How Students Learn Statistics Revisited: A Current Review of Research on Teaching and Learning Statistics

ARTICLE *in* INTERNATIONAL STATISTICAL REVIEW · DECEMBER 2007

Impact Factor: 1.19 · DOI: 10.1111/j.1751-5823.2007.00029.x · Source: RePEc

2 AUTHORS:

Joan Garfield
University of Minnesota Twin Cities

Dani Ben-Zvi
University of Haifa

International Statistical Review (2007), 75, 3, 372–396 doi:10.1111/j.1751-5823.2007.00029.x

How Students Learn Statistics Revisited: A Current Review of Research on Teaching and Learning Statistics

Joan Garfield[1] and Dani Ben-Zvi[2]

[1]*University of Minnesota, USA*
[2]*University of Haifa, Israel*

Summary

This paper provides an overview of current research on teaching and learning statistics, summarizing studies that have been conducted by researchers from different disciplines and focused on students at all levels. The review is organized by general research questions addressed, and suggests what can be learned from the results of each of these questions. The implications of the research are described in terms of eight principles for learning statistics from Garfield (1995) which are revisited in the light of results from current studies.

Key words: Statistics education; statistical reasoning; teaching and learning statistics.

1 Introduction: The Expanding Area of Statistics Education Research

Fifteen years ago the research related to teaching and learning statistics was reviewed and a subsequent paper was published in this journal (Garfield, 1995). In the years since that paper was published, there has been a proliferation of research studies across many disciplines, as well as new scientific conferences and publications devoted to research in statistics education. Therefore, it seems appropriate to revisit the 1995 paper, to see what new relevant knowledge has been accumulated since then, and to re-examine the links between research and teaching practice. This paper is an attempt to do just that.

Today, statistics education can still be viewed as a new and emerging discipline, when compared to other areas of study and inquiry. This new discipline has a research base that is often difficult to locate and build upon. For many people interested in reading this area of scholarship, statistics education research can seem to be an invisible, fragmented discipline, because studies related to this topic of interest have appeared in publications from diverse disciplines, and are more often thought of as studies in those disciplines (e.g. psychology, science education, mathematics education, or in educational technology) than in the area of statistics education. In 2002 the *Statistics Education Research Journal* (SERJ, http://www.stat.auckland.ac.nz/serj) was established, and the statistics education discipline now has its first dedicated scientific journal in which to publish high-quality research. This should make it easier for future researchers to become acquainted with the discipline and locate studies for literature reviews, and for teachers of statistics to look for research relevant to the teaching and learning of statistics.

In addition to SERJ, recent research studies related to statistics education have been published in conference proceedings such as The International Conference on the Teaching Statistics (ICOTS, http://www.stat.auckland.ac. nz/∼iase/conferences.php), International Group for the Psychology of Mathematics Education (PME, http://igpme.org), The Mathematics Education Research Group of Australasia (MERGA, http://www.merga.net.au), The International Congress on Mathematics Education meetings (ICME, http://www.mathunion.org/ICMI), and The International Statistical Institute (ISI, http://isi.cbs. nl). The numerous presentations and publications from these conferences reflect the fact that there now exists an active group of educators, psychologists, and statisticians who are involved in scholarship related to the teaching and learning of statistics. In addition, more graduate students are completing dissertations in various departments that relate to teaching and learning statistics. Over 44 doctoral dissertations have been reported since 2000 (see http://www.stat.auckland.ac. nz/iasedissert).

There is much to learn from the abundant studies in the current literature, which offers important contributions to understanding the nature and development of students' statistical reasoning and what it means to understand and learn statistical concepts. In this paper we provide an overview of the more current research (Section 2), summarizing a sample of studies that have been conducted by researchers from different disciplines (psychology, mathematics education, educational psychology, and statistics education). We organize these summaries according to the general research questions addressed, and suggest our view of what can be learned from the results.

We then provide a summary of a newer focus of research that examines the development of statistical literacy, reasoning and thinking (Section 3). This is followed by a brief summary of the research on reasoning about three "big ideas" in statistics to illustrate the contributions of these studies to the emerging knowledge about teaching and learning statistics (Section 4). The ideas we chose to focus on for this paper are three of the foundational statistical concepts: distribution, centre and variability. For more information on research on additional topics see Shaughnessy (2007) and Garfield & Ben-Zvi (2005). We also use this focus of research on just three core concepts to illustrate the complexity of studying and developing a student's reasoning about these concepts and provide implications for teaching these "big ideas". We include a discussion of general implications from the research in terms of teaching and assessing students and highlight eight principles for learning statistics (Section 5).

2 Recent Research on Teaching and Learning Statistics

2.1 What are Some of the Errors and Misconceptions in Reasoning about Statistics and Probability?

Much of the literature summarized in previous reviews (e.g. Garfield & Ahlgren, 1988; Shaughnessy, 1992; Garfield, 1995; Shaughnessy *et al*., 1996) summarized research conducted primarily by psychologists on how people make judgments and decisions when faced with uncertainty (e.g. Kahneman *et al*., 1982). A focus of these studies was the identification of common faulty heuristics, biases, and misconceptions found in college students and adults (e.g. the representativeness heuristic, availability heuristic, law of small numbers, gambler's fallacy, equiprobability bias, and correlation fallacy). Subsequent research focused on methods for training individuals to reason more correctly. Some critics (e.g. Gigerenzer, 1996; Sedlmeier, 1999) argued that the cause of many identified misconceptions was actually people's inability to use proportional reasoning, required by many of these problems that involved probabilities. They suggested using frequency approach (using counts and ratios rather than percentages and decimals) and observed that subjects performed better on similar tasks when using frequencies rather than fractions or decimals.

Recognizing these persistent errors, researchers have explored ways to help college students and adults correctly use statistical reasoning, sometimes using specific training sessions (e.g. Sedlmeier, 1999). Researchers still continue to examine errors and misconceptions related to statistical reasoning. Most of these studies focus on topics related to probability (e.g. Fast, 1997; O'Connell, 1999; Hirsch & O'Donnell, 2001; Batanero & Sánchez, 2005; Tarr & Lannin, 2005). However, other studies have examined misconceptions and errors related to additional topics such as contingency tables (Batanero *et al*., 1996), sampling distributions (Yu & Behrens, 1995), significance tests (e.g. Falk & Greenbaum, 1995), and a variety of errors in statistical reasoning (e.g. Garfield, 2003; Tempelaar *et al*., 2006).

2.1.1 What can we learn from these studies?

The main message from this body of research seems to be that inappropriate reasoning about statistical ideas is widespread and persistent, similar at all age levels (even among some experienced researchers), and quite difficult to change. There are many misconceptions and faulty intuitions used by students and adults that are stubborn and difficult to overcome, despite even the best statistics instruction. In addition, students' statistical reasoning is often inconsistent from item to item or topic to topic, depending on the context of the problem and students' experience with the context. Although some types of training seem to lead to positive results, there is no strong evidence that the results were sustained beyond the training sessions or could be generalized beyond the specific types of problems used.

2.2 How do School Students Come to Understand Statistics and Probability?

In contrast to studies on misconceptions and faulty heuristics that looked at particular types of training to overcome or correct these types of problems, another line of inquiry has focused on how to develop good statistical reasoning and understanding, as part of instruction in elementary and secondary mathematics classes. Researchers began to take an interest in studying how children understand basic concepts related to data analysis when these topics began to be added to the mathematics curricula for elementary and secondary schools in the US in the 1980s and 1990s (e.g. NCTM, 2000). These studies revealed the many difficulties students have with concepts that were believed to be fairly elementary such as the mean (Rubin *et al*., 1991; Shaughnessy, 1992, 2007; Mokros & Russell, 1995; Russell & Mokros, 1996; Konold *et al*., 1997; Bright & Friel, 1998). Not surprisingly, most of the research examining school children's understanding of data analysis has been conducted by mathematics education researchers who have focused their studies on foundational concepts and their interconnections, such as data, distribution, centre and variability (e.g. Cobb *et al*., 2003b; Bakker & Gravemeijer, 2004). The focus of these studies was to investigate how students begin to understand these ideas and how their reasoning develops when using carefully designed activities assisted by technological tools.

Studies focused on students in K-12 classes, investigating how they come to understand statistical ideas such as data (e.g. Ben-Zvi & Arcavi, 2001), distribution (Watson, 2005); Pfannkuch, 2006; Prodromou & Pratt, 2006; Reading & Reid, 2006), variability (Bakker, 2004; Ben-Zvi, 2004b; Hammerman & Rubin, 2004; Reading, 2004; delMas & Liu, 2005), and probability (e.g. Abrahamson *et al*., 2006; Pratt, 2007). Some involved teaching experiments (Steffe & Thompson, 2000; Cobb *et al*., 2003a) conducted over several weeks, where a team of researchers and teachers teach and/or closely observe the class to see how particular activities and tools help develop understanding of a statistical concept or set of concepts (e.g. Cobb, 1999;

Saldanha & Thompson, 2003; Shaughnessy *et al*., 2004). Section 4 of this paper summarizes much of the work related to developing students' reasoning about the core statistical ideas of distribution, centre and variability.

Interest in probability continues among mathematics educators, primarily documenting the difficulties students have in understanding these concepts at different grade levels, the common misconceptions about probability, and the role of computer tools to help students develop reasoning about chance and uncertainty (see Jones *et al*., 2007). In a recent compilation of research studies in this area, Jones (2005) remarks that there is a need to study the evolution and development of students' reasoning and to find ways to link ideas of chance and data, rather than studying probability as a formal mathematics topic. This type of work is currently underway by Konold who is developing the *Model Chance* software, and conducting studies using the software with young children (Konold *et al*., 2007).

2.2.1 What can we learn from these studies?

The studies focused on developing students' reasoning about data and chance suggest that these ideas are often more complex and difficult for students to learn than was assumed. Studies involving elementary and secondary school students that focus on understanding of particular concepts (e.g. Ben-Zvi & Amir, 2005; Cobb *et al*., 2003b; Pfannkuch, 2006) show that carefully designed sequences of activities using appropriate technological tools can help students improve reasoning and understanding over substantial periods of time (Ben-Zvi, 2000). These studies suggest some possible sequences of activities that can help students develop ideas of important concepts such as distribution, variability and co-variation, and offer implications for the types of instructional activities and technological tools that may facilitate students learning and reasoning. More implications are included in Section 4 of this paper.

2.3 How do Pre-Service and Practicing Teachers Develop an Understanding of Statistics?

A newer line of research that has also been the focus of studies by mathematics educators is the study of pre-service or practicing teachers' knowledge of statistics and probability, and how that understanding develops in different contexts (e.g. Makar & Confrey, 2005; Leavy, 2006; Pfannkuch, 2006). Some studies of pre-service K-12 teachers focus on undergraduate college students majoring in elementary or mathematics education and how they understand and reason about statistics (e.g. Groth & Bergner, 2005).

In one study, Groth & Bergner (2005) examined the use of metaphors as a way to reveal student understanding, and were disappointed to note that students (pre-service teachers) who had completed a course in statistics have limited and often incorrect notions of the idea of sample. Leavy (2006) examined pre-service teachers' reasoning about the concept of distribution. In her one group the pre-test and post-test design that took place during a semester-long mathematics methods course, participants worked in small groups on two statistical inquiry projects requiring the collection, representation, analysis and reporting of data involving comparing distributions. She found that many teachers appeared to be gaining in their ability to reason about distributions in comparing groups while others failed to use the relevant content they had learned when comparing groups of data (see also Ciancetta, 2007).

Stohl (2005) summarizes studies that examine teachers' understanding and teaching of probability. She addresses problems resulting from mathematics teachers' more computational approach to thinking about probability and suggests ways to better prepare teachers to understand and teach this challenging topic.

Some studies on practicing teachers examine how these "students" learn and reason about statistics as a result of workshops or in-service courses (e.g. Mickelson & Heaton, 2004). Makar & Confrey (2005) and Hammerman & Rubin (2004) suggest that teachers' understanding of basic statistical analysis, such as comparing two groups, can be very confused (e.g. wanting to compare individual data points rather than group trends). However, they have found that with carefully designed instruction using innovative visualization software (such as *Fathom*, Key Curriculum Press, 2006; http://www.keypress.com/x5656.xml, or *TinkerPlots*, Konold & Miller, 2005; http://www.keypress.com/x5715.xml) they can be guided to reason more statistically. Studies also focus on how teachers teach and how their knowledge affects their teaching of statistics (e.g. Canada, 2004, 2006; Makar & Confrey, 2005; Rubin *et al*., 2005; Pfannkuch, 2006).

2.3.1 What can we learn from these studies?

The studies focused on pre-service and in-service K-12 teachers suggest that both have many difficulties in understanding and teaching core ideas of probability and statistics. The studies suggest that further explorations are needed in the issues of developing teacher knowledge of statistics as well as methods of helping teachers to understand the big ideas of statistics. A current joint IASE-ICMI study is focused on this issue (see http://www.ugr.es/∼icmi/iase_study). Efforts such as the TEAM project (Franklin & Mewborn, 2006) have attempted to bring mathematics educators and statisticians together to create new ways to prepare future K-12 teachers of statistics, by making sure that these students have a course in statistics as part of their requirements, taught in methods that emphasize conceptual understanding, data exploration, and use of appropriate technology. How to help practicing teachers develop a better knowledge of statistics is still an area that needs to be further explored.

2.4 How do Tertiary Students Learn Statistics?

Researchers across many disciplines have long been interested in the teaching and learning of statistics in college classes perhaps because of the tremendous numbers of students who enrol in introductory statistics course as a requirement for their degree programs. Some of the studies on tertiary-level students examined a particular activity or intervention; others have looked at use of a technological tool or teaching method (e.g. Noll, 2007). Several statisticians who teach statistics have focused their attention on studying students' learning in their classes (e.g.Wild *et al*., 1997; Chance, 2002; Lee *et al*., 2002). Most of these studies involve the researchers' own classes, sometimes examining one class, or involving multiple classes at the same institution.

Because of the large number and variety of studies in tertiary settings, this section is subdivided into several subsections that correspond to important questions regarding the teaching and learning of statistics after secondary school.

2.4.1 How can technology be used to promote statistical reasoning?

One of the major areas of current interest is the role technological tools (such as computers, graphing calculators, software, and Internet) can play in helping students develop statistical literacy and reasoning. Research on simulation training indicates that even a well-designed simulation is unlikely to be an effective teaching tool unless students' interaction with it is carefully structured (Lane & Peres, 2006). Simulations, however, can play a significant role in enhancing students' ability to study random processes and statistical concepts (Lane & Tang, 2000; Mills, 2004; Lane & Peres, 2006).

Using a collaborative classroom research model that implemented activities and gathered data in three different institutions, delMas *et al*. (1999) studied the development of reasoning about sampling distributions, using a simulation program and research-based activities. They found that student performance on a specially designed post test to assess students' reasoning about sampling distributions improved as the activity was changed to imbed assessments within the activity and having students make and test conjectures about different empirical sampling distributions from various populations. Lunsford *et al*. (2006) replicated this study in a different type of undergraduate course and found similar results.

Lane & Tang (2000) compared the effectiveness of simulations for teaching statistical concepts to the effectiveness of a textbook, while Aberson *et al*. (2000) studied the impact of a Web-based, interactive tutorial used to present the sampling distribution of the mean on student learning.

In a study of students' reasoning about the standard deviation, delMas (2005) had students manipulate a specially designed software tool to create histograms with the highest or lowest possible standard deviation, given a set of fixed bars. He identified some common ways in which students understand and misunderstand the standard deviation, such as thinking of "spread" as spreading butter, being evenly distributed in a graph. He also found that students had difficulty reasoning about bars in a histogram having density, in that they represent several points on a particular interval on a graph.

2.4.2 How effective is online instruction?

Another topic of interest to statistics educators has been the use of online instruction either in a Web-based course or "hybrid/blended" course, in which a significant amount of the course learning activity has been moved online, making it possible to reduce the amount of time spent in the classroom. For example, Utts (2003) and Ward (2004) found no differences in course performance for students in a hybrid versus a traditional course, and concluded that hybrid courses were not resulting in decreased student performance, although Utts noted lower evaluations by students in the hybrid courses. However, no significant differences in course performance do not imply that there were no real differences in student outcomes for the compared instructional methods.

2.4.3 What do students remember after taking statistics?

Mathews & Clark (2003) and Clark *et al*. (2003) investigated high-achieving students (an A grade in their first statistics course) from four tertiary institutions on their understanding of the mean, standard deviation and the Central Limit Theorem. Their interviews of students within the first six weeks of the term after the completion of the statistics course revealed that students tended to have relatively unsophisticated understandings of the concepts of mean and standard deviation and fragmentary recall of the Central Limit Theorem.

2.4.4 How effective is active learning in teaching statistics?

Keeler & Steinhorst (1995), Giraud (1997), and Magel (1998) investigated different methods of cooperative learning in teaching statistics at their institutions, and found generally positive results. Keeler & Steinhorst (1995) found that when students worked in pairs, the final grades were higher and more students stayed in the course than in a previous semester. Giraud (1997) found that using cooperative groups in class to work on assignments led to higher test grades than students in a lecture class. Magel (1998) found that implementing cooperative groups in a large lecture class also led to improved test scores compared to grades from a previous semester that did not use group work.

Meletiou & Lee (2002) organized their curricula along a Project-Activities-Cooperative Learning-Exercises model emphasizing statistical thinking and reasoning and an orientation towards investigating conjectures and discovery of results using data. Students were assessed on their understanding at the beginning and end of the course. Increased understanding was observed on tasks requiring statistical reasoning such as deducing whether a set of data could have been drawn at random from a particular population.

2.4.5 How can formal statistical ideas be developed from informal ideas?

Building on collaborative classroom research methods, Garfield *et al*. (2007) used Japanese Lesson Study to design, test and revise a lesson to help students develop reasoning about variability, building formal ideas from informal ideas. Japanese Lesson Study builds on the idea that teachers can conduct their own classroom research by carefully examining a particular problem in their class, trying an activity or set of activities to develop student learning, and then to evaluate, reflect and revise the activity.

A group of novice and experienced teachers designed a lesson to help reveal and build on students' informal intuitions about variability, which was taught, observed, analysed and revised. Their study suggested a sequence of activities to help students develop a deep understanding of the concept of variability and measures such as range, interquartile range and standard deviation. Schwartz *et al*. (2007) used a similar approach to develop what they referred to as students' prior knowledge, using specific activities to motivate and engage students to develop more formal reasoning about particular statistical concepts.

2.4.6 Can training improve students' statistical problem solving?

In one type of study, students are trained in a particular type of procedure to see if this affects their performance on different outcome measures. For example, Quilici & Mayer (2002) taught college students to sort statistics word problems on the basis of structural features (i.e. whether the problem could be solved by *t*-test, correlation or chi-square statistics) rather than surface features (i.e. the problem's cover story). In this study, college students displayed a higher level of structural awareness (i.e. sorting word problems on the basis of structural features) at the end rather than the beginning of their first statistics course. Recognizing that the problem one is working on can be solved using the same method as a problem one already knows is an important skill in statistical problem solving. Lovett (2001) collected participants' talk-aloud protocols to find out what ideas and strategies students were using to solve data analysis problems. She found that feedback could be given to help students improve their ability to select appropriate data analyses. Meyer & Lovett (2002) developed a computer-based training program to provide scaffolding to guide students in analysing data to solve statistical problems.

2.4.7 What is the role of affect in learning statistics?

Several researchers have explored factors related to students' success in statistics classes. Most of these studies have examined non-cognitive variables, such as students' attitudes and anxiety about statistics, (e.g. Schau & Mattern, 1997; Suanpang *et al*., 2004). This work has sometimes included development of instruments to assess student outcomes (e.g. attitudes, anxiety and reasoning). Studies have also examined relationships among different student characteristics (e.g. mathematics background, statistics attitudes or anxiety) and course outcomes for students taking statistics in education or psychology courses (e.g. Elmore & Vasu, 1986; Wisenbaker &

Scott, 1997). In addition, some of these studies examined what graduate students in education, psychology or the social sciences know and understand while or after learning statistics (e.g. Huberty *et al*., 1993; O'Connell & Corter, 1993; Finney, 2000; Earley, 2001).

2.4.8 How does students' reasoning develop during a statistics course?

In a recent study, Zieffler (2006) studied the growth in the students' reasoning about bivariate data over an introductory statistics course. He found that most of the growth in this reasoning, as measured by four administrations of a bivariate reasoning scale, happened before students formally studied a unit on bivariate data. His results suggested that perhaps the course that was designed to help students develop their general statistical reasoning, was helping them reason well about distributions of bivariate data before they formally studied that topic. He recommended the use of similar longitudinal studies, administering a set of items at three or more points of time in order to model the growth of students' reasoning during instruction, a suggestion also included in the recent report on Statistics in Mathematics Education Research (Scheaffer, 2007). In a related study, Zieffler *et al*. (2007) explore the growth in students' statistical reasoning throughout a 14-week class that embedded a sequence of simulation activities designed to develop students' inferential reasoning. They found that students' reasoning did not develop in a consistent linear way throughout the course. They also found that in some cases students' reasoning about concepts (such as sampling distribution) developed before the formal study of that topic, supporting the previous results by Zieffler (2006) about bivariate reasoning.

2.5 What Can we Learn from These Studies?

The many studies that focus on teaching and learning statistics at the college level continue to point out the many difficulties tertiary students have learning, remembering and using statistics, and point to some modest successes. These studies also serve to illustrate the many practical problems faced by college statistics instructors such as how to incorporate active or collaborative learning in a large class, whether or not to use an online or "hybrid" course, or how to select one type of software tool as more effective than another.

Many of these studies set out to answer a question such as "which is better?" However, these studies reveal that it is difficult to determine the impact of a particular teaching method or instruction tool on students' learning in a course due to limitations in study design or assessments used. While teachers would like research studies to convince them that a particular teaching method or instructional tool leads to significantly improved student outcomes, that kind of evidence is not actually available in the research literature. The results of many of the comparative studies are usually limited to that particular course setting and cannot be generalized to other courses. For example, if one study compared a particular type of active learning to a "traditional" course, results cannot be generalized to active learning versus a "traditional" course, because of the variety of methods of implementing "active learning" and the variety of "traditional" courses.

Though not based on comparative experiments, some recent classroom research studies, while not trying to be comparative, suggest some practical implications for teachers. For example, developing a deep understanding of statistics concepts is quite challenging and should not be underestimated. Research suggests that it takes time, a well thought out learning trajectory, and appropriate technological tools, activities, and discussion questions to develop deep understanding. Good reasoning about important concepts can be developed very carefully using activities and tools given enough time and revisiting of these ideas.

The research studies on attitudes and anxiety suggest that there are few strong (or large) predictors of how well students do in a statistics course, that there is little change in attitudes from beginning to end of a first course in statistics (and sometimes negative changes) and that difficulties in students' reasoning and poor attitudes are fairly widespread. The evidence does not show that if students are good in mathematics or have good attitudes that they will be likely to succeed in statistics, which is contrary to many teachers' beliefs. Instead, students who may not be strong in mathematics may work hard, enjoy the subject matter, and do very well in an introductory statistics course. Variables such as motivation, conscientiousness and desire to learn may be better predictors.

These studies suggest the types of negative value judgments students place on the study of statistics, how difficult they perceive the subject to be, and how useful, as pertaining to either one's course or the field in general. Nevertheless, the studies suggest that teachers need to cultivate more positive beliefs about the value of statistics and statistical literacy by being aware that students come to statistics courses with a great variety of expectations and perspectives on what statistics is about, and their own ability or lack of ability to succeed in the course.

One consistent problem in many of the quantitative studies focused on tertiary students has to do with the lack of high quality and consistent measures used to assess student learning outcomes. It is very common for these studies to use final exam scores or course grades as outcome measures. These measures are often problematic because they are used without establishing evidence of validity and reliability and do not necessarily measure outcomes of general value to the wider community (Garfield, 2006). In the past few years new instruments have been carefully developed and studied (e.g. delMas *et al*., in press) which may lead to less reliance on teacher-made measures.

In recent years there has also been more attention paid to distinguishing and defining learning outcomes in introductory statistics courses, and the frequently used terms for these outcomes refer to statistical literacy, statistical reasoning and statistical thinking. Clarifying desired learning outcomes can also help researchers better develop and use appropriate measures in their studies, and to align these measures with learning goals valued by the statistics education community.

3 Distinguishing between Statistical Literacy, Reasoning and Thinking

Although statistics is now viewed as a unique discipline, statistical content is most often taught in the mathematics curriculum (at elementary and secondary school level) and in departments of mathematics (tertiary level). This has led to exhortations by leading statisticians, such as Moore (1998), about the differences between statistics and mathematics. These arguments challenge statisticians and statistics educators to carefully define the unique characteristics of statistics, and in particular, the distinctions between statistical literacy, reasoning and thinking (Ben-Zvi & Garfield, 2004). Garfield & Ben-Zvi (in press) present the following definitions:

Statistical literacy is a key ability expected of citizens in information-laden societies, and is often touted as an expected outcome of schooling and as a necessary component of adults' numeracy and literacy. Statistical literacy involves understanding and using the basic language and tools of statistics: knowing what basic statistical terms mean, understanding the use of simple statistical symbols, and recognizing and being able to interpret different representations of data (Garfield, 1999; Snell, 1999; Rumsey, 2002). There are other views of statistical literacy such as Gal's (2000, 2002), whose focus is on the data consumer: Statistical literacy is portrayed as the ability to interpret, critically evaluate, and communicate about statistical information and messages. Gal (2002) argues that statistically literate behaviour is predicated on the joint activation of five interrelated knowledge bases (literacy, statistical, mathematical, context and critical), together with a cluster of supporting dispositions and enabling beliefs.

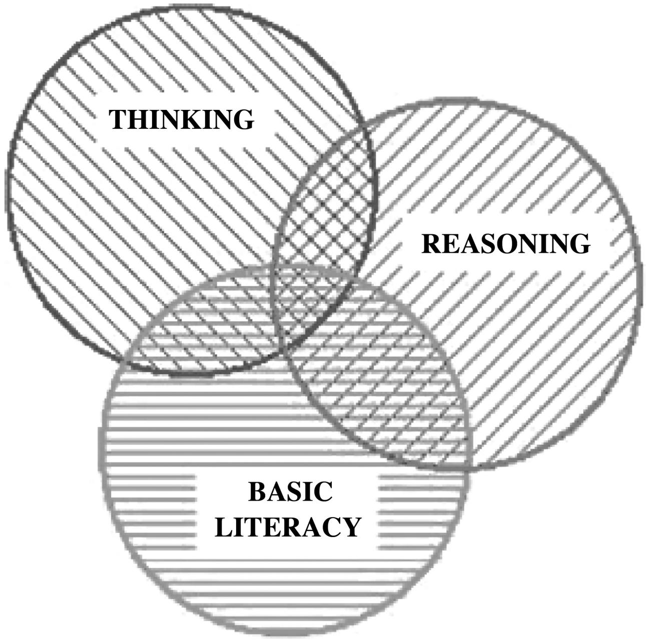

Figure 1. *The overlap and hierarchy of statistical literacy, reasoning and thinking*(Artist Website, https://app.gen.umn.edu/artist).

Statistical reasoning is the way people reason with statistical ideas and make sense of statistical information. Statistical reasoning may involve connecting one concept to another (e.g. centre and spread) or may combine ideas about data and chance. Statistical reasoning also means understanding and being able to explain statistical processes, and being able to interpret statistical results (Garfield, 2002).

Statistical thinking involves a higher order of thinking than statistical reasoning. Statistical thinking is the way professional statisticians think (Wild & Pfannkuch, 1999). It includes the knowing how and why to use a particular method, measure, design or statistical model; deep understanding of the theories underlying statistical processes and methods as well as understanding the constraints and limitations of statistics and statistical inference. Statistical thinking is also about understanding how statistical models are used to simulate random phenomena, understanding how data are produced to estimate probabilities, recognizing how, when, and why existing inferential tools can be used, and being able to understand and utilize the context of a problem to plan and evaluate investigations and to draw conclusions (Chance, 2002).

Statistical literacy, reasoning and thinking are unique areas but there is some overlap and a type of hierarchy, with statistical literacy providing the foundation for reasoning and thinking (see Figure 1). A summary of additional models of statistical reasoning and thinking can be found in Jones *et al*. (2004).

There is now a growing network of researchers interested in studying the development of students' statistical literacy, reasoning and thinking (e.g. SRTL – *The International Statistical Reasoning, Thinking, and Literacy Research Forums*, http://srtl.stat.auckland.ac.nz/). The topics of these research studies conducted by members of this community reflect the shift in emphasis in statistics instruction, from focusing on procedural understanding, i.e. statistical techniques, formulas, computations and procedures, to developing conceptual understanding and statistical literacy, reasoning and thinking.

Current research studies address this shift by focusing on some core ideas of statistics, often referred to as the "big ideas". This research focus is parallel to the increasing attention that is being paid in the educational research community to the need to clearly define and focus both research and instruction, and therefore, assessment, on the "big ideas" of a discipline (Bransford *et al*., 2000; Wiggins, 1998). The following sections offer a brief summary of some of the research on reasoning about three of the key "big ideas": distribution, centre and variability. The amount of research in these areas illustrates the complexity of studying and developing students' reasoning about these ideas. For more summaries of research on other "big ideas" as well as implications for teaching these big ideas, see Garfield & Ben-Zvi (in press).

4 Research on Developing Reasoning about Distribution, Centre and Variability

4.1 How Can We Develop Students' Reasoning about Distribution?

The research literature provides a strong case that understanding of distributions, even in the simplest forms, is much more complex and difficult than many statistics teachers believe. Although little of the research includes tertiary students, the results of studies on pre-college-level students and pre-college-level teachers demonstrate the difficulty of learning this concept, some common misconceptions, and incomplete or shallow understandings that we believe also apply to college students (see Bakker & Gravemeijer, 2004; Ben-Zvi & Arcavi, 2001).

A major outcome of several studies on how students solve statistical problems is that they tend not to see a data set (statistical distribution) as an aggregate, but rather as individual values (e.g. Hancock *et al.,* 1992). Students need to make a major conceptual leap to move from seeing data as a group of individuals—each with its own characteristics—to seeing the data as an aggregate, a group with emergent properties that often are not evident in any individual member (Konold & Higgins, 2003). To be able to think about the data as an aggregate, the aggregate must be constructed by the student (Hancock *et al*., 1992, p. 355).

Several other studies focused on how students come to conceive of shape, centre and spread as characteristics of a distribution and look at data with a notion of distribution as an organizing structure or a conceptual entity. For example, a study by Mokros & Russell (1995) suggests that students need a notion of distribution before they can sensibly choose between measures of centre and perceive them as a "representatives" of a distribution.

One of the difficulties in learning about graphical representations of distributions is confusion between case-value plots, where a bar or line represents an individual case, and histograms, where a bar represents multiple cases. These differences can cause confusion by students, leading them to try to describe shape, centre and spread of case-value plots (see delMas *et al*., 2005) or to think that bars in a histogram indicate the magnitude of single values (Bright & Friel, 1998). These studies as well as others have suggested that students should be given repeated opportunities to compare and reason about multiple representations of the same data set (e.g. Bakker & Hoffmann, 2005).

Research studies also suggest that students tend to see and use graphs as illustrations rather than as reasoning tools to learn something about a data set or gain new information about a particular problem or context (Wild & Pfannkuch, 1999; Konold & Pollatsek, 2002). Current research on student statistical understanding of distribution (e.g. Ben-Zvi & Amir, 2005; Watson, 2005; Pfannkuch, 2006; Pfannkuch & Reading, 2006; Prodromou & Pratt, 2006; Reading & Reid, 2006) recommends a shift of instructional focus from drawing various kinds of graphs and learning graphing skills to making sense of the data, for detecting and discovering patterns, for confirming or generating hypotheses, for noticing the unexpected and for unlocking the stories in the data.

The research studies summarized reveal the complexities of developing a conceptual understanding of distribution as an entity with characteristics, and view a set of data as an aggregate rather than as individual cases. It is suggested that instructors move away from teaching students how to construct graphical representations of data to a focus on what these graphs tell us with regard to a question of interest and how we may examine different graphs or manipulate a graph to learn more. These studies also point to the need to spend ample time on the topic of distribution as an important, foundational concept in statistics (Wild, 2006).

4.2 How Can We Develop Students' Reasoning about Centre?

How students understand ideas of centre has been of central interest in the research literature. Research on the concept of average or mean was at first the most common topic studied on learning statistics at the school level (see Shaughnessy, 1992; 2003; Konold & Higgins, 2003). The studies suggested that the concept of the average is quite difficult to understand by children, college students and even elementary school pre-service and in-service teachers (Russell, 1990; Groth & Bergner, 2006).

What do students remember about the mean? In general, it appears that many students, who complete introductory statistics courses, are unable to understand the idea of the mean. Mathews & Clark (2003) analysed audio-taped clinical interviews with eight college freshmen immediately after they completed an elementary statistics course with a grade of "A". The point of these interviews was not to see how quickly isolated facts could be recalled, nor was the point to see how little students remember. Rather, the goal was to determine as precisely as possible the conceptions of mean, standard deviation and the Central Limit Theorem, which the most successful students had, shortly after having completed a statistics course. The results are alarming since these top students demonstrated a lack of understanding of the mean, and could only state how to find it arithmetically. Interviewing along the same lines a larger ($n = 17$) and more diverse sample of college students from four distinct campuses, Clark *et al*. (2003) found overall the same disappointing results.

Difficulties in determining the medians of data sets have also been documented by research. Elementary school teachers have trouble finding out the medians of data sets presented graphically (Bright & Friel, 1998). Only about one-third of 12th-grade students in the US who took the National Assessment of Education Progress (NAEP) test were able to determine the median when given a set of unordered data (Zawojewski & Shaughnessy, 2000).

Another focus of research has been on the challenge of choosing an appropriate measure of centre to represent a data set. NAEP data confirm that school students frequently make poor choices in selecting measures of centre to describe data sets (Zawojewski & Shaughnessy, 2000). Choosing an appropriate measure of centre was also a challenge for high school students' school statistics course (Groth, 2003). A problem-solving clinical interview was used where students were given a problem that involved determining the typical value within a set of incomes. A different problem was set in a signal-versus-noise context (Konold & Pollatsek, 2002). Several patterns of thinking emerged in the responses to each task as students used either formal measures or more informal strategies and displayed varying amounts of attention to both data and context in formulating responses to both problems.

Similar results were found by Callingham (1997) who administered an item containing a data set structured so that the median would be a better indicator of centre than the mean, to a group of pre-service and in-service teachers. Callingham reports that most of them calculated the mean instead of the more appropriate median.

In a study on the statistical reasoning of pre-service elementary school teachers when comparing and contrasting mean, median and mode, Groth & Bergner (2006) describe four reasoning levels (unistructural, multistructural, relational, and extended abstract). Their study illustrates that attaining a deep understanding of these seemingly easy statistical concepts is a non-trivial matter and that there are complex conceptual and procedural ideas that need to be carefully developed.

The *typical value* interpretation of the arithmetic mean has received a great deal of attention in curriculum materials and in research literature (Konold & Pollatsek, 2002). Several studies have provided insights about students' thinking with regard to typical value problems. Mokros & Russell (1995) studied the characteristics of fourth through eighth graders' constructions of "average" as a *representative* number summarizing a data set. Twenty-one students were interviewed, using a series of open-ended problems that called on children to construct their own notion of mathematical representativeness. They reported that students may respond to *typical value* problems by: (i) locating the most frequently occurring value; (ii) executing an algorithm; (iii) examining the data and giving a reasonable estimate; (iv) locating the midpoint of the data; or (v) looking for a point of balance within the data set. These approaches illustrate the ways in which school students are (or are not) developing useful, general definitions for the statistical concept of average, even after they have mastered the algorithm for the mean.

Patterns of thinking about average in different contexts were investigated by Groth (2005) who studied 15 high school students. He used problems set in two different contexts: determining the typical value within a set of incomes and determining an average set in a signal-versus-noise context (Konold & Pollatsek, 2002). Analysis of the problem-solving sessions showed that some students attempted to incorporate formal measures, while others used informal estimating strategies. Students displayed varying amounts of attention to both data and context in formulating responses to both problems.

The research studies cited suggest a need for teachers to be conscious of the difficulties students have in understanding and reasoning about measures of centre, beyond a computational level. The studies imply that it is important for teachers to recognize and build on students' statistical intuitions about data and context in developing students' reasoning about measures of centre, helping them move from informal notions of these ideas to more formal ideas of mean and median. Furthermore, it is recommended that teachers provide opportunities to connect measures of centre in relation to other core concepts such as distribution, comparing groups, sampling and informal ideas of inference.

4.3 How Can We Develop Students' Reasoning About Variability?

Recent discussions in the statistics education community have drawn attention to the fact that statistics text books, instruction, public discourse, as well as research have been overemphasizing measures of centre at the expense of variability (e.g. Shaughnessy, 1997). Instead, there is a growing consensus to emphasize general distributional features such as shape, centre and spread and the connections among them in students' early experiences with data. It is also suggested to focus students' attention on the nature and sources of variability of data in different contexts, such as variability in a particular data set, outcomes of random experiments, and sampling (Shaughnessy *et al.*, 1999; Gould, 2004). These views are supported by a review of several studies by Konold & Pollatsek (2002) that has shown that "the notion of an average understood as a central tendency is inseparable from the notion of spread"

(p. 263). Their metaphor for data as *signal and noise* implies that students should come to see statistics as "the study of noisy processes – processes that have a signature, or signal" (p. 260).

Despite the widespread belief in the importance of this concept, current research demonstrates that it is extremely difficult for students to reason about variability and that we are just beginning to learn how reasoning about variability develops (Garfield & Ben-Zvi, 2005). Understanding variability has both informal and formal aspects, moving from understanding that data vary (e.g. differences in data values) to understanding and interpreting formal measures of variability (e.g. range, interquartile range and standard deviation). While students can learn how to compute formal measures of variability, they rarely understand what these summary statistics represent, either numerically or graphically, and do not understand their importance and connection to other statistical concepts. Two additional factors make the understanding of the concept even more complex: (a) variability may sometimes be desired and be of interest, and sometimes be considered error or noise (Konold & Pollatsek, 2002; Gould, 2004); (b) the multi-faceted interconnectedness of variability to concepts of distribution, centre, sampling and inference (Cobb *et al*., 2003b).

These difficulties in understanding variability are evident, in some interview studies of introductory statistics students' conceptual understanding of the standard deviation, (Matthews & Clark, 2003; delMas & Liu, 2005). DelMas & Liu's study included a computer environment designed to promote students' ability to coordinate characteristics of variation of values about the mean with the size of the standard deviation as a measure of that variation. delMas & Liu found that students moved from simple, one-dimensional understandings of the standard deviation that did not consider variation about the mean to more mean-centred conceptualizations that coordinated the effects of frequency (density) and deviation from the mean.

A variety of contexts have been used in statistics education to study students' reasoning about variability at all age levels. For example, in a study of elementary students, Lehrer & Schauble (2007) contrast students' reasoning about variability in two contexts: (a) measurement and (b) "natural" (biological). When fourth-graders were engaged in measuring the heights of a variety of objects, distribution emerged as a coordination of their activity. They were able to invent statistics as indicators of stability (e.g. centre corresponded to "real" length) and variation of measure (e.g. spread corresponded to sources of error such as tool, person, trial-to-trial). In the context of natural variation activity (growth of plants), these same students (now fifth-graders) had difficulties handling sources of natural variation and related statistics. Activities that promoted investigations of sampling (e.g. "what would be likely to happen to the distribution of plant heights if we grew them again"), and comparing distributions (e.g. how one might know whether two different distributions of height measurements could be considered "really" different) were found useful in developing students' understanding of variability.

The advantage in discussing ideas of variability in connection with ideas of centre was described by Garfield *et al*. (2007). In this study with undergraduate students, results indicated that students could develop ideas of a lot or a little variability when asked to make and test conjectures about a series of variables measuring minutes per day spent on various activities (e.g. time spent studying, talking on the phone, eating, etc.). They also found that by having students reason about the distributions of these variables informally they could move them to comparisons of formal measures of variability (e.g. standard deviation, range and interquartile range).

Other contexts where students' reasoning about variability were examined include variability in a univariate data set (Konold & Pollatsek, 2002; Petrosino *et al*., 2003; Ben-Zvi, 2004a; Groth, 2005), bivariate relationships (Cobb *et al*., 2003b; Hammerman & Rubin, 2003), comparing groups (Biehler, 2001; Lehrer & Schauble, 2002; Ben-Zvi, 2004b; Makar & Confrey, 2005;

Ciancetta, 2007), measures of spread such as the standard deviation (delMas & Liu, 2005), and sampling contexts (Chance *et al*., 2004; Reading & Shaughnessy, 2004; Watson, 2004).

The research studies on understanding variability suggest ways to help students develop reasoning about this important idea across an entire course, moving from informal ideas to more formal ideas and measures. Carefully designed activities are needed to lead students throughout this process and to help students come to view variability as a fundamental idea underlying statistics and to recognize the different "faces" of variability, such as overall spread in a data set, variability between two data sets, variability as measurement error, etc.

4.4 What can We Learn from the Research on Reasoning about Distribution, Centre and Variability?

What has been striking over 25 years of research is the difficulty encountered by students of all ages in understanding concepts related to distribution, centre and variability. Although students may be able to construct graphs and compute simple arithmetic means and standard deviations, they need help in developing an understanding of what these concepts actually mean and how to reason about them in an integrated way.

The research suggests that careful attention be paid to developing these concepts first in informal and intuitive ways, leading to more formal notions (see Garfield & Ben-Zvi, in press). For example, first introducing students informally to ideas of sampling variability and sampling distribution, in the context of exploring data and designing and carrying out simple experiments, before later formally studying more formal ideas of sampling distribution in the context of the Central Limit Theorem. Another example is having students informally compare groups using boxplots, where students are asked to make informal inferences, before later learning formal methods of making inferences based on two samples of data (Zieffler *et al*., 2007).

Another implication is that these concepts be encountered by students together, combining ideas of distribution, centre and spread, rather than as separate, isolated topics. Examples of sequences of ideas, or learning trajectories that lead from informal to formal understanding of these and other concepts, as well as activities designed to guide students to a better understanding of these concepts, can be found in Garfield & Ben-Zvi (in press).

5 General Implications of the Research Literature for Teaching Statistics

Research studies across the disciplines that relate to statistics education provide valuable information for teachers of statistics. For example, some of the studies reveal the types of difficulties students have when learning particular topics, so that teachers may be aware of where errors and misconceptions might occur and how students' statistical reasoning might develop, but also what to look for in their informal and formal assessments of their learning.

We think that the research literature is especially important to consider because it contradicts many informal or intuitive beliefs held by teachers. For example, that students earning a grade of A in a statistics class understand the basic ideas of statistics (e.g. Clark *et al*., 2003; Mathews & Clark, 2003), or that students' reasoning about statistics is consistent from topic to topic (e.g. Konold, 1995). In addition, even the most clever and carefully designed technological tool or good instructional activity will not necessarily lead students to correctly understand and reason about an abstract statistical concept (e.g. Chance *et al*., 2004).

5.1 Principles for Learning Statistics

After reviewing the research related to teaching and learning statistics over a decade ago, Garfield (1995) proposed 10 principles for learning statistics. Despite the increased number of studies since that paper was published, we believe that these principles are still valid. They are also consistent with recent publications on student learning, such as *How People Learn* (Bransford *et al*., 2000), which focus on promoting learning for understanding and developing student-centred learning environments and rethinking what is taught, how it is taught and how it is assessed. These principles have been regrouped into eight research-supported statements about student learning of statistics.

5.1.1 Students learn by constructing knowledge

Teaching is not telling, learning is not remembering. Regardless of how clearly a teacher or book tells them something, students will understand the material only after they have constructed their own meaning for what they are learning. Moreover, ignoring, dismissing, or merely "disproving" the students' current ideas will leave them intact—and they will outlast the thin veneer of course content (Bakker & Gravemeijer, 2004; Lehrer & Schauble, 2007).

Students do not come to class as "blank slates" or "empty vessels" waiting to be filled, but instead, approach learning activities with significant prior knowledge. In learning something new, they interpret the new information in terms of the knowledge they already have, constructing their own meanings by connecting the new information to what they already believe (Bransford *et al*., 2000). Students tend to accept new ideas only when their old ideas do not work, or are shown to be inefficient for purposes they think are important.

5.1.2 Students learn by active involvement in learning activities

Research suggests that students learn better if they are engaged in, and are motivated to struggle with their own learning. For this reason, if no other, students appear to learn better if they work cooperatively in small groups to solve problems and learn to argue convincingly for their approach among conflicting ideas and methods (e.g. Keeler & Steinhorst, 1995; Giraud, 1997; Magel, 1998). Small-group activities may involve groups of three or four students working in class to solve a problem, discuss a procedure, or analyse a set of data. Groups may also be used to work on an in-depth project outside of class. Group activities provide opportunities for students to express their ideas both orally and in writing, helping them become more involved in their own learning. However, just being active and having a good time is not enough to ensure learning. Good learning activities are carefully designed and the teacher has an important role to listen, probe, sum up and assess the main points (e.g. Courtney *et al*., 1994; Potthast, 1999; Perkins & Saris, 2001; Chick & Watson, 2002).

5.1.3 Students learn to do well only what they practice doing

Practice may mean hands-on activities, activities using cooperative small groups, or work on the computer. Students also learn better if they have experience applying ideas in new situations. If they practice only calculating answers to familiar, well-defined problems, then that is all they are likely to learn. Students cannot learn to think critically, analyse information, communicate ideas, make arguments and tackle novel situations, unless they are permitted and encouraged to do those things over and over in many contexts. Merely repeating and reviewing tasks is unlikely to lead to improved skills or deeper understanding (e.g. Watson, 2004; Watson & Shaughnessy, 2004; Pfannkuch, 2005).

5.1.4 It is easy to underestimate the difficulty students have in understanding basic concepts of probability and statistics

Many research studies have shown that ideas of probability and statistics are very difficult for students to learn and often conflict with many of their own beliefs and intuitions about data and chance (Shaughnessy, 1992, 2007; delMas *et al*., in press; Jones *et al*., 2007).

5.1.5 It is easy to overestimate how well students understand basic concepts

A few studies have shown that although students may be able to answer some test items correctly or perform calculations correctly, they may still misunderstand basic ideas and concepts. Also, students who receive top grades in a class may not understand and remember the basic ideas of statistics (e.g. Clark *et al*., 2003; Mathews & Clark, 2003).

5.1.6 Learning is enhanced by having students become aware of and confront their errors in reasoning

Several research studies in statistics as well as in other disciplines show that students' errors in reasoning (sometimes appearing to be misconceptions) are often strong and resilient—they are slow to change, even when students are confronted with evidence that their beliefs are incorrect (Bransford *et al*., 2000).

Students seem to learn better when activities are structured to help students evaluate the difference between their own beliefs about chance events and actual empirical results. If students are first asked to make guesses or predictions about data and random events, they are more likely to care about and process the actual results. When experimental evidence explicitly contradicts their predictions, students should be helped to evaluate this difference. In fact, unless students are forced to record and then compare their predictions with actual results, they tend to see in their data confirming evidence for their misconceptions of probability (e.g. Shaughnessy, 1977; Konold, 1989; Jones *et al*., 1999). Research in physics instruction also points to this method of testing beliefs against empirical evidence (e.g. Clement, 1987).

5.1.7 Technological tools should be used to help students visualize and explore data, not just to follow algorithms to pre-determined ends

Technology-based instruction appears to help students learn basic statistics concepts by providing different ways to represent the same data set (e.g. going from tables of data to histograms to boxplots) or by allowing students to manipulate different aspects of a particular representation in exploring a data set (e.g. changing the shape of a histogram to see what happens to the relative positions of the mean and median). Instructional software may be used to help students understand abstract ideas. For example, students may develop an understanding of the Central Limit Theorem by constructing various populations and observing the distributions of statistics computed from samples drawn from these populations (e.g. Ben-Zvi, 2000). The computer can also be used to improve students' understanding of probability by allowing them to explore and represent statistical models, change assumptions and parameters for these models and analyse data generated by applying these models (Biehler, 1991; Jones *et al*., 2007).

Innovative new visualization software, such as *Fathom* (Key Curriculum Press, 2006; http://www.keypress.com/x5656.xml) and *TinkerPlots* (Konold & Miller, 2005; http://www.keypress.com/x5715.xml) are available to students at all levels to explore data and learn to reason statistically.

5.1.8 Students learn better if they receive consistent and helpful feedback on their performance

Learning is enhanced if students have ample opportunities to express ideas and get feedback on their ideas. Feedback should be analytical, and come at a time when students are interested in it (see Garfield & Chance, 2000). There must be time for students to reflect on the feedback they receive, make adjustments, and try again. For example, evaluation of student projects may be used as a way to give feedback to students while they work on a problem during a course, not just as a final judgment when they are finished with the course. Since statistical expertise typically involves more than mastering facts and calculations, assessment should capture students' ability to reason, communicate and apply their statistical knowledge. A variety of assessment methods should be used to capture the full range of students' learning, e.g. written and oral reports on projects, minute papers (e.g. Angelo & Cross, 1993) reflecting students' understanding of material from one class session, or essay questions included on exams. Teachers should become proficient in developing and choosing appropriate methods that are aligned with instruction and key course goals, and should be skilled in communicating assessment results to students (e.g. delMas *et al*., 1999).

6 Summary

There has been a tremendous increase in research studies focused on teaching and learning statistics and probability over the past 15 years. These studies continue to span many different disciplines and differ in focus, theory, methodology and supporting literature. However, when reviewed together, they suggest the difficulties students have in learning statistics and the need to revise traditional methods of teaching. The most recent studies on the development of particular types of learning outcomes and reasoning about special topics offer many implications for changes in curriculum and teaching methods. However, there are still many open questions and much work is needed to offer more specific guidance to teachers of statistics. Since research is now elucidating some conceptual foundations (e.g. notion of distribution, variability, sampling and statistical inference) for statistics education, the consequence is that statistics education is emerging as a discipline in its own right, not an appendage to mathematics education (Garfield & Ben-Zvi, in press).

We find the most helpful results to come from collaborative research projects, and encourage future researchers to find collaborators, ideally from different disciplines, to combine expertise in the content area (statistics) student learning (education and/or psychology) and assessment. Eventually, we hope to see larger studies on particular questions of interest conducted across several institutions, using high-quality measurement instruments. We are particularly interested in seeing studies that use newer methods of analysis (e.g. hierarchical linear modelling, analysis of longitudinal data) that allow the careful study of the growth of reasoning and learning over different instructional settings and/or over a period of time. The new guidelines for using statistical methods in mathematics education research (see Scheaffer, 2007), offer many useful suggestions for improving the growing field of statistics education research as well. We look forward to the wealth of results from new studies that will be available 15 years from now that will inform our knowledge about how students learn statistics.

Acknowledgements

We gratefully acknowledge the valuable contributions of Beth Chance, Shira Yuval Adler, and two anonymous reviewers to the preparation of this paper.

References

Aberson, C.L., Berger, D.E., Healy, M.R., Kyle, D. & Romero, V.L. (2000). Evaluation of an interactive tutorial for teaching the Central Limit Theorem. *Teaching Psychol.*, **27**, 289–291.

Abrahamson, D., Janusz, R.M. & Wilensky, U. (2006, March). There once was a 9-block... – A middle-school design for probability and statistics. *J. Stat. Educ.*, **14**(1). Retrieved July 15, 2007, from http://www.amstat.org/publications/jse/v14n1/abrahamson.html

Angelo, T. & Cross, K. (1993). *A Handbook of Classroom Assessment Techniques for College Teachers* (2nd edition). San Francisco: Jossey-Bass.

Bakker, A. (2004, November). Reasoning about shape as a pattern in variability. *Stat. Educ. Res. J.*, **3**(2), 64–83. Retrieved July 15, 2007, from http://www.stat.auckland.ac.nz/∼iase/serj/SERJ3(2)_Bakker.pdf

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

Bakker, A. & Hoffmann, M. (2005). Diagrammatic reasoning as the basis for developing concepts: A semiotic analysis of students' learning about statistical distribution. *Educ. Studies Math.*, **60**, 333–358.

Batanero, C., Estepa, A., Godino, J.D. & Green, D.R. (1996). Intuitive structures and preconceptions about association in contingency tables. *J. Res. Math. Educ.*, **27**(2), 151–169.

Batanero, C. & Sánchez, E. (2005). What is the nature of high school student's conceptions and misconceptions about probability? In *Exploring Probability in School: Challenges for Teaching and Learning*, Ed. G. Jones, pp. 260–289. Dordrecht: Kluwer Academic.

Ben-Zvi, D. (2000). Toward understanding the role of technological tools in statistical learning. *Math. Thinking Learning*, **2**(1&2), 127–155.

Ben-Zvi, D. (2004a). Reasoning about data analysis. In *The Challenge of Developing Statistical Literacy, Reasoning, and Thinking*, Eds. D. Ben-Zvi and J. Garfield, pp. 121–145. Dordrecht: Kluwer Academic.

Ben-Zvi, D. (2004b, November). Reasoning about variability in comparing distributions. *Stat. Educ. Res. J.*, **3**(2), 42–63. Retrieved July 15, 2007, from http://www.stat.auckland.ac. nz/∼iase/serj/SERJ3(2)_BenZvi.pdf

Ben-Zvi, D. & Amir, Y. (2005). How do primary school students begin to reason about distributions? In *Reasoning about distribution: A collection of current research studies*. *Proceedings of the Fourth International Research Forum on Statistical Reasoning, Thinking, and Literacy* (SRTL-4), Ed. K. Makar, University of Auckland, New Zealand, 2–7 July, 2005. Brisbane: University of Queensland.

Ben-Zvi, D. & Arcavi, A. (2001). Junior high school students' construction of global views of data and data representations. *Educ. Studies Math.*, **45**(1–3), 35–65.

Ben-Zvi, D. & Garfield, J. (2004). Statistical literacy, reasoning, and thinking: goals, definitions, and challenges. In *The Challenge of Developing Statistical Literacy, Reasoning, and Thinking*, Eds. D. Ben-Zvi and J. Garfield, pp. 3–16. Dordrecht: Kluwer Academic.

Biehler, R. (1991). Computers in probability education. In *Chance Encounters: Probability in Education*, Eds. R. Kapadia and M. Borovcnik, pp. 169–211. Amsterdam: Kluwer Academic.

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

Bransford, J., Brown, A.L. & Cocking, R.R. (Eds.). (2000). *How People Learn: Brain, Mind, Experience, and School*. Washington, DC: National Academy Press.

Bright, G.W. & Friel, S.N. (1998). Graphical representations: Helping students interpret data. In *Reflections on Statistics: Learning, Teaching, and Assessment in Grades K-12*, Ed. S.P. Lajoie, pp. 63–88. Mahwah, NJ: Lawrence Erlbaum.

Callingham, R.A. (1997). Teachers' multimodal functioning in relation to the concept of average. *Math. Educ. Res. J.*, **9**(2), 205–224.

Canada, D. (2004). *Pre-Service Elementary Teachers' Conceptions of Variability*. Unpublished doctoral dissertation, Portland State University, Portland, OR.

Canada, D. (2006, May). Elementary pre-service teachers' conceptions of variation in a probability context. *Stat. Educ. Res. J.*, **5**(1), 36–64. Retrieved July 15, 2007, from http://www.stat.auckland.ac. nz/∼iase/serj/SERJ5(1)_Canada.pdf

Chance, B.L. (2002, November). Components of statistical thinking and implications for instruction and assessment. *J. Stat. Educ.*, **10**(3). Retrieved July 15, 2007, from http://www.amstat.org/publications/jse/v10.n3/chance.html

Chance, B., delMas, R. & Garfield, J. (2004). Reasoning about sampling distributions. In *The Challenge of Developing Statistical Literacy, Reasoning, and Thinking*, Eds. D. Ben-Zvi and J. Garfield, pp. 295–323. Dordrecht: Kluwer Academic.

Chick, H.L. & Watson, J.M. (2002). Collaborative influences on emergent statistical thinking – a case study. *J. Math. Beh.*, **21**, 371–400.

Ciancetta, M. (2007). *Undergraduate and graduate statistics students' reasoning when comparing distributions*. Unpublished doctoral dissertation, Portland State University, Portland, OR.

Clark, J., Karuat, G., Mathews, D. & Wimbish, J. (2003). *The Fundamental Theorem of Statistics: Classifying Student understanding of basic statistical concepts*. Retrieved October 20, 2007, from http://www1.hollins.edu/faculty/clarkjm/stat2c.pdf

Clement, J. (1987). Overcoming students' misconceptions in physics: The role of anchoring intuitions and analogical validity. *Proceedings of the Second International Seminar, Misconceptions and Educational Strategies in Science and Mathematics*. Ithaca, NY: Cornell University.

Cobb, P. (1999). Individual and collective mathematical development: The case of statistical data analysis. *Math. Thinking Learning*, **1**(1), 5–43.

Cobb, P., Confrey, J., diSessa, A., Lehrer, R. & Schauble, L. (2003a). Design experiments in educational research. *Educ. Res.*, **32**(1), 9–13.

Cobb, P., McClain, K. & Gravemeijer, K.P.E. (2003b). Learning about statistical covariation. *Cogn. Instruct.*, **21**(1), 1–78.

Courtney, D.P., Courtney, M. & Nicholson, C. (1994). The effect of cooperative learning as an instructional practice at the college level. *College Student J.*, **28**, 471–477.

delMas, R. (2005). *Activities to promote an understanding of variability*. Paper presented at the Annual Meeting of the American Mathematics Association of Two-Year Colleges (AMATYC). San Diego, CA, November 11.

delMas, R.C., Garfield, J. & Chance, B.L. (1999, November). A model of classroom research in action: developing simulation activities to improve students' statistical reasoning. *J. Stat. Educ.*, **7**(3). Retrieved July 15, 2007, from http://www.amstat.org/publications/jse/secure/v7 n3/delmas.cfm

delMas, R., Garfield, J. & Ooms, A. (2005). *Using assessment items to study students' difficulty with reading and interpreting graphical representations of distributions*. Presented at the Fourth Forum on Statistical Reasoning, Thinking, and Literacy (SRTL-4), July 6, 2005. Auckland, New Zealand.

delMas, R., Garfield, J., Ooms, A. & Chance, B. (in press). Assessing students' conceptual understanding after a first course in statistics. *Stat. Educ. Res. J.*

delMas, R. & Liu, Y. (2005, May). Exploring students' conceptions of the standard deviation. *Stat. Educ. Res. J.*, **4**(1), 55–82. Retrieved July 15, 2007, from http://www.stat.auckland.ac. nz/∼iase/serj/SERJ4(1)_delMas_Liu.pdf

Earley, M.A. (2001). *Improving statistics education through simulations: the case of the sampling distribution*. Paper presented at the Annual Meeting of the Mid-Western Educational Research Association. Chicago, Illinois, October 24–27.

Elmore, P.B. & Vasu, E.S. (1986). Model of statistics achievement using spatial ability, feminist attitudes and mathematics-related variables as predictors. *Educ. Psychol. Measure.*, **46**, 215–222.

Falk, R. & Greenbaum, C. (1995). Significance tests die hard. *Theory Psych.*, **5**(1), 75–98.

Fast, G.R. (1997). *Using analogies to produce long term conceptual change: Overcoming high school mathematics students' probability misconceptions*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL.

Finney, S.J. (2000). *The meaning of success for students in statistical methods courses: A phenomenological study*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Franklin, C.A. & Mewborn, D.S. (2006). The statistical education of grades Pre-K-2 teachers: A shared responsibility. In *Thinking and Reasoning with Data and Chance: Sixty-eighth Yearbook*, Ed. G.F. Burrill, Reston, VA: National Council of Teachers of Mathematics (NCTM).

Gal, I. (Ed.) (2000). *Adult Numeracy Development: Theory, Research, Practice*. Cresskill, NJ: Hampton Press.

Gal, I. (2002). Adults' statistical literacy: Meaning, components, responsibilities. *Int. Stat. Rev.*, **70**(1), 1–25.

Garfield, J. (1995). How students learn statistics. *Int. Stat. Rev.*, **63**, 25–34.

Garfield, J. (1999). *Thinking about statistical reasoning, thinking, and literacy*. Paper presented at First International Research Forum on Statistical Reasoning, Thinking and Literacy (STRL-1), Kibbutz Be'eri, Israel.

Garfield, J. (2002, November). The challenge of developing statistical reasoning. *J. Stat. Educ.*, **10**(3). Retrieved July 15, 2007, from http://www.amstat.org/publications/jse/v10n3/garfield.html

Garfield, J. (2003). Assessing statistical reasoning. *Stat. Educ. Res. J.*, **2**(1), 22–38.

Garfield, J. (2006). Collaboration in statistics education research: Stories, reflections, and lessons learned. In *Proceedings of the Seventh International Conference on Teaching Statistics*, Eds. A. Rossman and B. Chance, [CD-ROM]. Voorburg, The Netherlands: International Statistical Institute.

Garfield, J. & Ahlgren, A. (1988). Difficulties in learning basic concepts in probability and statistics: implications for research. *J. Res. Math. Educ.*, **19**(1), 44–63.

Garfield, J. & Ben-Zvi, D. (2005, May). A framework for teaching and assessing reasoning about variability. *Stat. Educ. Res. J.*, **4**(1), 92–99. Retrieved July 15, 2007, from http://www.stat.auckland.ac.nz/∼iase/serj/SERJ4(1)_Garfield_BenZvi.pdf

Garfield, J. & Ben-Zvi, D. (in press). *Developing Students' Statistical Reasoning: Connecting Research and Teaching Practice*. Dordrecht: Springer.

Garfield, J. & Chance, B. (2000). Assessment in statistics education: Issues and challenges. *Math. Thinking Learning*, **2**, 99–125.

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

Gigerenzer, G. (1996). On narrow norms and vague heuristics: A reply to Kahneman and Tversky (1996). *Psychol. Rev.*, **103**(3), 592–596.

Giraud, G. (1997, November). Cooperative learning and statistics instruction. *J. Stat. Educ.*, **5**(3). Retrieved July 15, 2007, from http://www.amstat.org/publications/jse/v5n3/giraud.html

Gould, R. (2004, November). Variability: One statistician's view. *Stat. Educ. Res. J.*, **3**(2), 7–16. Retrieved July 15, 2007, from http://www.stat.auckland.ac.nz/∼iase/serj/SERJ3(2)_Gould.pdf

Groth, R.E. (2003). High school students' levels of thinking in regard to statistical study design. *Math. Educ. Res. J.*, **15**, 252–269.

Groth, R.E. (2005). An investigation of statistical thinking in two different contexts: Detecting a signal in a noisy process and determining a typical value. *J. Math. Beh.*, **24**(2), 109–124.

Groth, R.E. & Bergner, J.A. (2005, November). Preservice elementary school teachers' metaphors for the concept of statistical sample. *Stat. Educ. Res. J.*, **4**(2), 27–42. Retrieved July 15, 2007, from http://www.stat.auckland.ac.nz/∼iase/serj/SERJ4(2)_groth_bergner.pdf

Groth, R.E. & Bergner, J.A. (2006). Preservice elementary teachers' conceptual and procedural knowledge of mean, median, and mode. *Math. Thinking Learning*, **8**, 37–63.

Hammerman, J. & Rubin, A. (2003). Reasoning in the presence of variability. In *Reasoning About Variability: A Collection of Current Research Studies*, Ed. C. Lee, Dordrecht: Kluwer Academic.

Hammerman, J.K. & Rubin, A. (2004, November). Strategies for managing statistical complexity with new software tools. *Stat. Educ. Res. J.*, **3**(2), 17–41. Retrieved July 15, 2007, from http://www.stat.auckland.ac.nz/∼iase/serj/SERJ3(2)_Hammerman_Rubin.pdf

Hancock, C., Kaput, J.J. & Goldsmith, L.T. (1992). Authentic inquiry with data: Critical barriers to classroom implementation. *Educ. Psychol.*, **27**(3), 337–364.

Hirsch, L.S. & O'Donnell, A.M. (2001, July). Representativeness in statistical reasoning: Identifying and assessing misconceptions. *J. Stat. Educ.* **9**(2). Retrieved July 15, 2007, from http://www.amstat.org/publications/jse/v9n2/hirsch.html

Huberty, C.J., Dresden, J. & Bak, B. (1993). Relations among dimensions of statistical knowledge. *Educ. Psychol. Measure.*, **53**, 523–532.

Jones, G.A. (2005). Reflections. In *Exploring Probability in School: Challenges for Teaching and Learning*, Ed. G.A. Jones, pp. 367–372. New York: Springer.

Jones, G.A., Langrall, C.W. & Mooney, E.S. (2007). Research in probability: Responding to classroom realities. In *The Second Handbook of Research on Mathematics*, Ed. F.K. Lester, pp. 909–956. Reston, VA: National Council of Teachers of Mathematics (NCTM).

Jones, G.A., Langrall, C.W., Mooney, E.S. & Thornton, C.A. (2004). Models of development in statistical reasoning. In *The Challenge of Developing Statistical Literacy, Reasoning, and Thinking*, Eds. D. Ben-Zvi and J. Garfield, pp. 97–117. Dordrecht: Kluwer Academic.

Jones, G.A., Langrall, C.W., Thornton, C.A. & Mogill, A.T. (1999). Students' probabilistic thinking in instruction. *J. Res. Math. Educ.*, **30**, 487–519.

Kahneman, D., Slovic, P. & Tversky, A. (Eds.). (1982). *Judgment Under Uncertainty: Heuristics and Biases*. New York: Cambridge University Press.

Keeler, C.M. & Steinhorst, R.K. (1995, July). Using small groups to promote active learning in the introductory statistics course: A report from the field. *J. Stat. Educ.*, **3**(2). Retrieved July 15, 2007, from http://www.amstat.org/publications/jse/v3n2/keeler.html

Key Curriculum Press (2006). *Fathom Dynamic Data™ Software* (Version 2.03) [Computer software]. Emeryville, CA: Key Curriculum Press. Available: http://www.keypress.com/x5656.xml

Konold, C. (1989). Informal conceptions of probability. *Cogn. Instruct.*, **6**(1), 59–98.

Konold, C. (1995, March). Issues in assessing conceptual understanding in probability and statistics. *J. Stat. Educ.*, **3**(1). Retrieved July 15, 2007, from http://www.amstat.org/publications/jse/v3n1/konold.html

Konold, C. & Higgins, T. (2003). Reasoning about data. In *A Research Companion to Principles and Standards for School Mathematics*, Eds. J. Kilpatrick, W.G. Martin and D.E. Schifter, pp. 193–215. Reston, VA: National Council of Teachers of Mathematics (NCTM).

Konold, C., Kazak, S., Dahlstrom-Hakki, I., Lehrer, R. & Kim, M.J. (2007, August). *Students building data models*. Paper presented at the Fifth International Research Forum on Statistical Reasoning, Thinking and Literacy (SRTL-5). The University of Warwick, UK.

Konold, C. & Miller, C.D. (2005). *TinkerPlots: Dynamic Data Exploration*™ (Version 1.0) [Computer software]. Emeryville, CA: Key Curriculum Press. Available: http://www.keypress. com/x5715.xml

Konold, C. & Pollatsek, A. (2002). Data analysis as the search for signals in noisy processes. *J. Res. Math. Educ.*, **33**(4), 259–289.

Konold, C., Pollatsek, A., Well, A. & Gagnon, A. (1997). Students analyzing data: research of critical barriers. In *Research on the Role of Technology in Teaching and Learning Statistics*, Eds. J. Garfield and G. Burrill, pp. 151–168. Voorburg, The Netherlands: International Statistical Institute.

Lane, D.M. & Peres, S.C. (2006). Interactive simulations in the teaching of statistics: Promise and pitfalls. In *Proceedings of the Seventh International Conference on Teaching Statistics*. [CD-ROM], Eds. A. Rossman and B. Chance. Voorburg, The Netherlands: International Statistical Institute.

Lane, D.M. & Tang, Z. (2000). Effectiveness of simulation training on transfer of statistical concepts. *J. Educ. Comput. Res.*, **22**(4), 383–396.

Leavy, A. (2006, November). Using Data comparisons to support a focus on distribution: Examining pre-service teachers' understandings of distribution when engaged in statistical inquiry. *Stat. Educ. Res. J.*, **5**(2), 89–114. Retrieved July 15, 2007, from http://www.stat.auckland.ac.nz/~iase/serj/SERJ5(2)_Leavy.pdf

Lee, C., Zeleke, A. & Wachtel, H. (2002). Where do students get lost? The concept of variation. In *Developing a Statistically Literate Society: Proceedings of the Sixth International Conference on Teaching Statistics* (CD-ROM), Ed. B. Phillips. Voorburg, The Netherlands: International Statistical Institute.

Lehrer, R. & Schauble, L. (2002). Distribution: A resource for understanding error and natural variation. In *Proceedings of the Sixth International Conference on Teaching Statistics: Developing a Statistically Literate Society*, Ed. B. Phillips. Voorburg, The Netherlands: International Statistical Institute.

Lehrer, R., & Schauble, L. (2007). Contrasting emerging conceptions of distribution in contexts of error and natural variation. In *Thinking with data*, Eds., M. Lovett and P. Shah, pp. 149–176. Mahwah, NJ: Erlbaum.

Lovett, M.C. (2001). A Collaborative convergence on studying reasoning processes: A case study in statistics. In *Cognition and Instruction: Twenty-Five Years of Progress*, Eds. S. Carver and D. Klahr, pp. 347–384. Mahwah, NJ: Erlbaum.

Lunsford, M.L. Rowell, G.H. & Goodson-Espy, T. (2006, November). Classroom research: Assessment of student understanding of sampling distributions of means and the central limit theorem in post-calculus probability and statistics classes. *J. Stat. Educ.*, **14**(3). Retrieved July 17, 2007, from www.amstat.org/publications/jse/v14n3/lunsford.html

Magel, R.C. (1998, November). Using cooperative learning in a large introductory statistics class. *J. Stat. Educ.*, **6**(3). Retrieved July 15, 2007, from http://www.amstat.org/publications/jse/v6 n3/magel.html

Makar, K. & Confrey, J. (2005). Using distributions as statistical evidence in well-structured and ill-structured problems. In *Reasoning About Distribution: A Collection of Current Research Studies. Proceedings of the Fourth International Research Forum on Statistical Reasoning, Thinking, and Literacy (SRTL-4)*, Ed. K. Makar, University of Auckland, New Zealand, 2–7 July. Brisbane: University of Queensland.

Mathews, D. & Clark, J. (2003). *Successful Students' Conceptions of Mean, Standard Deviation and the Central Limit Theorem*. Unpublished paper. Retrieved October 20, 2007, from http://www1.hollins.edu/faculty/clarkjm/stats1.pdf.

Meletiou, M. & Lee, C. (2002, December). Teaching students the stochastic nature of statistical concepts in an introductory statistics course. *Stat. Educ. Res. J.*, **1**(2), 22–37. Retrieved July 15, 2007, from http://www.stat.auckland.ac.nz/~iase/serj/SERJ1(2).pdf

Meyer, O. & Lovett, M. (2002). Implementing a computerized tutor in a statistical reasoning course: Getting the big picture. In *Proceedings of the sixth international conference on teaching statistics: Developing a statistically literate society*, Ed. B. Phillips. Voorburg, The Netherlands: International Statistical Institute.

Mickelson, W.T. & Heaton, R.M. (2004). Primary teachers' statistical reasoning about data. In *The Challenge of Developing Statistical Literacy, Reasoning, and Thinking*, Eds. D. Ben-Zvi and J. Garfield, pp. 295–324. Dordrecht: Kluwer Academic.

Mills, J.D. (2004). Learning abstract statistics concepts using simulation. *Educ. Res. Q.*, **28**(4), 18–33.

Mokros, J. & Russell, S.J. (1995). Children's concepts of average and representativeness. *J. Res. Math. Educ.*, **26**, 20–39.

Moore, D. (1998). Statistics among the liberal arts. *J. Amer. Stat. Assoc.*, **93**, 1253–1259.

National Council of Teachers of Mathematics (NCTM). (2000). *Principles and Standards for School Mathematics*. Reston, VA: National Council of Teachers of Mathematics, NCTM.

Noll, J. (2007). *Statistics teaching assistants' statistical content knowledge and pedagogical content knowledge*. Unpublished doctoral dissertation, Portland, OR.: Portland State University.

O'Connell, A.A. (1999). Errors in probability problem-solving. *Educ. Res. Evalu.*, **5**(1), 1–21.

O'Connell, A.A. & Corter, J.E. (1993). Student misconceptions in probability problem-solving. *Proceedings of the AERA Conference*, pp. 1–15. Atlanta, GA.

Perkins, D.V. & Saris, R.N. (2001). A "jigsaw classroom" technique for undergraduate statistics courses. *Teaching Psychol.*, **28**, 111–113.

Petrosino, A.J., Lehrer, R. & Schauble, L. (2003). Structuring error and experimental variation as distribution in the fourth grade. *Math. Thinking Learning*, **5**, 131–156.

Pfannkuch, M. (2005). Probability and statistical inference: How can teachers enable learners to make the connection? In *Exploring Probability in School: Challenges for Teaching and Learning*, Ed. G.A. Jones, pp. 267–294. New York: Springer.

Pfannkuch, M. (2006, November). Comparing Box plot distributions: A teacher's reasoning. *Stat. Educ. Res. J.*, **5**(2), 27–45. Retrieved July 15, 2007, from http://www.stat.auckland.ac.nz/∼iase/serj/SERJ5(2)_Pfannkuch.pdf

Pfannkuch, M. & Reading, C. (2006, November). Reasoning about distribution: A complex process. *Stat. Educ. Res. J.*, **5**(2), 4–9. Retrieved July 15, 2007, from http://www.stat.auckland.ac.nz/∼iase/serj/SERJ5(2).pdf

Potthast, M.J. (1999). Outcomes of using small-group cooperative learning experiences in introductory statistics courses. *College Student J.*, **33**, 34–42.

Pratt, D. (2007). How do teachers foster students' understanding of probability? In *Exploring Probability in School: Challenges for Teaching and Learning*, Ed. G.A. Jones, pp. 182–201. Dordrecht: Kluwer Academic.

Prodromou, T. & Pratt, D. (2006, November). The role of causality in the co-ordination of two perspectives on distribution within a virtual simulation. *Stat. Educ. Res. J.*, **5**(2), 69–88. Retrieved July 15, 2007, from http://www.stat.auckland.ac.nz/∼iase/serj/SERJ5(2).pdf

Quilici, J.L. & Mayer, R.E. (2002). Teaching students to recognize structural similarities between statistics word problems. *Appl. Cogn. Psychol.*, **16**(3), 325–342.

Reading, C. (2004, November). Student description of variation while working with weather data. *Stat. Educ. Res. J.*, **3**(2), 84–105. Retrieved July 15, 2007, from http://www.stat.auckland.ac.nz/∼iase/serj/SERJ3(2)_Reading.pdf

Reading, C. & Reid, J. (2006, November). An emerging hierarchy of reasoning about distribution: From a variation perspective. *Stat. Educ. Res. J.*, **5**(2), 46–68. Retrieved July 15, 2007, from http://www.stat.auckland.ac.nz/∼iase/serj/SERJ5(2)_reading_reid.pdf

Reading, C. & Shaughnessy, M. (2004). Reasoning about variation. In *The Challenge of Developing Statistical Literacy, Reasoning, and Thinking*, Eds. D. Ben-Zvi and J. Garfield, pp. 201–226. Dordrecht: Kluwer Academic.

Rubin, A., Bruce, B. & Tenney, Y. (1991). Learning about sampling: Trouble at the core of statistics. In *Proceedings of the Third International Conference on Teaching Statistics*, Ed. D. Vere-Jones, Vol. 1, pp. 314–319. Voorburg, The Netherlands: International Statistical Institute.

Rubin, A., Hammerman, J., Cambell, C. & Puttick, G. (2005). The effect of distributional shape on group comparison strategies (SRTL-4). In *Reasoning about Distribution: A Collection of Current Research Studies. Proceedings of the Fourth International Research Forum on Statistical Reasoning, Thinking, and Literacy (SRTL-4)*, Ed. K. Makar. University of Auckland, New Zealand, 2–7 July. Brisbane: University of Queensland.

Rumsey, D.J. (2002, November). Statistical literacy as a goal for introductory statistics courses. *J. Stat. Educ.*, **10**(3). Retrieved July 15, 2007, from http://www.amstat.org/publications/jse/v10 n3/rumsey2.html

Russell, S.J. (1990). Issues in training teachers to teach statistics in the elementary school: A world of uncertainty. In *Training Teachers to Teach Statistics: Proceedings of the International Statistical Institute Round Table Conference*, Ed. A. Hawkins, pp. 59–71. Voorburg, The Netherlands: International Statistical Institute.

Russell, S.J. & Mokros, J. (1996). What do children understand about average? *Teaching Children Math.*, **2**, 360–364.

Saldanha, L. & Thompson, P. (2003). Conceptions of sample and their relationship to statistical inference. *Educ. Studies Math.*, **51**, 257–270.

Schau, C. & Mattern, N. (1997). Assessing students' connected understanding of statistical relationships. In *The Assessment Challenge in Statistics Education*, Eds. I. Gal and J.B. Garfield, pp. 91–104. Amsterdam, The Netherlands: IOS Press.

Scheaffer, R. (2007). *Using Statistics effectively in mathematics education research*: *A report from a series of workshops organized by the American Statistical Association*. Retrieved October 20, 2007, from: http://www.amstat.org/research_grants/pdfs/SMERReport.pdf

Schwartz, D.L., Sears, D. & Chang, J. (2007). Reconsidering prior knowledge. In *Thinking with Data*, Eds. M.C. Lovett and P. Shah, pp. 319–344. Mahwah, NJ: Erlbaum. Available: http://aaalab.stanford.edu/papers/Schwartz_Reconsidering_Prior_K.pdf

Sedlmeier, P. (1999). *Improving Statistical Reasoning: Theoretical Models and Practical Implications*. Hillsdale, NJ: Erlbaum.

Shaughnessy, J.M. (1977). Misconceptions of probability: An experiment with a small-group, activity-based, model building approach to introductory probability at the college level. *Educ. Studies Math.*, **8**, 285–316.

Shaughnessy, J.M. (1992). Research in probability and statistics: Reflections and directions. In *Handbook of Research on Mathematics Teaching and Learning*, Ed. D. Grouws, pp. 465–494. New York: Macmillan.

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

Shaughnessy, J.M. (2003). Research on students' understanding of probability. In *A Research Companion to Principles and Standards for School Mathematics*, Eds. J. Kilpatrick, W.G. Martin, and D. Schifter, pp. 216–226. Reston, VA: National Council of Teachers of Mathematics (NCTM).

Shaughnessy, J.M. (2007). Research on statistics learning and reasoning. In *The Second Handbook of Research on Mathematics*, Ed. F.K. Lester, pp. 957–1010. Reston, VA: National Council of Teachers of Mathematics (NCTM).

Shaughnessy, J.M., Ciancetta, M., Best, K. & Canada, D. (2004, April). *Students' attention to variability when comparing distributions*. Paper presented at the Research Pre-session of the 82nd annual meeting of the National Council of Teachers of Mathematics, Philadelphia, PA.

Shaughnessy, J.M., Garfield, J. & Greer, B. (1996). Data handling. In *International Handbook of Mathematics Education*, Eds. A.J. Bishop, K. Clements, C. Keitel, J. Kilpatrick, and C. Laborde, Vol. 1, pp. 205–237. Dordrecht: Kluwer Academic.

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

Snell, L. (1999). *Using chance media to promote statistical literacy*. Paper presented at the 1999 Joint Statistical Meetings, Dallas, TX.

Steffe, L.P. & Thompson, P.W. (2000). Teaching experiment methodology: Underlying principles and essential elements. In *Handbook of Research Design in Mathematics and Science Education*, Eds. A.E. Kelly and R.A. Lesh, pp. 267–307. Hillsdale, NJ: Erlbaum.

Stohl, J. (2005). Probability in teacher education and development. In *Exploring Probability in School: Challenges for Teaching and Learning*, Ed. G.A. Jones, pp. 345–366. New York: Springer.

Suanpang, P., Petocz, P. & Kalceff, W. (2004). Student attitudes to learning business statistics: comparison of online and traditional methods. *Educ. Technol. Soc.*, **7**(3), 9–20.

Tarr, J.E. & Lannin, J.K. (2005). How can teachers build notions of conditional probability and independence? In *Exploring Probability in School: Challenges for Teaching and Learning*, Ed. G.A. Jones, pp. 215–238. New York: Springer.

Tempelaar, D.T., Gijselaers, W.H. & van der Loeff, S.S. (2006, March). Puzzles in statistical reasoning. *J. Stat. Education*, **14**(1). Retrieved July 15, 2007, from http://www.amstat.org/publications/jse/v14n1/tempelaar.html

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

Ward, B. (2004, November). The best of both worlds: A hybrid statistics course. *J. Stat. Educ.*, **12**(3), 74–79. Retrieved July 15, 2007, from http://www.amstat.org/publications/jse/v12 n3/ward.html

Watson, J. (2004). Developing reasoning about samples. In *The Challenge of Developing Statistical Literacy, Reasoning and Thinking*, Eds. D. Ben-Zvi and J. Garfield, pp. 277–294. Dordrecht: Kluwer Academic.

Watson, J. (2005). Developing an awareness of distribution. In *Reasoning about distribution: A collection of research studies*. *Proceedings of the Fourth International Forum on Statistical Reasoning, Thinking, and Literacy (SRTL-4)*, Ed. K. Makar. University of Auckland, New Zealand, 2–7 July. Brisbane: University of Queensland [CD-ROM].

Watson, J.M. & Shaughnessy, J.M. (2004). Proportional reasoning: Lessons from research in data and chance. *Math. Teaching Middle School*, **10**, 104–109.

Wiggins, G. (1998). *Understanding by design*. Alexandria, VA: ASCD Press.

Wild, C. (2006, November). The concept of distribution. *Stat. Educ. Res. J.*, **5**(2), 10–26. Retrieved July 15, 2007, from http://www.stat.auckland.ac.nz/~iase/serj/SERJ5(2)_Wild.pdf

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

Wild, C., Triggs, C. & Pfannkuch, M. (1997). Assessment on a budget: Using traditional methods imaginatively. In *The Assessment Challenge in Statistics Education*, Eds. I. Gal and J.B. Garfield. Amsterdam: IOS Press.

Wisenbaker, J. & Scott, J. (1997). *Modeling aspects of students' attitudes and achievement in introductory statistics courses*. Paper presented at annual meeting of the American Educational Research Association (AERA), Chicago, IL.

Yu, C.H. & Behrens, J.T. (1995). Applications of multivariate visualization to behavioral sciences. *Behav. Res. Methods, Instrum. Comput.*, **27**(2), 264–271.

Zawojewski, J.S. & Shaughnessy, J.M. (2000). Mean and median: Are they really so easy? *Math. Teaching Middle School*, **5**(7), 436–440.
Zieffler, S.A. (2006). A longitudinal investigation of the development of college students' *reasoning about bivariate data during an introductory statistics course*. Unpublished Ph.D. Thesis. University of Minnesota. Retrieved July 15, 2007, from http://www.stat.auckland.ac.nz/∼iase/publications/dissertations/06.Zieffler.Dissertation.
Zieffler, S.A., Garfield, J., del Mas, R. & Gould, R. (2007). *Studying the development of college students' reasoning about statistical inference*. Paper presented at the Fifth International Research Forum on Statistical Reasoning, Thinking and Literacy (SRTL-5). University of Warwick, UK.

Résumé

Cet article présente une vue d'ensemble de la recherche actuelle sur l'enseignement et l'étude de la statistique. Il résume les analyses qui ont été menées par des chercheurs de disciplines différentes. Le travail s'organise autour de questions de recherche générales qui ont été adressées, et suggère ce que l'on peut apprendre à partir de ces résultats au sujet de chacune de ces questions. Les implications de la recherche sont décrites en termes de huit principes sur l'étude de la statistique de Garfield (1995) qui sont repris selon les résultats d'études actuelles.

[*Received March 2007, accepted September 2007*]