she bought at the book fair.

I made a data collection sheet for 10 blocks. I showed her how to use the scale and showed her about grams. She weighed 10 blocks and wanted to keep going and weigh them all. She extended the columns, added the numbers 11-15 and weighed all the blocks. Then she weighed her Big Fat Gummy Bear eraser. It’s remarkable how engaging weighing blocks and marbles and recording their weights is for the students.

Carla’s favorite part was weighing the blocks. She also liked balancing the blocks although she found it hard.

Case 1.7 Session 3

Carla did the Dice Roll, balanced the blocks easily, found 10 statistics, made and labeled a boxplot.

I brought a printed 4-column data collection sheet. I labeled the columns and she entered the block number, weighed the blocks and recorded the weight. I showed her how to round the weights to the nearest gram. She rounded the numbers and I wrote them in the data collection sheet. The third column was for rounded weights.

I labeled the column headings of a second data collection sheet for the Marble Grab. She performed 5 trials comparing left and right hands. Then I gave her a verbose and clumsy explanation of variation.

Pairs of Students

Case 2.1: Second grade girl and second grade boy, 1 session

This first session of the study with Victor and Andra did not produce useful data. I intended to begin the study with one-on-one sessions. However, when two students came due to a miscommunication, I decided to proceed with both. The dynamics of the interactions were not conducive to a productive session.

Case 2.2: Two second grade girls, 1 session

Carla showed Andra the Dice Roll. Carla had already had three sessions and Andra had one previous session where she did a Free Form block lesson. Andra was also in the first session with Victor. They took turns rolling the dice and placing the blocks. Blocks fell off as they moved it. Carla did not know what *vertical* meant. They took turns putting X's in the boxes to make the X-plot. The shape reminded Carla of a hammer and reminded Andra of Minecraft. They worked together to find the statistics. Carla remembered what the mode was. When Carla picked up the ruler to balance it, the blocks fell off. When Andra adjusted the ruler on the pencils, the blocks fell off. Carla did not remember how to make fractions. Before finding the quartiles, we took a break and talked about Field Day. They seemed tired and I ask if they wanted to stop:

Tutor: Are you guys tired of this? [28:00] Or do you want to finish it?

Carla: Finish it?

Andra: I want to finish it.

It takes 10 minutes to find the quartiles and make a boxplot. That's a long time. And the activity becomes tedious. We take a break and eat some golden berries and talk about food.

I showed Carla the worksheet from our previous session. It has block weights and rounded block weights. Carla showed Andra how to set up the scale and weigh blocks. I coached Carla to put the blocks on the ruler according to the distribution of block weights from the previous session. The blocks are on 10, 11, and 12, with one on 9. I asked her to predict where it would balance:

Tutor: There you go. Yeah, you can work together. Where do you think it would balance?

Carla: Right here. [Carla gestures toward the middle of the ruler around 6.]

Tutor: Why would it balance there?

Carla: Like right here or something [Carla points to near the 9].

Tutor: Why would it balance there?

Carla: I don't know. I'm just guessing.

They worked together to balance the blocks. They found the balance point. The blocks fell off.

Both Andra and Carla said they had fun. Andra liked making the X-plot. Carla said she liked everything. Carla said it was hard to get the blocks to balance. Andra said Q1, Q3, and IQR were hard.

Case 2.3: Second grade girl tutors kindergarten girl, 1 session

Mary showed Olga (kindergarten) the Dice Roll and the Marble Grab. The Dice Roll lesson took just under 19 minutes. It included this exchange where Mary showed Olga fractions:

Mary: Ok. Now it balances. So, it's not quite at the seven, but not quite at the eight. So, you call that seven and a half. And then [7:00] you would write 7½. Right [Mary grabs the worksheet and poises the pencil above the worksheet.] So, you would write 15 right there [N=15]. And then you would write 7½ right [Mary studies the list of statistical variables then turns to the tutor] I think right there? [tentatively pointing to the X-bar]

Tutor: Yeah.

Mary: Ok. So, you would write 7½ right here. And you write 7½ like this. So, you could either go like that, or you could go like this. Ok?

Olga: Ok

Mary: So why don't you try writing it [Mary erases what she wrote.]

Olga: Ok. [Olga writes 7½]

Next, Mary decided to do the Marble Grab. She took a blank sheet of lined paper and made a data collection sheet. Olga grabbed the marbles and Mary recorded the data.

Olga	
LH	RH
1. 13	1. 8
2. 10	2. 13
3. 10	3. 12
4. 10	4. 10
5. 11	5. 15

After they completed the data collection, I sat between them to debrief what they had done. I showed them how a 3-column data collection sheet allows them to have a separate column for the trial number and only one piece of data in each box.

Trial	LH	RH
1	13	8
2	10	13
3	10	12
4	10	10
5	11	15

I then wrote the symbols for the statistics and for each hand they found N, min, max, R, and median. To find the average, I wrote in pencil on the yellow tape on the ruler a scale from 7 to 16. Then they found the average. For the sake of the exercise, they pooled the marble counts for the left and right hands to make a dataset of 10 counts. They placed 10 blocks on the ruler and found the average where the ruler balanced at 12½.

Olga said the marbles and the blocks were fun. Mary liked the marbles and making and using the data collection sheet. Mary said the boxplot was hard.

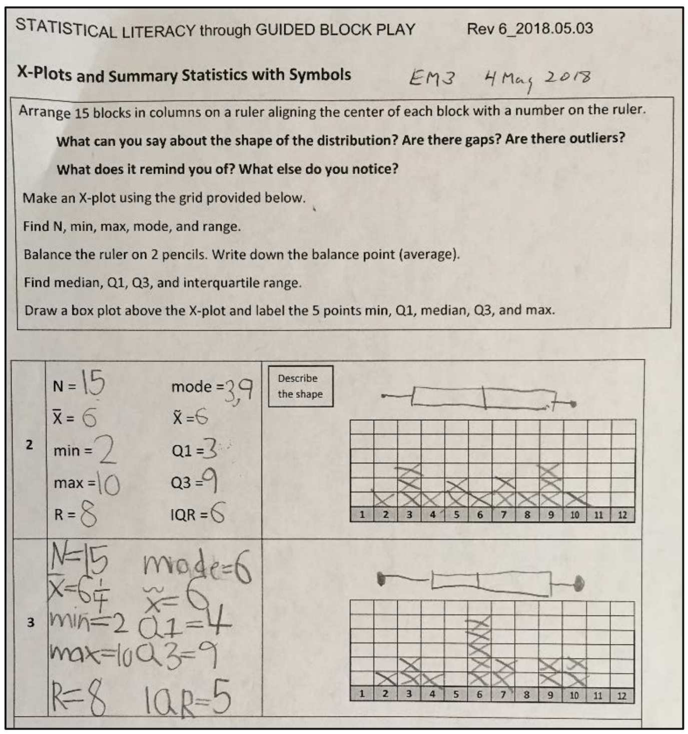
Appendix B: Interaction Analysis Transcripts

Interaction Analysis Session 1

8 Aug 2018

Adult observer participants: 3 elementary Montessori teachers; one PhD psychologist.

Participants observed video EM3 4May18 001. They observed the Dice Roll lesson beginning +17:00 minutes and ending +32:00 minutes. The worksheet for this exercise is shown in Figure B.1. This was the third session for this first-grade boy. He first created the symmetric block design shown as Lesson 2, made the X-plot, and found the statistics. Then he copied the statistical symbols into the corresponding area for the Lesson 3. Adult observer participants were given a copy of the completed worksheet and the transcript for reference during the Interaction Analysis session.


STATISTICAL LITERACY through GUIDED BLOCK PLAY Rev 6_2018.05.03

X-Plots and Summary Statistics with Symbols EM3 4 May 2018

Arrange 15 blocks in columns on a ruler aligning the center of each block with a number on the ruler.

What can you say about the shape of the distribution? Are there gaps? Are there outliers?

What does it remind you of? What else do you notice?

Make an X-plot using the grid provided below.

Find N, min, max, mode, and range.

Balance the ruler on 2 pencils. Write down the balance point (average).

Find median, Q1, Q3, and interquartile range.

Draw a box plot above the X-plot and label the 5 points min, Q1, median, Q3, and max.

2	N = 15 mode = 3,9 $\bar{X}$ = 6 $\tilde{X}$ = 6 min = 2 Q1 = 3 max = 10 Q3 = 9 R = 8 IQR = 6	Describe the shape 1 2 3 4 5 6 7 8 9 10 11 12
3	N=15 mode=6 $\bar{X}$=6 1/4 $\tilde{x}$=6 min=2 Q1=4 max=10 Q3=9 R=8 IQR=5	1 2 3 4 5 6 7 8 9 10 11 12

Figure B.1. This worksheet is from Edgar's third session. He made a Free Form symmetric arrangement of blocks, an X-plot, found the statistics, and made a boxplot. On his own initiative in the second lesson, he copied the symbols before doing the Dice Roll.

Participant Feedback. Edgar is not learning independently. The lesson did not match his developmental level of reasoning. Although he could find statistics under the direction of an adult, they had no meaning for him. He lacked prior knowledge; he had no reference or context. He had no prior knowledge of fractions. The value of this lesson at this stage of development is questionable. However, it might be valuable in 4th or 5th grade with students working in small groups. The lesson is not sufficiently user-friendly to engage a younger child naturally, spontaneously, and autonomously so that he develops mastery. The child needs to be two or three years older before his reasoning mind matches the level of the lesson. Ideas like mode and range he can see, but median, mean, Q1, and Q3 are too abstract. He can see the shape, gaps or spaces in the arrangement of blocks, but the boxplot is too abstract to hold any meaning for him.

Histograms are found in the curriculum in 3rd and 4th grade. This lesson might work with children as young as third grade. 5th grade is a good year for surveys and for introducing technical terms. 6th grade for box and whisker plots. The lesson goes too quickly to a level of abstraction beyond the reach of the first or second grader. The level of complexity and abstraction should increase more gradually. It's important to track age by the month, not just the year. Drawing the blocks, rather than doing an X-plot is more concrete and might have more meaning for younger learners.

It would be interesting to do a longitudinal study to see if there is an effect on attitudes toward statistics later in life.

Interaction Analysis Session 2

19 Aug 2018

Adult observer participants: 2 retired elementary teachers and one retired specialist in early childhood education.

The researcher (R) presented an overview of the study and the research questions. Participants observed the first 34 minutes of video FB3 24Apr18 001. The worksheet for this exercise is shown in Figure B.2. This was Fiona's third session.

Participants P1, P2, and P3 observed the student create a design of blocks on the ruler, balance the blocks on the ruler, find the median, Q1, Q3, and make a boxplot. The following excerpts from the transcript illustrate the perspectives of these observers.

P2 … although you did say play, she said she liked to work, and I think it's good for them to know that that's their work. [48:00] That playing is their work. … Because there's a purpose in it. And I think that's really important, for kids to know why they're doing something. … I was surprised that she didn't put the blocks evenly on the ruler so the ends were the same.

P2 Well, some adults think that when children are not doing something with a pencil and a paper that they're not learning. [55:00] And other people think that the more the kids manipulate objects, and in order to kind of accomplish something, even if the kid doesn't know what it is, if the teacher knows and sees them and gives them a little guidance and sees that they can do that. … This is work to her. But it looks like play. And it's just so important for kids to learn that they are learning when they're playing. I like to call it work too. I often use that word. It's time to do our work now, not just when they're using a pencil and paper, even when they're relating to each other.

P1 Even in preschool we use that language too: “It's time to do your work.” Then we'll go play or go outside or whatever, but we call it work. [56:00]

P3 But their work really is, I mean, it's what adults think looks like play.

P2 She’s getting tired [59:38] [she has been in the session for 30 minutes]

P1 She’s fatigued.

P2 She’s getting tired. She wants to get it right. She knows what she's doing. Even when you have to remind her, she remembers. But I definitely see fatigue. …

P3 How long has it been going?

R At this point, we've been at this for a little over 30 minutes.

P1 Oh, that’s a long time.

R Maybe 32 minutes.

P2 That’s unusual to me. … And so, to me, she's amazing.

P1 Well, remember, he's giving her that one-on-one. If you were to do this with the whole class, it would be almost impossible. And then you’d have the kid that's going bonkers. It would have to be a small group instruction. And maybe at a table with maybe three or four kids …

P1 I wonder if there's a way you can do this, do the block activity without using the vocabulary, without doing the recording, and just letting them manipulate and play. Say, can you just find the balance. Where does it balance it? What number or just you know, baby step it into you know in kindergarten. In first grade, maybe start doing some documentation maybe in second grade, you know kind of build on it. I think it's a lot to ask a kindergartner to do this kind of problem solving with this kind of vocabulary and I just don't think it's developmentally appropriate for kids to be doing this kind of activity at this young of an age. But there are things you can do to build up to it, to give that concrete experience, and like you say, scaffold it and sequence it up into the upper grades. I know, I think it starts in the second or third grade where they're doing a lot more with the median and the mode and all that stuff, but it's not until you get … past early childhood, basically through second grade. But give them the experiences of variables and different things without the actual plotting it and …

R That was 34 minutes.

P2 She did really great.

P1 Yeah. I’m gonna tell you: she’s a teacher’s dream. [laughter]

P2 She is.

P1 “I like doing worksheets” [laughter] [quoting Fiona]

P2 What I found … [1:13:00] It doesn't address this specific thing, but I found that when working with the little kids, that repetition is not boring to them. And if you found that the repetition of doing something is not boring to them because it becomes so that they could say “Oh, this is easy!”

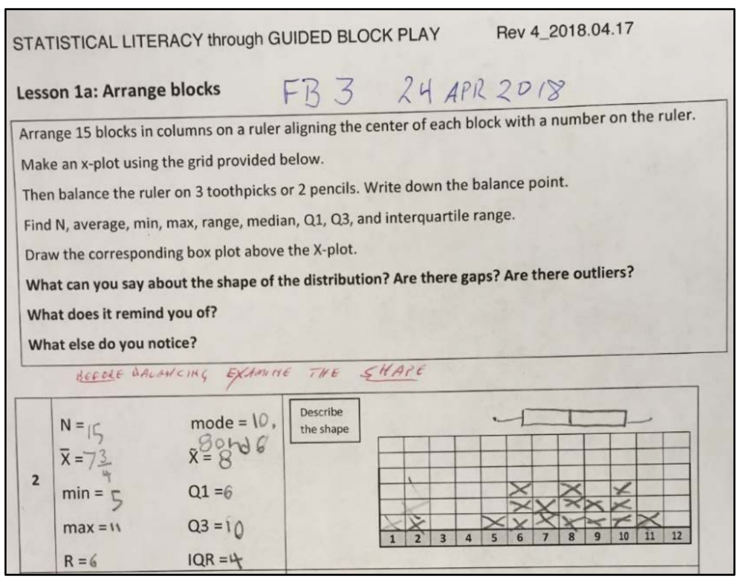
P1 Right?

P2 And I think … by the end of all the things you do, it would be second nature to her. When I was teaching little kids, every morning I would repeat what we did the day before. Five minutes, this is what we did yesterday. Do you remember? “Yes, Yes, Yes.” Or “No, No, No.” But so many wouldn’t remember.

R [We’re targeting] two key concepts: … the idea of distribution and the idea of variation. … one of the questions is [1:15:00] How well is this preparing a child, and is it … [appropriate] for the age or the development level. How do you see that?

P1 This activity does not, it's not part of her world. It doesn't make sense. She can't transfer this to another activity. It's pretty much isolated to the activity you're doing with her. It's not something that's part of her world. So, it has no meaning to her at this point. She's just memorizing these, you know, mode, median, Q1. But it has no real meaning or relevance to her. So that's my only concern with an activity like this at such a young age. Instead of doing the worksheet and doing the questioning and learning these principles, just doing the foundational work, the concrete work that will lead up to later on in second or third grade when you start introducing mode and median. They're like, “Oh, yeah! I remember back when we played with blocks and we were trying to find … it makes sense, because they've had that experiential base to build on.

R But that's the question: Does it prepare her for understanding that later on?


STATISTICAL LITERACY through GUIDED BLOCK PLAY Rev 4_2018.04.17

Lesson 1a: Arrange blocks FB 3 24 APR 2018

Arrange 15 blocks in columns on a ruler aligning the center of each block with a number on the ruler.
Make an x-plot using the grid provided below.
Then balance the ruler on 3 toothpicks or 2 pencils. Write down the balance point.
Find N, average, min, max, range, median, Q1, Q3, and interquartile range.
Draw the corresponding box plot above the X-plot.
What can you say about the shape of the distribution? Are there gaps? Are there outliers?
What does it remind you of?
What else do you notice?

BEFORE BALANCING EXAMINE THE SHAPE

2

N = 15 mode = 10, 8 or 6
$\bar{X}$ = 7 3/4 $\tilde{X}$ = 8
min = 5 Q1 = 6
max = 11 Q3 = 10
R = 6 IQR = 4

Describe the shape

Figure B.2. Fiona arranged the blocks on the ruler, made the X-plot, and found the statistics. The tutor made the box of the boxplot and Fiona made the whiskers.

P1 You know, one of the things that we learned in our programs with Early Childhood is: just because they can do it doesn't mean they should be doing it. So, you could take these baby Einstein's and you can show the flash cards of presidents or states and they can memorize it and you think, “Oh my God so smart.” But is that what a toddler is supposed to be doing? Memorizing states or facts? They should be playing in the mud, and getting their hands in the dirt and building …

P3 And playing with blocks.

P1 Playing with blocks and … so just because they can do it doesn't mean they should be doing it or that is appropriate. But once they reach that stage where they can go from the concrete to the symbolic, that doesn't happen until the second grade or third grade where they really … I know some countries they're not even introducing, letters and numbers to kids until they're like first or second grade. Way past what we do here in the United States because they're finding that they're just not ready for that symbolic representation of concepts; that they still need to be doing the concrete, a lot of manipulating objects, and experimenting, and problem solving. And then they introduced letters and sounds and numbers and it just clicks with them because it reaches them. You know, it's birth through age 8 is early childhood and they're still in a world of fantasy, they still believe in Santa Claus and princesses and fairies and castles, yeah magical thinking. So, to them, that's their world of make-believe, of fun and play. And then once that is left behind, it's kind of sad. It's like, "Oh my God, they don't like the tooth fairy anymore and Santa Claus" … because now they're into a more symbolic stage. They can go on to those more abstract concepts. [1:18:10]

P1 If you get those balance scales? I know they have a ball with those. How many teddy bears does an apple weigh? They love doing things like that.

R So, what I'm hearing is that they can begin to engage these concepts variation and distribution.

P1 Absolutely.

R But maybe not boxplots and quartiles.

P1 Right.

R Something short of that.

P2 I don't think it hurts them, but [1:21:00] I think it's good to hear the words, that they hear the words. I don't think that's a problem. But, if you apply these things to a regular setting, for instance, if you're going to do this in a classroom, it would be really hard.

P2 I myself do not feel comfortable even with the terminology, you know, this is all new kind of … it's a new language to me that I haven't been teaching.

R What about adults playing with blocks?

P3 That would be good.

P2 Yeah, I think so too.

P3 I'd like to do that. [1:32:00]

Reflections

There is a unity of play and work at this age. What looks like play to adults is their work.

How much could they do just playing with the materials by themselves?

Even if this is not developmentally appropriate, what could be done to build them up to it?

Maybe more repetition should be built into the activity.

There may be a tendency for teachers to associate what they see with the grade level at which some of these concepts are introduced in the standard curriculum.

How much of their experience and understanding can transfer and contribute to their growing understanding of data and variation? What meaning does the activity have for them?

Two games were mentioned by P2: Pancake and Water Drops on a Penny. There was a suggestion to use the balance scales. Could these be incorporated into these lessons?

There is a concern that at least part of the activity, such as quartiles and boxplots, may be beyond what is appropriate for this age. This material may be more appropriate for older students.

There may be a tendency for teachers to rely on what they’re used to or comfortable with or what they’ve read in their professional sources. This approach is unfamiliar. In fact, the field of statistics is unfamiliar to many teachers.

Appendix C: Notes on the Redesign of Learning Activities

Balance Blocks Model of Performance: Developing Knowledge and Skills of Statistical Inquiry

Performance → Ability → Component → Step → Skill / Knowledge

Course → Module→ Lesson → Learning Interaction → Teaching Point

Target Population: K-5 with ability to

- Count to twenty
- Recognize and write numerals from one to twenty
- Recognize and write letters
- Read a ruler

Design Innovations from the study

- X-plot
- Marble bowl
- Marble plot
- Two-pencil fulcrum
- Yellow tape ruler
- Revised worksheets

Re-sequence of learning interactions suggested by the findings

- First do a one-handed marble count activity (N=11) twice with each hand then do a two-handed marble count exercise and compare boxplots.

Proposed innovations

- Design an embodied boxplot: telescoping box, sliding whiskers, and sliding median
- Design an embodied MAD method using length (maybe spaghetti)
- Revise MAD method to account for zeros
- A large group exercise to find confidence intervals using dice roll. Form small groups.
    - One group rolls the dice 5 times; use blocks to find the average; repeat as many times as time allows.
    - One group rolls the dice 15 times; use blocks to find average; repeat.
    - One group rolls the dice 30 times; use blocks to find average; repeat.
    - Compare the spread of sample averages among the three groups.
    - Analyze the system and compare empirical findings to the theoretical distribution
    - Introduce relevant principles of probability

End Notes

[i] Wild & Pfannkuch (1999) trace the roots of statistical thinking to John Graunt (1620-1674). Historically, data were simply recorded and stored. But Graunt saw how to think and reason with data in a process of inquiry. His insights enabled him to investigate the spread of the plague and to estimate the population of London based on knowledge of the birth rate. Graunt's work reflected a deeper shift in thinking with respect to the nature of knowledge and of evidence – a shift away from arbitrary authority toward empirical observation and procedures that could be replicated by any observer, building on foundations laid by Copernicus, Galileo (Einstein & Infeld, 1938), Kepler, Francis Bacon (Eisley, 1962), and others. This represented a shift from an absolutist to a probabilistic view, from determinism to indeterminism with respect to the natural world and social conditions. Quetelet (1796-1874) discovered the significance of patterns of variation in aggregate phenomena. He discovered that these patterns could be modeled by theoretical distributions and provide the basis for probabilistic predictions of the future. William Playfair (1759-1823) developed visual representations of data that had traditionally been contained in tables but could now be used as tools for thinking and communicating – an innovation that was put to good use by Florence Nightingale (Figure 1.2) and regrettably, *not* used in the communications leading up to the Challenger disaster (Figure 1.3). Galton discovered the principle of regression to the mean – given a stable pattern of "common cause variation" (Deming, 1993)**,** extreme values are followed by less extreme values giving rise to the illusion that action taken in response to these values *caused* the movement back toward the mean value.

[ii] In 1858, Florence Nightingale became the first woman elected to the Royal Statistical Society and later became an honorary member of the American Statistical Association.

[iii] *Transforming Our World: the 2030 Agenda for Sustainable Development* (United Nations General Assembly, 2015) is a plan of global action that seeks to strengthen universal peace, prosperity and freedom in collaborative partnerships between all nations and peoples. It is a continuation of the Millennium Development Goals (MDGs) that guided UN action from 2000 to 2015. This plan of action puts forth a global vision of economic, environmental, and social justice where life can thrive free of hunger and want, poverty and disease, fear and violence. The goals of the plan include: significant progress in securing gender equity, racial equality, the rights of Indigenous peoples, and the care and protection of children. The plan further calls for transparency and accountability in governance at all levels; educating people for full participation in society; a well-educated workforce; sustainable patterns of production and consumption; sustainable industrial development; and strengthening scientific, technological, and innovative capabilities, especially in the so-called "least developed countries." This far-reaching global vision calls for thinking in terms of decades and generations. It calls for statistically sound indicators to track progress toward these Sustainable Development Goals (SDGs). These indicators will require consistent, valid, and representative data to establish baselines and serve as a foundation for data-driven decisions leading to measurable improvements. The plan calls for "strengthening data collection and capacity-building in Member States … to better inform the measurement process" (p. 13). To enact effective follow-up, the plan calls for stronger evaluation programmes and for "strengthening the capacity of national statistical offices and data systems to ensure access to high-quality, timely, reliable and disaggregated data" (p. 32). Developing and implementing meaningful measures of social, environmental, and economic processes calls for collaboration, transparency, and statistical literacy (UNESCO, 2006; United Nations Economic Commission for Europe, 2012)**.**

[iv] What are the "Big Ideas" or fundamental concepts of statistics? (J. Watson, Fitzallen, Fielding-Wells, & Madden, 2018). Ben-Zvi and Garfield (2004, 2008) presented eight: data, distribution, trend, variability, models, association, samples and sampling, inference, and comparing groups. Crites & St. Laurent (2015) present five Big Ideas that summarize 24 essential understandings expected of students in grades 9-12: (1) data consist of structure and variability; (2) distributions describe variability; (3) hypothesis tests answer the question, "Do I think that this could have happened by chance?"; (4) The way in which data are collected matters; (5) evaluating an estimator involves considering bias, precision, and the sampling method (pp. 127-128). There is general agreement that variation and distribution are perhaps the most fundamental of statistical ideas.

[v] Reasoning about variation is foundational to statistical literacy (Biehler *et al.*, 2018, p. 145). Garfield and Ben-Zvi (2005) proposed seven facets of understanding variation. Building on their work, Reading and Reid (2010) proposed nine. Peters (2011) proposed three perspectives on "robust understanding of variation" Bakker & Gravemeijer (2004) concluded that "distribution is the central concept for thinking about variability" (p. 155). Wild (2006) saw distribution as the lens through which to view variation (p. 11). Shaughnessy (2007) summarized a trajectory of conceptions of variability and found evidence of distributional reasoning (Noll & Shaughnessy, 2012). Ben-Zvi (2004) called it "global perception of a distribution" (p. 160). Reading and Reid (2006) found that distributional reasoning depends on understanding variation. Several researchers have provided evidence for conceptual trajectories of student reasoning about distribution: Canada (2006), Ciancetta (2007), and Noll (2011). The authors suggest that design experiments might elucidate "hypothesized reasoning transitions about distributions into their teaching, to build teaching-learning trajectories for the classroom and test them" (p. 162).

[vi] Methods of *lean manufacturing* adopted from the Toyota Production System define a current state and a desired future state then employ tools and methods such as value stream mapping and cycle time reduction to plot a course toward a more robust, more efficient, less expensive operation.

[vii] This question of transfer and "deep learning" is the subject of an extensive study by the National Research Council (U.S.) (2012), *Education for life and work: Developing transferable knowledge and skills in the 21st century*.

[viii] Csikszentmihalyi identified the following conditions of the flow experience:

1. Goals are clear – one knows at every moment what one wants to do.
2. Feedback is immediate – One knows at every moment how well one is doing.
3. Skills match challenges – The opportunities for action in the environment are in balance with the person's ability to act.
4. Concentration is deep – Attention is focused on the task at hand.
5. Problems are forgotten – Irrelevant stimuli are excluded from consciousness.
6. Control is possible – In principle, success is in one's hands.
7. Self-consciousness disappears – One has the sense of transcending the limits of one's ego.
8. The sense of time is altered – Usually it seems to pass much faster.
9. The experience becomes autotelic – It is worth having for its own sake. (Csikszentmihalyi, 2014, p. 133)

[ix] Deming (1986) also advocated elimination of rewards and punishments in education. This practice carries over to the workplace where merit pay and performance appraisals rob workers of their right to pride in workmanship and isolate the individual from the larger system to the detriment of both. Deming claimed that the effects of performance appraisals and merit ratings are devastating. They promote short-term performance at the expense of long-term planning; they cultivate fear, rivalries, and politics. "Merit rating rewards people that do well in the system. It does not reward attempts to improve the system" (p. 102). As a predictor of performance, it is meaningless except where people fall outside the variability attributable to the system. Traditional appraisal systems increase the variability of human performance, increase turnover, and diminish the overall stability of the system. This was one of the lessons of the Red Bead Game (The Deming Institute, 1980).

[x] Tukey (1962) referred to the mechanical application of statistical techniques as commonly taught in basic statistics courses as "cookbookery." He was instrumental in making exploratory data analysis (EDA) and statistical thinking more prominent features of a first course in statistics.

[xi] Wilkinson (1999) identifies specifically where the problems are and makes the following exhortations: make clear at the outset the type of study you're doing; clearly define the population; describe the sampling procedure, including inclusion/exclusion criteria and rationale for stratification; describe how random assignment was achieved; where random assignment is not feasible, explicitly state, test and justify assumptions

about the effects of covariates; describe methods used to attenuate sources of bias; explicitly define variables, how they relate to the goals of the study, and how they were measured; summarize psychometric properties of instruments in relation to the specific context sufficiently to allow replication; indicate how attrition may have affected generalizability; provide the rationale for sample size decision; document effect sizes; spell out sampling and measurement assumptions; describe results using confidence intervals; report complications, protocol violations, and other unanticipated events in data collection and describe how your analysis took these into account; inspect data graphically; do not choose a complex analytic technique when a simpler one will do; understand how your chosen software computes and don't let the software shape your thinking; assess underlying assumptions and examine residuals graphically; always provide effect size when reporting p-value; place effect sizes in a practical and theoretical context; provide interval estimates for effect sizes; provide external support for claims of causality, especially in nonrandom designs; provide both tables and figures including graphical representations of interval estimates and the shape of the dataset; include credibility, generalizability, and robustness in the interpretation of the data. In summary, there is no substitute for thought and understanding of context; we need to be transparent with respect to our assumptions; and graphical methods are useful for both presenting data and for testing our assumptions about it.

[xii] Prior knowledge plays a central role in comprehension; when prior knowledge is incorrect, it can be extremely difficult to correct (Bransford & Johnson, 1972). True learning must produce "activatable" knowledge for future use. Correct prior knowledge at an early age can help avoid many of the conceptual difficulties commonly found in students who struggle (and often fail) to grasp statistical concepts. The goal of the teacher is to move the student along a continuum toward a stage of self-directed learning where the student becomes an autonomous, life-long learner; this is the most important outcome of formal education (Dewey, 1938a; Grow, 1991). However, statistics education is notorious for engendering life-long anxiety rather than life-long learning.

[xiii] She was the daughter of the tutor/researcher.

[xiv] ISO 9000 is an international standard for the design and operation of a management system. It incorporates many of the ideas Deming taught such as focus on the customer, statistical methods to improve processes, and employee engagement.

[xv] The entire *Da Xue* is a single page of Chinese characters. It predates Confucius and goes as follows:

> *Wishing to bring order to the world, the ancient Sages first sought to govern their States well.*
> *Wishing to govern their States well, they first sought to regulate their families.*
> *Wishing to regulate their families, they first sought to cultivate themselves.*
> *Wishing to cultivate themselves, they first sought to clarify their thinking.*
> *Wishing to clarify their thinking, they first sought to extend their knowledge.*
> *Wishing to extend their knowledge, they engaged in the investigation of things.*
>
> *Things being investigated, knowledge was extended.*
> *Their knowledge being extended, their thinking became clear.*
> *Their thinking being clear, their hearts were rectified.*
> *Their hearts being rectified, their persons were cultivated.*
> *Their persons being cultivated, their families were regulated.*
> *Their families being regulated, their States were rightly governed.*
> *Their States being rightly governed, the entire world was at peace.*

[xvi] The Laplacean fallacy states, "We may regard the present state of the universe as the effect of its past and the cause of its future. An intellect which at a certain moment would know all forces that set nature in motion, and all positions of all items of which nature is composed, if this intellect were also vast enough to submit these data to analysis, it would embrace in a single formula the movements of the greatest bodies of the universe and those

of the tiniest atom; for such an intellect nothing would be uncertain and the future just like the past would be present before its eyes." (Pierre-Simon Laplace, *Essai philosophique sur les probabilités*, 1814)

[xvii] The learners are developing facility with the mechanics of making data displays. Students commonly struggle with simultaneously learning the mechanics and the concepts. This usually leads to focus on the visible output – the data display – and this display is commonly seen as a static representation rather than as a tool to aid statistical reasoning (Garfield & Ben Zvi, 2007). By learning the mechanics as a game without the expectation of immediate conceptual understanding, the mechanical skill can develop and later serve as a foundation for conceptual understanding. In gaining direct embodied experience with the elements (data points) and the aggregate (frequency distribution), numerical, spatial, and perceptual relationships become clear. The shape, spread, and center of a dataset as three dimensions of an aggregate together become a tangible representation of variation.

[xviii] At a macro level, the conceptual edifice of statistical literacy ties in to other areas of science and might help to illuminate some of the bigger questions in the philosophy of science. diSessa contrasts Kuhn's notion of shifting incommensurable paradigms with Toulmin's challenge to presumptions of coherence within paradigms. Toulmin sees the assumptions of logico-mathematical coherence as unfounded. He suggests that we consider the content of natural science as a conceptual aggregate with only "localized pockets of logical systematicity" (Toulmin, 1972, p. 128). Perhaps statistics could be viewed in a similar light. Toulmin recommends replacing the "snapshot" account with an historical, "moving picture" account of conceptual change.