PUBLISHING IN *SERJ*: AN ANALYSIS OF PAPERS FROM 2002–2009

ANDREW ZIEFFLER
University of Minnesota
zief0002@umn.edu

JOAN GARFIELD
University of Minnesota
jbg@umn.edu

ROBERT C. DELMAS
University of Minnesota
delma001@umn.edu

LAURA LE
University of Minnesota
free0312@umn.edu

REBEKAH ISAAK
University of Minnesota
isaak009@umn.edu

AUDBJORG BJORNSDOTTIR
University of Minnesota
bjrns001@umn.edu

JIYOON PARK
University of Minnesota
parkx666@umn.edu

ABSTRACT

SERJ has provided a high quality professional publication venue for researchers in statistics education for close to a decade. This paper presents a review of the articles published to explore what they suggest about the field of statistics education, the researchers, the questions addressed, and the growing knowledge base on teaching and learning statistics. We present a detailed analysis of these articles in order to address the following questions: What is being published and why, who is publishing research in SERJ, how is the research being carried out, and what do the results suggest about future research? Implications for future directions in statistics education research are suggested.

Keywords: *Statistics education research; Content analysis; Methodology*

Statistics Education Research Journal, 10(2), 5-26, http://www.stat.auckland.ac.nz/serj

1. INTRODUCTION

The *Statistics Education Research Journal* (*SERJ*) is a peer-reviewed electronic journal, which "aims to advance research-based knowledge to improve the teaching, learning, and understanding of statistics or probability at all educational levels and in both formal (classroom-based) and informal (out-of-classroom) contexts" (*SERJ*, n.d.). *SERJ*, which is published twice a year, has been freely available since its inception in 2002. It is currently an official journal of the International Statistical Institute. Whereas research on the teaching and learning of statistics continues to be published in a variety of other journals in disciplines such as mathematics education and psychology (see Zieffler, Garfield, Alt, Dupuis, Holleque, & Chang, 2008), as well as in statistics education journals with a wider focus (e.g., *Teaching Statistics*, *Journal of Statistics Education*, *and Technology Innovations in Statistics Education*), *SERJ* is the sole journal that is dedicated to publishing only statistics education research, and its creation in 2002 was an important event for the relatively new field of statistics education (Garfield & Ben-Zvi, 2008).

Research in statistics education has been conducted in a variety of disciplines and has been heavily influenced by the diverse background and perspectives of the scholars trained in these different disciplines (Garfield & Ben-Zvi, 2007). Now that *SERJ* has been in existence for close to a decade, it is useful to review the articles published to determine what they suggest about the field of statistics education research: the questions addressed, the methods used, the types of researchers involved, and the growing knowledge base. To this end, this paper presents a set of analyses of the articles published in *SERJ* in order to address the following questions: What is being published and why (on what literature is this research based), who is publishing research in *SERJ*, how is the research being carried out, and what do these results suggest about future research?

2. BACKGROUND

Research on teaching and learning statistics has been increasing over the past several decades. There has been growing interest in the challenges of helping students learn statistics due to the increased numbers of students studying statistics at all levels (Garfield & Ben-Zvi, 2007, 2008). Unlike research on other aspects of education, statistics is perhaps unique because it is taught in a wide variety of departments, and to diverse groups of students (e.g., liberal arts students with little mathematics background; students in science, mathematics and engineering with strong quantitative skills). Although statisticians have gone to great lengths to point out that statistics is a discipline separate from mathematics, it is still included in the elementary and secondary mathematics curricula in many countries today, creating a challenge for teachers who may have had little if any formal study of statistics nor experience analyzing data (Cobb & Moore, 1997).

Several reviews of the literature on teaching and learning statistics over the past twenty years provide analyses of difficult concepts, types of reasoning, challenges to teaching, and technology issues (Garfield, 1995; Garfield & Ben-Zvi, 2007; Konold & Higgins, 2003; Shaughnessy, 1992, 2007). Becker (1996) published a meta-analysis on research on teaching statistics, based on 375 articles, and found that although "extensive resources [were] currently available to instructors," less than 30% of the literature she examined reported results from empirical studies. She also found at that time that most of the literature "on the teaching of statistics is largely anecdotal and comprises mainly recommendations for instruction based on the experiences and intuitions of individual instructors" (p. 71, 85). More recently, van der Merwe and Wilkinson (2011) reviewed and categorized scholarship related to the teaching and learning of statistics finding that much of it focused on students' difficulties and challenges related to statistical reasoning.

A consistent finding in these reviews has been the diversity among researchers in terms of their backgrounds, focus, and publication venues. This has led to a lack of coherence in the field, as many articles do not cite relevant work by other researchers because it was not published in particular journals or disciplines. This problem was partially addressed by the creation of *SERJ* in 2002, the first journal devoted exclusively to research in this domain. Such a journal offers the promise of making the research more accessible, connecting researchers across disciplines, and using a consistent set of standards for papers published. It also provides the opportunity to analyze what has been published in this journal and what changes may be evident in the brief period that *SERJ* has been in existence.

2.1. METHODS USED TO ANALYZE A BODY OF RESEARCH

Analyzing the contents of a journal is a challenging, yet important task. One method of quantitatively examining a set of research articles in a particular domain of research is meta-analysis (Rosenthal, 1984). This type of analysis examines a set of related studies that typically provide quantitative measures of effects of experimental treatments, and methods are used to code and synthesize the effects across studies (such as Becker's 1996 meta-analysis). Because meta-analysis focuses on studies that address a single broad question (e.g., what is the effect of cooperative learning on achievement) and uses research results from the *entire published and unpublished* corpus of work, it is not a useful method for analyzing the content of a single journal.

A more traditional method used for analyzing a single journal is content analysis (see Järvelin & Vakkari, 1993; Petrina, 1998). Content analysis is a scholarly method used to study the nature of communications, which can include publications and articles (Carney, 1972). Lasswell (1951), the originator of content analysis, drew on his model of communication to describe the core questions in a content analysis as: *who says what, to whom, why, to what extent, and to what effect.*

Text mining is a newer and more computational method of analyzing content. Text mining typically uses technology and techniques from the disciplines of information processing, computer science, and statistics. The different text mining methods are used to process, extract, and analyze information from text data, such as generating and graphing frequencies of words used, or the co-occurrence of words used. Some of the methods of text mining are referred to as text classification, text clustering, taxonomy creation, document summarization, and latent corpus analysis (Feinerer, Hornik, & Meyer, 2008)

The methods described above are a few approaches that summarize and synthesize a set of related research studies. In addition, there are qualitative methods to prepare a review of the literature on a given topic such as those used in dissertations and thesis papers in education. An analysis of research studies in statistics education offers the unique opportunity to apply the tools of this quantitative discipline (i.e., statistics) to the analysis of a mostly qualitative set of educational research studies. Therefore, the analysis used in this paper combines methods of content analysis and text analysis, as described in the *Methods* section.

2.2. PURPOSE AND FRAMEWORK FOR THE PRESENT STUDY

The purpose of the present study was to review the articles published in *SERJ* from 2002 (the first issue) through 2009, an eight-year period, with a focus on determining where the statistics education research program is today (or at least, at the end of 2009, which is already two years past). The questions offered by Lasswell (1951) in the previous section provided a framework for the analysis of papers published in *SERJ* and simple methods of text mining were used to examine data to analyze these questions:

1. *Who says*: Who are the primary authors publishing in *SERJ*? What is their orientation and discipline? Whom are they working with and where are they working?
2. *What:* What are they saying? What research questions are being studied?
3. *To whom*: What populations are they drawing inferences about?
4. *Why:* What studies are being used as background and theory for the studies? Who are the key people whose work is influencing the studies?
5. *To what extent and to what effect:* How are the studies being conducted? What methods are used? To what extent are they exploratory and descriptive as opposed to studies that can be generalized to larger populations?

To address this last category of questions on research methods, a report written collaboratively by educators and statisticians was drawn upon. In 2007, the American Statistical Association (ASA) published a report written for mathematics education researchers that outlined a set of guidelines for evaluating and reporting mathematics education research. The report, entitled *Using Statistics Effectively in Mathematics Education Research* (*SMER*; ASA, 2007), emerged from a series of three NSF-funded workshops that brought together both statisticians and mathematics education researchers in an effort to improve the quality of mathematics education research through the contributions of modern statistical methods.

3. METHODS

The initial data set consisted of 72 articles published in *SERJ* from 2002–2009 (editorials and announcements were not included as articles). Because the focus of this paper is on the research being published in *SERJ*, eight articles were omitted that were considered to be non-research studies. These included three articles that were introductions to special issues of *SERJ* (Ben-Zvi & Garfield, 2004; Pfannkuch & Reading, 2006; Pratt & Ainley, 2008); three articles from the special issues that were position papers (Gould, 2004; Rossman, 2008; Wild, 2006); as well as two articles that were basically bibliographies appearing in the first volume of *SERJ* while it was making its transition from the *Statistics Education Research Newsletter* (Holmes, 2002; Melitou, 2002). The final useable sample size consisted of 64 articles.

Table 1 shows the data that were collected and coded to address each of the five research questions outlined above. Additionally, data for each of the major research questions are displayed by year to examine what changes were evident across the years that *SERJ* has been published. For each research question, one person did the initial coding of data. All of the coded data were then cross-validated by multiple authors of this paper to help ensure accuracy and minimize errors. When there was disagreement, decisions were made through a consensus building process involving review of the disputed coding by an additional author and a joint decision about the ultimate coding.

Data from each of the 64 articles were recorded in an eXtensible Markup Language (XML) file. XML (see http://en.wikipedia.org/wiki/Extensible_Markup_Language) is a flexible mechanism for representing structured data. It is a generalized version of hypertext markup language (HTML) in which data and metadata are stored in an

Table 1. Research questions and the data collected to answer each of them

Research Questions	Data Collected
1. *Who says*: Who are the primary authors publishing in *SERJ*? What is their orientation and discipline? Whom are they working with and where are they working?	▪ The primary authors ▪ Departments of the primary authors ▪ Institutions of the primary authors ▪ Countries of the primary authors ▪ Collaborations (single/multiple authors) ▪ Collaborations across departments ▪ Collaborations across countries
2. *What:* What are they saying? What research questions are being studied?	▪ Research questions and/or goals ▪ Topics ▪ Keywords used
3. *To whom*: What populations are they drawing inferences about?	▪ Populations being studied
4. *Why:* What studies are being used as background and theory for the studies? Who are the key people whose work is influencing the studies?	▪ References cited
5. *To what extent and to what effect:* How are the studies being conducted? What methods are used? To what extent are they exploratory and descriptive as opposed to studies that can be generalized?	▪ Research classification (*SMER* framework) ▪ Constructs/Instruments ▪ Research methods

electronic text document using tags and attributes. Unlike HTML, the tags and attributes defined in XML are intended to be self-describing of the data embedded in them, and are thus meaningful for information retrieval.

Using the online abstracts and papers, data were collected and entered in the XML file. These data included information about the primary and subsequent authors (name, institution and department affiliation, and country of the institution), the research reported on in the article (research question(s), population being studied, methodology, construct(s) measured, instruments used, and level of *SMER* framework) and metadata regarding the article itself (year of publication, language, abstract, and references cited). More extensive information about how these data were coded and methodologies used are included with the results for ease of their interpretations.

The data from the XML file were extracted using the XML package (Lang, 2009) in R and then analyzed to provide answers to each of the five major research questions regarding statistics education research being published in *SERJ*.

3.1. *SMER* FRAMEWORK CLASSIFICATION

Each of the 64 articles was classified using the five research phases presented in *SMER*: *Generate*, *Frame*, *Examine*, *Generalize*, or *Extend.* The seven authors examined both the abstracts and papers for each article and came to consensus about the classifications through discussion. It was agreed that each article would be classified according to the highest level of the framework that was appropriate, so that only one classification could be used for each article reviewed. The characteristics used to make these classifications and examples are listed below. ASA (2007) provides more detail and examples for each of the categories.

- ***Generate:*** An article generates some ideas about the phenomenon of interest. The ideas might emerge from theoretical considerations, previous research, or observations of practice. An example that falls into this category is Beyth-Marom, Fidler, & Cumming (2008) which introduces and describes statistical cognition as an area of research in statistics education.
- ***Frame:*** An article involves clarification of the goals of the research and definitions of the constructs it entails, development and procedures of the measurement of those constructs, and consideration of the feasibility in putting the ideas into practice. As an example, Zieffler, Garfield, delMas, and Reading (2008) provided the components of a framework to support research on informal inferential reasoning, and the types of tasks suggested by this framework for its development.
- ***Examine:*** An article examines the phenomena more systematically. The purpose of research is to understand the phenomena better and to get indicators of what might work under which conditions. For example, Collins and Mittag (2005) studied the effect of graphing calculators on conceptual understanding in introductory statistics.
- ***Generalize:*** An article seeks to generalize what has been found addressing questions of scale (studying different populations or sites, using more comprehensive measures, examining different implementation conditions), or refining the theory or reframing the entire research. Examples of research in this category include the Student/Teacher Achievement Ratio (STAR) study carried out in Tennessee in which over 7,000 students in 79 schools were randomly assigned into one of three interventions and followed for four years. (See http://www.heros-inc.org/star.htm#Overview for more information on this study.)
- ***Extend:*** An article yielding some generalizable outcomes is extended in a variety of ways—synthesizing multiple studies, examining long-term effects, developing policies for effective implementation. One example from mathematics education research at this level is Reys, Reys, Lapan, Holliday, and Wasman (2003), a multi-institutional follow-up study that compared the mathematics achievement of eighth grade students using NSF-funded Standards-based materials with that of students using traditional materials.

We recorded the different constructs assessed in the study that were classified as *Examine*, *Generalize*, and *Extend.* Where possible, these were taken verbatim from the paper, but often were inferred from our reading of the articles. In addition, each instrument (e.g., surveys, interview protocols, assessments, etc.) that was used and reported in the *SERJ* papers was examined and classified based on whether the instrument was developed *locally* for use in the particular research study, or whether it was developed from previous studies—*existing*. Finally, the online abstracts and methodology sections for these articles were used to categorize the research paradigm employed as qualitative, quantitative, or mixed-methods.

4. RESULTS

Each of the following sections reports the analysis for one of the five research questions used to structure the analyses.

4.1. WHO SAYS?

To examine the primary authors who have published in *SERJ*, data on the names and affiliations (departmental, institutional, and country) of each author were extracted from the XML file. Each of the authors' affiliated departments was categorized into one of seven department types. When a department fit multiple categories, an online investigation was used to aid in the categorization decision based on the primary type of research or work being done in that department. For example, the University of Minnesota's Educational Psychology department was categorized as "education" rather than "psychology."

This section presents the analysis in two parts. The first set of analyses uses the data for only the first (or sole) author of the 64 articles published in *SERJ*. The first author is often the person who contributes the most intellectually to a paper, so it was felt that this information was important to consider in and of itself. We also felt it was important to examine data entailing information about the full set of people contributing to the research presented in *SERJ* (particularly as there are cases where the "main" author, the person who is intellectually driving the research and publication, has his or her name last in the list). Thus, the second set of analyses reported uses data from all authors for the articles. All results represent author data reported in *SERJ* at the time of publication.

First authors There are 52 unique individuals who were first or sole author of a research article in *SERJ*. Because Arthur Bakker has been published while at institutions in both the Netherlands and England the following percentages are based on 53 authors. These authors represent 15 different countries. A majority of these first authors (64%) are from institutions in the English speaking countries of the United States, Australia, the United Kingdom, and New Zealand. The Netherlands, Cyprus, Israel, and Spain are also well represented in *SERJ* (representing an additional 23% of the first authors). Argentina, Belgium, France, Canada, Greece, Ireland, and Mexico are each represented by a single first author. The authors are affiliated with 45 unique institutions. The institutional affiliations most frequently associated with first authors are the University of Minnesota, University of Auckland, University of Tasmania, and University of New England. These four institutions constitute 38% of the institutions affiliated with first authors.

The first authors are affiliated with seven different types of departments. Fifty-three percent of the work published in *SERJ* is by first authors who are affiliated with an education department (including mathematics education) and 17% are from a statistics department.

All authors There have been 112 unique authors published in *SERJ* when all authors (first author, second author, etc.) are considered. These authors represent 15 different countries, where the majority of authors are from institutions in the United States (38%), Australia (10%), and England (13%). Authors from institutions in the Netherlands, New Zealand, and Spain constitute an additional 19%. Argentina, Belgium, Cyprus, France, Israel, Canada, Greece, Ireland, and Mexico make up the rest. It is noted that there is a lack of authors from developing countries. This is in stark contrast to researchers who present their work at conferences such as ICOTS or the IASE/ICME roundtables.

These authors represent 63 unique institutions. Of these institutions, the University of Minnesota (17%), Maastricht University (14%) and the University of New England (13%) were the most represented institutions. Whereas Harvard University represented 17% of institutions, all of the (eleven) authors affiliated with Harvard University were authors on one single article in *SERJ*. Within these institutions, these authors were affiliated with 52 unique departments, the majority of which were categorized as Education (34%) or Statistics (16%).

Collaborations among authors Of the 64 articles analyzed, 66% included multiple authors. Table 2 shows the percentage of single and multiple authored articles by publication year. This table shows that outside of the years 2003 and 2004, the proportion of single author articles and of multiple author articles seems to have stayed pretty stable, with *SERJ* publishing more multiple authored papers. We note that one reason that single authored publications might have been more prevalent in the beginning years of *SERJ* was because the editors invited particular people to submit articles in order to get the journal off the ground.

Table 2. Frequencies (proportions) of SERJ articles each year (2002–2009) that were single or multiple authored

	Publication year								
Author	2002	2003	2004	2005	2006	2007	2008	2009	Total
Single Author	1 (.25)	4 (.57)	6 (.67)	3 (.27)	3 (.38)	1 (.14)	2 (.22)	2 (.22)	22 (.34)
Multiple Authors	3 (.75)	3 (.43)	3 (.33)	8 (.73)	5 (.62)	6 (.86)	7 (.78)	7 (.78)	42 (.66)

Of the multi-authored articles, 45% included cross-departmental collaboration with many of those collaborations (47%) involving Education departments. There were no articles that included cross-departmental collaboration between more than two departments. Thirty-three percent of the multi-authored articles included cross-institutional collaborations.

There was some evidence of collaborations across countries in the authoring of published *SERJ* articles, as 19% of the multi-authored articles included international collaboration with most of these collaborations between authors from Australia and the United States. None of the articles that included international collaboration included authors from more than two countries.

4.2. WHAT?

This section reports the results from examining data related to what is being published in *SERJ*. The research question(s) and/or goal(s) for each article were recorded verbatim from each paper. We opted not to paraphrase or infer what the research question was if it was not directly stated in the article. Because of this, only 44 of the 64 articles could be included in this part of the analysis. Research questions and goals were then categorized into broader categories based on the general topic or types of questions being investigated. A research question/goal that addressed two or more topics (such as using technology and developing statistical reasoning) was categorized under all relevant categories. Figure 1 shows a display of the five categories that emerged during this analysis: *Reasoning/Understanding, Teaching and learning, Affect, Technology*, and *Other*.

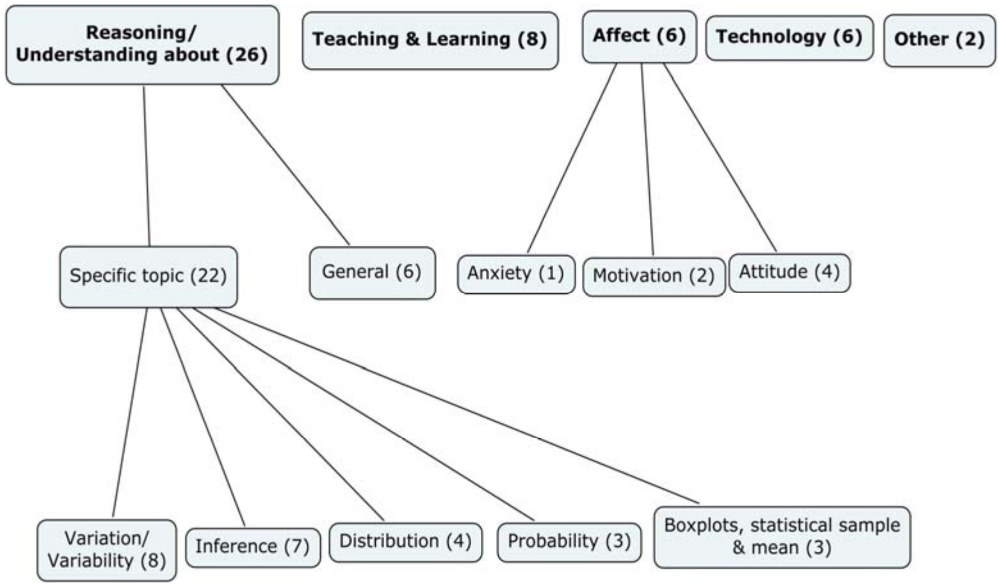

Figure 1. Categories and sub-categories used to group research questions and goals from n = 44 papers in SERJ. Note that a research question/goal that addressed two or more topics was categorized into all related areas and therefore sums of categories and subcategories may exceed totals.

The category with the largest number of related research questions and goals was *Reasoning/Understanding*, which included six sub-categories (five of which were based on specific topics). For example the research goal posed by Reading (2004), "How students describe variation during an inference task" (p. 88), is an example of a research question in this category classified under the specific topic of "variation." The *General* sub-category included research questions or goals about students' general reasoning and understanding that did not deal with a specific statistical concept.

The *Teaching and Learning* category encompassed research questions and goals related to teaching methods and students' learning. An example of a research question in this category was, "Can active or cooperative learning be successfully implemented and accepted in graduate introductory statistics classes?" (Vaughn, 2009, p. 107).

Research questions and goals categorized as *Affect* were related to the examination of students' attitudes, anxiety, or motivation while studying statistics. For example, one research goal categorized under the *Affect* was "to determine if students' level of achievement motivation was a moderating factor on the relationship between students' anxiety and performance" (Keeley, Zayac, & Correia, 2008, p. 6).

The *Technology* category included research questions/goals related to the use of technology in the teaching of statistics. For example, "to see the effect of graphing calculators on conceptual understanding in introductory statistics" (Collins & Mittag, 2005, p. 8) was categorized under this topic. This research question was also categorized under the *General* sub-category within *Reasoning/Understanding*.

The category of *Other* included two research questions that did not fit in any of the other four major categories. The goal of both of the studies categorized as *Other* was to gather information about statistical training in several different specialties in the workplace (e.g., to gather information regarding the status of statistics and its teaching in agricultural colleges).

What keywords are being used to describe the articles? The keywords that the *SERJ* authors used to describe their papers were obtained from the online abstracts and extracted from the XML file. The 64 *SERJ* papers used 303 keywords as descriptors, of which 193 were unique keywords. The most cited keyword—appearing as a descriptor for every article we examined—was "Statistics education research." (Note: Including the keyword "Statistics education research" is *SERJ* policy in order to aid web searches, especially those conducted outside the *SERJ* website.) After the keywords "Statistics," "Statistics education," and "Statistics education research" were excluded from the analysis, the remaining 248 non-unique keywords were categorized using a qualitative strategy suggested by Creswell (1998). Seventeen categories emerged from the data. Table 3 displays the frequency of keywords in each of the 17 categories along with examples of keywords that were categorized. Keywords that were related to subjects or setting represented 14% of the keywords; followed by key words related to a type of research methodology (12%). Among the statistical topics used as keywords, the most frequently used was related to variability (20%).

Table 3. Categorization, examples, and frequencies of the author selected keywords used to describe the 64 SERJ articles

Category	Example keywords	Frequency
Subjects or setting (what or who is being studied)	School students; introductory statistics course	35
Methods	Classroom experiment; interviews	29
Statistical Thinking/ Reasoning/Literacy	Statistical reasoning; statistical literacy; informal reasoning	25
Learning	Active learning; cooperative learning	17
Assessment	Question format; task design	15
Conceptions	Conceptual understanding; misconceptions	13
Affect	Attitudes; motivation	13
Teaching	Classroom instruction; statistics teaching	10
Technology	Graphing calculator; software tools	9
Curriculum	Instructional activities; topic sequencing	8
Statistical topics		61
Variability	Standard deviation; describing variation	12
Inference	Informal statistical inference; hypothesis testing	11
Probability	Venn diagrams; outcome listings	10
Sampling	Sample size; type of sample	3
Randomness	Randomness	2
Other topics	Distribution; context	23
Other	Gender; team projects	13

4.3. TO WHOM?

Data were also collected and extracted on both the sampling unit (e.g., students, teachers) and the level of the sampling unit (e.g., primary, secondary, university-level). For example a study exploring the influence of an instructional software environment on the association between student beliefs and subsequent course performance that used 172 students in an introductory statistics course (Alldredge & Brown, 2006), would be classified as taking place at the tertiary level where the unit of study would be students. Two of the 64 articles were more theoretical in nature and did not have sampling units. Table 4 shows the cross-tabulation of sampling units by level (based on the U.S. schooling system) for the remaining 62 *SERJ* articles. Although students were the most frequently used subjects in the studies reviewed, teachers and textbooks were also the focus of statistics education research studies published in *SERJ*. Other sampling units studied included adults, workers, statistical products (graphs, tables, etc.), instruments, and assessments. It is readily apparent that the largest group of the research papers has focused on tertiary students (39%) followed by 16 articles (26%) at the school (primary and secondary) level. Eighteen percent of the articles examined sampling units across multiple levels.

Table 4. The frequencies of SERJ research articles for each type of sampling unit crossed with level

	Sampling Unit				
Level of Unit	Students	Teachers	Textbooks	Other	Total
Primary (Pre-K to 5th)	4				4
Secondary (6th – 12th)	10	1	1		12
Tertiary	22	1		1	24
Graduate	6	1		1	8
Other				3	3
Multiple Levels	7	1	2	1	11
Total	49	4	3	6	62

4.4. WHY?

Data from each of the 2,466 references cited in the 64 *SERJ* articles from 2002 through 2009 were extracted from the XML file. These data included first author, publication year, title, and publication/conference. Each of these references was then categorized into one of 12 publication types (book, chapter, presentation, proceedings, journal article, manuscript in preparation, unpublished manuscript, master's thesis, Ph.D. thesis, software, web resource, and other [e.g., newsletters, software, personal communications, grants, dissertations, reports, etc.]) based on the status of the reference at the time of citation. Because we did not want to make judgments about whether particular journals, conferences, and so forth were more "prestigious" than others, each reference was given equal weight in this analysis.

The frequencies for these 12 categorizations are shown in Table 5. Not surprisingly the most frequently used category or reference was a journal article, followed by books

Table 5. Publication categorizations and frequencies of the 2,466 references cited in SERJ

Publication type	Frequency
Journal Article	1,076
Book	462
Chapter	338
Proceedings	293
Presentation	112
Ph.D. thesis	48
Software	33
Unpublished manuscript	7
Web resource	6
Master's thesis	5
Manuscript in preparation	2
Other	84

and book chapters.

Figure 2 shows the proportion of references per year for the six most common categories of publication. Three findings emerge from this figure: (1) Every year *SERJ* has been published, journal articles have been the most cited type of reference and the proportion of references that are categorized as journal articles has increased over time; (2) The proportion of references that are categorized as proceedings seems to be decreasing over time; (3) Apart from journal articles and proceedings, the proportion of all other publication types referenced seem to be roughly the same from year to year.

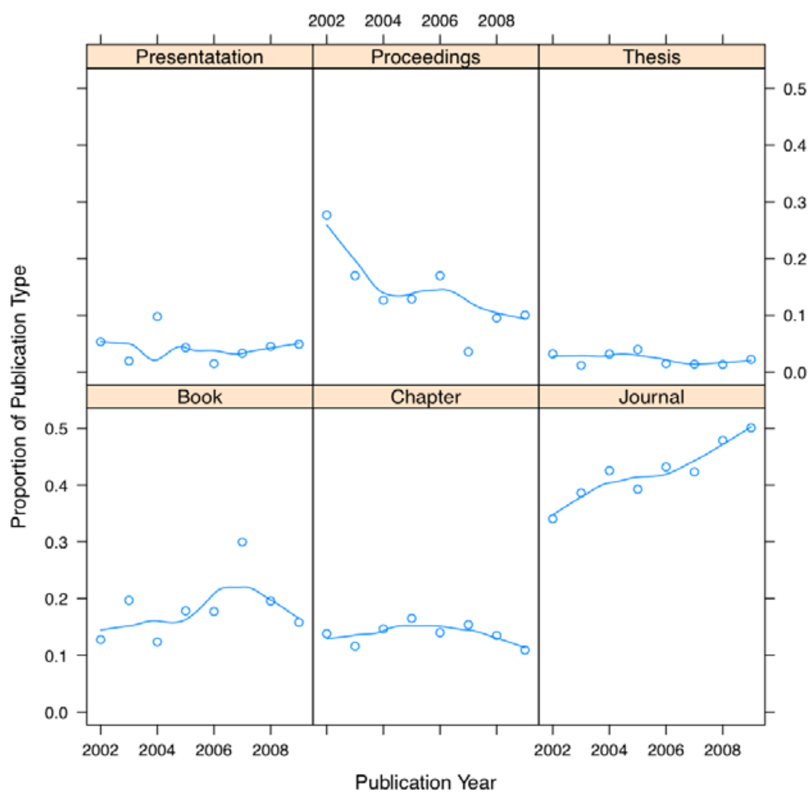

Figure 2. Proportion of references per year categorized by type of publication. The "Thesis" category combines both Ph.D. and master's theses. A LOESS smoother is also plotted to help identify the trend over time.

The references were also classified by discipline of the publication using information obtained through examination of the publication/conference websites, missions, and so forth. For example, an article that was published in the journal *Cognitive Psychology* would be classified under the discipline of psychology. Due to the multi-disciplinary aims and difficulty in specifying a single discipline for certain publication types—such as books, chapters, unpublished manuscripts, manuscripts in preparation, theses, software, and web resources—only journal articles, presentations, and proceedings were classified.

Figure 3 shows the proportion of references (journal articles, presentations, and proceedings) per year classified as a particular discipline. The key features to emerge from this figure are that mathematics education and statistics education seem to be the most cited disciplines in *SERJ* from year to year; mathematics references are rarely cited; there appears to be a slight increasing trend in citations of psychology references; and there appears to be a decreasing trend over time in the proportion of statistics references being cited. These trends may indicate that the discipline of statistics education has close ties to mathematics education, and that psychological aspects of statistics learning are evident in more recent studies. The erratic frequency of education citations is surprising given that it would be reasonable to expect educational and learning theories and principles to form the foundation of statistics education studies.

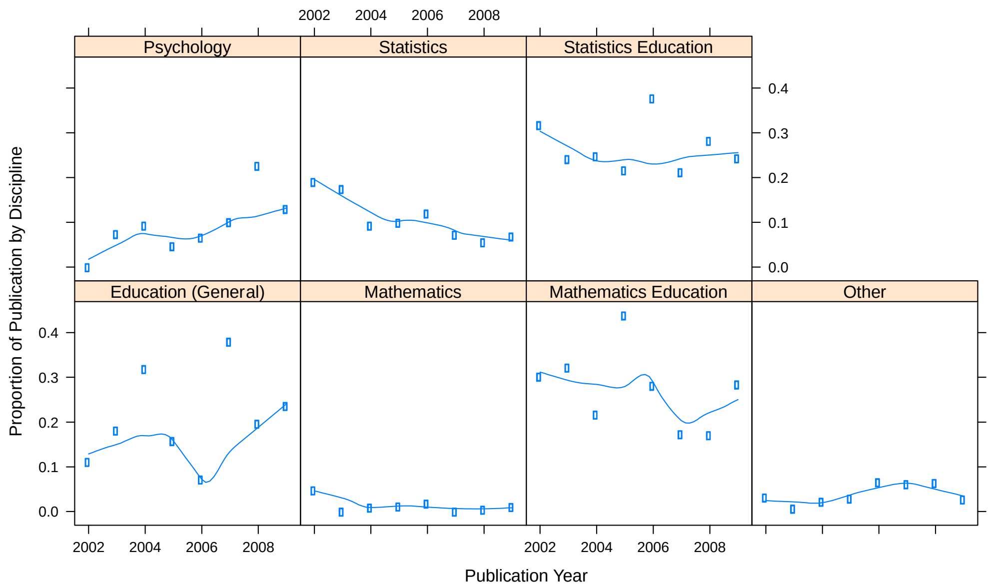

Figure 3. Proportion of references per year classified as a particular discipline. A LOESS smoother is also plotted to help identify the trend over time.

Table 6 presents the ten most cited journals in *SERJ*. These ten journals constitute roughly 48% of the references categorized as journal articles in *SERJ*. Five of the ten most cited journals are classified as statistics education or mathematics education oriented journals. Readers are probably not astonished at the list of journals that are most cited in *SERJ*. The *Journal of Statistics Education* had been in existence nearly 10 years prior to the inception of *SERJ* and was undoubtedly the primary outlet at that time for authors publishing research in statistics education, especially at the tertiary level. Mathematics

Table 6. Most frequently cited journals in SERJ

Journal Name	Discipline	Frequency
Journal of Statistics Education	Statistics Education	94
Journal for Research in Mathematics Education	Mathematics Education	83
Statistics Education Research Journal	Statistics Education	79
Educational Studies in Mathematics	Mathematics Education	66
International Statistical Review	Statistics	56
The American Statistician	Statistics	34
Mathematical Thinking and Learning	Mathematics Education	30
Educational and Psychological Measurement	Education	28
Cognition and Instruction	Psychology	25
Journal of Educational Psychology	Education	18

educators, on the other hand, might have been more inclined to publish their work in either the *Journal for Research in Mathematics Education*, or *Educational Studies in Mathematics*, those being flagship journals in that discipline. *Mathematical Thinking and Learning*, like *SERJ*, is a relatively young journal (started in 1999) so its presence in the top 10 most cited journals may be considered noteworthy. Lastly, we note the diversity of the disciplines represented by these journals.

The five most cited conference proceedings, constituting 73% of the references cited in *SERJ* categorized as proceedings, are presented in Table 7. These conferences, like the journals, are probably not surprising given that they are conferences associated with mathematics and statistics education. There are some conference proceedings that are also noted for their absence in the top five including: *ASA Section on Statistics Education*, *International Conference on Mathematics Education Roundtables*, and the *American Educational Research Association* (AERA). Statistics education research has been reported at each of these conferences.

Table 7. Most frequently cited proceedings in SERJ

Proceedings	Discipline	Frequency
International Conference on Teaching Statistics	Statistics Education	107
IASE Round Table Conference	Statistics Education	39
International Research Forum on Statistical Reasoning, Thinking and Literacy	Statistics Education	27
Conference of the Mathematics Education Research Group of Australasia	Mathematics Education	26
Conference of the International Group for the Psychology of Mathematics Education	Mathematics Education	16

We speculate that one reason for the omission of some proceedings and the inclusion of others in the top five may be related to access. For example, although many statistics education papers have been presented at the AERA conference, there is not a formal published proceedings volume. Other proceedings volumes, such as those from ASA, are primarily available to members of the association and those who attended the conference. The proceedings from the *International Conference on Teaching Statistics*, on the other hand, have been made freely accessible online, which may explain its appearance as the most cited conference proceedings in *SERJ*.

Table 8 presents the ten most frequently cited first authors from these 2,466 references. These authors account for about 20% of the references cited in *SERJ*. There

Table 8. Most frequently cited first authors

Author Name	Frequency
Konold, Clifford	80
Watson, Jane	74
Garfield, Joan	72
Ben-Zvi, Dani	45
Shaughnessy, J. Michael	44
Moore, David	39
Pfannkuch, Maxine	36
Bakker, Arthur	34
Gal, Iddo	33
Cobb, Paul	31

are an additional 20 first authors whose work has been cited ten or more times. Together these 30 authors make up about 33% of the references cited in *SERJ*. These authors have all been contributing to the field of statistics education for many years. The most cited organization as author was the National Council of Teachers of Mathematics (23 citations), followed by the National Research Council (5 citations), and the American Statistical Association (4 citations).

4.5. TO WHAT EXTENT AND TO WHAT EFFECT?

To examine the extent and effect of the published research in *SERJ*, the organizing framework presented in *SMER* was used to situate each of the 64 research studies. Most of the papers published in *SERJ* were categorized into either the second (*Frame*, 72%) or third (*Examine*, 25%) component of this framework. Only 3% of the research published in *SERJ* was classified into the first component of *Generate*, and no articles were considered to be at the levels of *Generalize* or *Extend*. Table 9 displays the frequency of *SERJ* articles by year classified as *Generate, Frame*, or *Examine.*

Table 9. Frequency of SERJ articles per year classified as Generate, Frame, or Examine based on the organizing framework presented in SMER

	2002	2003	2004	2005	2006	2007	2008	2009	Total
Generate		1					1		2
Frame	4	4	9	9	6	4	7	3	46
Examine		2		2	2	3	1	6	16
Total	4	7	9	11	8	7	9	9	64

Using only the 16 research studies that were classified at the *Examine* level, further analysis of the constructs measured was conducted, as well as an analysis of the instruments and research methods that were employed in those articles. (Articles at this level were further analyzed because this level represents the components of a research program in which more systematic studies are conducted.) Below, the abstracts for two articles classified as *Examine* are provided as examples.

> Despite the rapidly growing population of English language learners in U.S. colleges and schools, very little research has focused on understanding the challenges of English language learners specifically in statistics education. At a university near the United States-México border, the authors conducted an exploratory qualitative case study of issues of language in learning statistics for pre-service teachers whose first (and stronger) language is Spanish. The two strongest findings that emerged from cross-case analysis of the interviews were the importance of the role of context (the setting in which information is communicated) and the confusion among registers (subsets of language). This paper overviews and synthesizes relevant literature and offers resources and recommendations for teaching and future research. (Lesser & Winsor, 2009)

> This study examined students’ development of reasoning about quantitative bivariate data during a one-semester university-level introductory statistics course. There were three research questions of interest: (1) What is the nature, or pattern of change in students' development in reasoning throughout the course?; (2) Is the sequencing of quantitative bivariate data within the course associated

> with differences in the pattern of change in reasoning?; and (3) Are changes in reasoning about foundational concepts of distribution associated with differences in the pattern of change? Covariational and distributional reasoning were measured four times during the course, across four cohorts of students. A linear mixed-effects model was used to analyze the data, revealing some interesting trends and relationships regarding the development of covariational reasoning. (Zieffler & Garfield, 2009)

For each of the 16 articles at the *SMER* classification of *Examine*, the research/or analytic method that was employed was also recorded. Using the online abstracts and the paper's methodology section, the research paradigm employed was classified as qualitative, quantitative, or mixed-methods. For example, Lesser and Winsor (2009) was classified as employing a qualitative research paradigm, whereas Zieffler and Garfield (2009) was classified as employing a quantitative research paradigm. In these studies, quantitative methods (50%) were employed more often than qualitative methods (31%) and very few studies utilized mixed-methods (19%).

The instruments utilized in the 16 articles classified as *Examine* were primarily used to measure student attributes, such as reasoning, affect, and demographic variables. These consisted of questionnaires, evaluations, interview protocols, rating scales, and course assessments. The analysis showed an emphasis on instruments used to measure student reasoning and understanding (also seen earlier in Figure 1) followed by a large number of instruments used to examine some aspect of student affect or the relationship between student characteristics and learning statistics.

Each instrument used in these articles was classified as *new or newly developed*—for use in the particular research study or at a particular institution; or *existing*—if it was developed during previous studies. Of the 16 articles, 8 used only *new or newly developed* instruments, 3 used only *existing* instruments, and 3 used a combination of *new or newly developed* and *existing* instruments. (Note: There were 2 articles in which no instrument was used. These were classified as NA.) In the examples above, Lesser and Winsor (2009) used a *new or newly developed* instrument and Zieffler and Garfield (2009) used *existing* instruments.

5. DISCUSSION

This review of research articles published in *SERJ* was undertaken to explore what the *SERJ* articles reveal about the field of statistics education. The analysis focused on who is conducting research in the field, the types of questions addressed by research, and the methodology used in the studies. The five core questions from content analysis (Lasswell, 1951) and the organizing framework of the *SMER* report (ASA, 2007) guided the type of information collected.

Although some of the results of these analyses seem consistent with expectations, they serve as baseline measures of where publication in *SERJ*, and maybe the field of statistics education research, is at currently. We feel that this is a critical first step in what we hope will be an ongoing evaluation of the discipline. We also recognize that there are several potential limitations associated with using the analysis of a single journal, albeit even a flagship journal, to take the pulse of the field. Inherent in considering only the research that has been published in *SERJ* is the possibility of bias on the part of reviewers (and even editors). As an anonymous reviewer of this manuscript noted, "one hopes it doesn't happen, but it has to be acknowledged."

We also recognize that the eight to nine year span of the journal may be a limitation to what types of studies may be published. For example, the lack of studies that fit into either the *generalize* and *extend* components of the *SMER* framework may be an artifact of the fact that research studies at these levels require a prodigious amount of time to conduct. Furthermore, in order for studies to be classified in these categories, randomization is often a requirement. Randomization is often difficult, if not impossible to accomplish in educational settings, and can have ethical implications (see Holcomb, 2002).

Lastly, it is noted that this review has an ending point of 2009. Even as we made changes to this manuscript, the calendar moved on to 2011. In the process, a gap between what has been published and what is included in the manuscript suddenly appeared. This in some sense is unavoidable in any research that undertakes to review what has been done. In order to avoid the infinite loop of continually adding to the review and analysis and never publishing what has been found, a stopping point was identified. It is, however, important to note that in 2010, *SERJ* published a special issue on qualitative methods.

Despite the limitations, some interesting characteristics of the statistics education research published in *SERJ* emerge from the analysis.

Statistics education research is a collaborative and multidisciplinary endeavor.

Unlike many other research domains, there are no departments of statistics education and few courses taught in this specific subject. The research related to statistics education has emerged from different disciplines and has been published in journals in various disciplines that are not connected in any way other than some aspect related to statistical teaching and learning. The review of the authors of *SERJ* articles reflects the different disciplinary affiliations of the authors, as well as a growing collaboration among researchers across departments. A recent report on the development of graduate programs in statistics education (see http://www.causeweb.org/research/programs/) in fact strongly encourages such collaborations in the preparation of future researchers conducting doctoral dissertations. Having *SERJ* as a recognized and prestigious venue for publishing research allows researchers from different disciplines to be become more aware of the scholarship going on across disciplines so that they may draw on and cite a more diverse set of relevant studies in their own research. The amount of collaboration as indicated by the number of co-authored papers is a positive indication of the recognition of the value of collaboration in statistics education research.

Statistics education research is becoming a more coherent, visible domain of inquiry.

We believe the existence of *SERJ* has contributed significantly to the coherence and visibility of research in statistics education. This coherence and visibility is built on the research that is published in *SERJ*, which is strengthened through the peer-review process, the high standards and vision of the *SERJ* editors, and the diverse disciplines and perspectives of the researchers. In addition to the diverse disciplines and perspectives of the researchers, the research published in *SERJ* reveals a great diversity in research approaches and methods. Types of studies range from purely theoretical, to qualitative, to purely quantitative. The diversity in research approaches has consequently led to a variety of ways in which research questions and goals are presented. This diversity can also, in some cases, be a weakness when studies adhering to a certain research paradigm lead to an omission of relevant studies in the literature reviews on that topic. For example, a quantitative study rarely drew on results from a qualitative study in the literature review. *SERJ* would benefit from more research utilizing mixed-methods.

There does seem to be an increase in recent years in articles drawing on particular theoretical frameworks or a body of work on a particular topic such as statistical thinking or reasoning about one important statistical topic, which is helping to make the research more coherent and accessible. With respect to the references cited in research articles, there has been an increasing trend towards a higher proportion of articles cited that have been published in journals with a relative decrease in the citation of conference presentations and proceedings. More recent articles in *SERJ* refer to articles published in statistics education and mathematics education research journals, and fewer refer to journals in other disciplines.

There is, however, a need to broaden the geographic background of the authors in *SERJ*. Authors from only 15 out of around 225 countries in the world have been published in *SERJ*. This is in stark contrast to researchers who present their work at conferences such as ICOTS or the IASE/ICME roundtables. As a comparison, at ICOTS-8 alone, authors and presenters from over 35 distinct countries were represented.

As pointed out earlier, there is a distinct lack of scholarship published in *SERJ* from authors from developing or third world countries. The perspective such authors may bring to the table would be useful in understanding the domain and its areas of inquiry. One way to help remedy this omission would be to identify, approach, and encourage scholars from these backgrounds who present research at ICOTS and the IASE roundtables to write and submit their work to *SERJ*. (However, this may require resources in terms of help with writing in English—or French/Spanish.)

There is a need for more foundational studies to generate questions, frame constructs, and develop assessments.

The analysis of the type of research based on the *SMER* categories indicated that none of the research reported in *SERJ* has been at the *Generalize* or *Extend* levels. This is not surprising, and is probably fitting, given the nascent status of statistics education as a field. It also may be that research that would be classified into these categories may be being published in “high profile” journals in other disciplines, such as the *Journal for Research in Mathematics Education* or *Cognition & Instruction,* rather than in *SERJ*. As *SERJ* continues to grow and become more established, it is likely that it will become a more attractive and prestigious journal to publish the results of these more highly regarded studies.

Most of the research was categorized at the *Frame* or *Examine* levels. Of the 16 studies at the *Examine* level, the largest number of studies investigated statistical reasoning and thinking. Studies that examined the affective domain comprised the second largest number of articles. The lack of articles at the two highest stages and the few articles at the *Examine* level indicates that the field does not currently have a deep understanding of factors that affect statistical learning, thinking, and reasoning. Clearly more systematic research at the *Examine* level is needed before large-scale studies can be launched at the *Generalize* and *Extend* levels with respect to statistical reasoning and thinking or student affect.

There is also a need for instrument development. The analysis of the articles at the *Examine* level suggests that few studies used common or existing research instruments and most seemed to use author-developed tests. Although admittedly, in young fields, author-developed instruments may be a necessity for a time period, the use of course specific instruments (e.g., final exams, course ratings, etc.) is still prevalent. Future authors could really contribute to the field by developing assessments, rating scales, and so forth that could be used across studies, or by using previously developed instruments.

Very few studies published in *SERJ* were at the *Generate* level. One would expect more studies at this level given that statistics education research is still developing as a discipline. However, it may be that the current set of research builds on generative studies that were published prior to the inception of *SERJ*. It is also noteworthy that about half of the studies used only quantitative methods and only about a fifth used a mix of quantitative and qualitative methods. Again, given that statistics education research is still developing as a field, more studies of a qualitative nature might be expected in order to generate ideas about students' learning, reasoning, thinking, and affect that would lead to studies at higher stages within the *SMER* framework. (Recognizing the importance of qualitative research, *SERJ* published a special edition on qualitative approaches to statistics education research in November 2010.)

6. IMPLICATIONS FOR FUTURE RESEARCH

The creation of *SERJ* as a flagship research journal along with its prestige as an official journal of the International Statistical Institute suggest the emergence of statistics education research as a unique and credible discipline. However, there is still much to study, learn, and publish about how students reason, think, and learn about statistics. In particular, there is a compelling need for good ways to describe and measure understanding and reasoning about statistical concepts and problems, as well as promising methodologies to study the developments of these ways of reasoning and thinking. These foundational studies are needed eventually to be able to conduct informative studies at the *Generalize* and *Extend* levels.

About one-third of the research articles did not provide a direct statement of the research question that motivated the study. As a result, the type of research questions addressed in *SERJ* may not be adequately characterized by this study. However, there is also an implication for the future of the discipline. Future research could better build on previous research if the goals and purpose of earlier research were more clearly stated. Arguably, clear research questions are needed to advance the field and our understanding of how to improve knowledge of teaching and learning statistics. Abstracts for articles that clearly state research questions or goals, the use of some common, high quality assessment instruments, and drawing on relevant literature and expertise across disciplines as well as within this newly emerging field, will help make statistics education research a more recognized and coherent discipline.

Lastly, it is noted that the *SERJ* special issues have been very important in leading to coherence and visibility in the research world. Part of this is due to the connections of many of these special issues to a small, international research community (SRTL), which has stimulated and fostered connections among researchers and research studies. Having connected collections of papers in a special issue along with comments to introduce and discuss the papers is an important contribution of these special issues. We encourage additional special issues focused on important ideas and practices in statistics education research.

ACKNOWLEDGEMENTS

We gratefully acknowledge the reviews and feedback given to us on early drafts of this manuscript from Michelle Everson, Beth Chance, Laura Ziegler, and Matthew

Beckman. Furthermore, the feedback from Peter Petocz and that from two anonymous reviewers was instrumental in improving this manuscript.

REFERENCES

Alldredge, J. R., & Brown, G. R. (2006). Association of course performance with student beliefs: An analysis by gender and instructional software environment. *Statistics Education Research Journal*, *5*(1), 64–77.
[Online: www.stat.auckland.ac.nz/~iase/serj/SERJ5%281%29_Alldredge_Brown.pdf]

American Statistical Association. (2007). *Using statistics effectively in mathematics education research* (*SMER Report*). Alexandria: Author.
[Online: www.amstat.org/education/pdfs/UsingStatisticsEffectivelyinMathEdResearch.pdf]

Becker, B. J. (1996). A look at the literature (and other resources) on teaching statistics. *Journal of Educational and Behavioral Statistics*, *21*(1), 71–90.

Ben-Zvi, D., & Garfield, J. (2004). Research on reasoning about variability: A forward. *Statistics Education Research Journal*, *3*(2), 4–6.
[Online: http://www.stat.auckland.ac.nz/~iase/serj/SERJ3%282%29_forward.pdf]

Beyth-Marom, R., Fidler, F., & Cumming, G. (2008). Statistical cognition: Towards evidence-based practice in statistics and statistics education. *Statistics Education Research Journal*, *7*(2), 20–39.
[Online: www.stat.auckland.ac.nz/~iase/serj/SERJ7%282%29_Beyth-Maron.pdf]

Carney, T. F. (1972). *Content analysis: A technique for systematic inference from communications*. Winnipeg, Canada: University of Manitoba Press.

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

Collins, L., & Mittag, K. (2005). Effect of calculator technology on student achievement in an introductory statistics course. *Statistics Education Research Journal*, *4*(1), 7–15.
[Online: http://www.stat.auckland.ac.nz/~iase/serj/SERJ4%281%29_Collins_Mittag.pdf]

Creswell, J. W. (1998). *Qualitative inquiry and research design: Choosing among five traditions*. Thousand Oaks, CA: SAGE Publications.

Feinerer, I., Hornik, K., & Meyer, D. (2008). Text mining infrastructure in *R*. *Journal of Statistical Software*, *25*(5), 1–54.

Garfield, J. (1995). How students learn statistics. *International Statistical Review*, *63*(1), 25–34.

Garfield, J., & Ben-Zvi, D. (2007). How students learn statistics revisited: A current review of research on teaching and learning statistics. *International Statistical Review*, *75*(3), 372–396.

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

Gould, R. (2004). Variability: One statistician's view. *Statistics Education Research Journal*, *3*(2), 7–16.
[Online: http://www.stat.auckland.ac.nz/~iase/serj/SERJ3%282%29_Gould.pdf]

Holcomb, J. (2002). The ethics of comparison: A statistician wrestles with the orthodoxy of a control group. In P. Hutchings (Ed.), *Ethics of inquiry: Issues in the scholarship of teaching and learning* (pp. 19–26). Menlo Park, CA: Carnegie Foundation for the Advancement of Teaching.

Holmes, P. (2002). Some basic references for the teaching of undergraduate statistics. *Statistics Education Research Journal*, *1*(2), 49–53.
[Online: http://www.stat.auckland.ac.nz/~iase/serj/SERJ1%282%29.pdf]

Järvelin, K., & Vakkari, P. (1993). The evolution of library and information science 1965–1985: A content analysis of journal articles. *Information Processing & Management*, *29*(1), 129–144. doi:10.1016/0306-4573(93)90028-C.

Keeley, J., Zayac, R., & Correia, C. (2008). Curvilinear relationships between statistics anxiety and performance among undergraduate students: Evidence for optimal anxiety. *Statistics Education Research Journal*, *7*(1), 4–15.
[Online: http://www.stat.auckland.ac.nz/~iase/serj/SERJ7%281%29_Keeley.pdf]

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

Lang, D. T. (2009). XML: Tools for parsing and generating XML within *R* and *S-Plus*. *R* package version 3.1-0.
[Online: http://CRAN.R-project.org/package=XML]

Lasswell, H. D. (1951). *The political writings of Harold D. Lasswell.* Glencoe, IL: Free Press.

Lesser, L., & Winsor, M. S. (2009). English language learners in introductory statistics: Lessons learned from an exploratory case study of two pre-service teachers. *Statistics Education Research Journal*, *8*(2), 5–32.
[Online: http://www.stat.auckland.ac.nz/~iase/serj/SERJ8%282%29_Lesser_Winsor.pdf]

Meletiou, M. (2002). Teaching students the stochastic nature of statistical concepts in an introductory statistics course. *Statistics Education Research Journal*, *1*(2), 22–37.
[Online: http://www.stat.auckland.ac.nz/~iase/serj/SERJ1%282%29.pdf]

Petrina, S. (1998). The politics of research in technology education: A critical content and discourse analysis of the journal of technology education. *Journal of Technology Education*, *10*(1), 27–57.

Pfannkuch, M., & Reading, C. (2006). Reasoning about distribution: A complex process. *Statistics Education Research Journal*, *5*(2), 4–9.
[Online: http://www.stat.auckland.ac.nz/~iase/serj/SERJ5%282%29_GuestEd.pdf]

Pratt, D., & Ainley, J. (2008). Introducing the special issue on informal inferential reasoning. *Statistics Education Research Journal*, *7*(2), 3–4.
[Online: http://www.stat.auckland.ac.nz/~iase/serj/SERJ7%282%29_Pratt_Ainley.pdf]

Reading, C. (2004). Student description of variation while working with weather data. *Statistics Education Research Journal*, *3*(2), 84–105.
[Online: http://www.stat.auckland.ac.nz/~iase/serj/SERJ3%282%29_Reading.pdf]

Reys, R., Reys, B., Lappan, R., Holliday, G., & Wasman, D. (2003). Assessing the impact of standards-based middle grades mathematics curriculum material on student achievement. *Journal for Research in Mathematics Education*, *34*(1), 74–95.

Rosenthal, R. (1984). *Meta-analytic procedures for social research.* Beverly Hills, CA: Sage.

Rossman, A. (2008). Reasoning about informal statistical inference: One statistician's view. *Statistics Education Research Journal*, *7*(2), 5–19.
[Online: http://www.stat.auckland.ac.nz/~iase/serj/SERJ7%282%29_Rossman.pdf]

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

Shaughnessy, J. M. (2007). Research on statistics learning and reasoning. In F. K. Lester (Ed.), *The second handbook of research on mathematics teaching and learning* (pp. 957–1010). Charlotte, NC: Information Age Publishing.

Statistics Education Research Journal. (n.d.). International Association for Statistics Education.
[Online: http://www.stat.auckland.ac.nz/~iase/publications.php?show=serj]

van der Merwe, L., & Wilkinson, A. (2011). Mapping the field of statistics education research in search of scholarship. *International Journal for the Scholarship of Teaching and Learning*, *5*(1).
[Online: http://academics.georgiasouthern.edu/ijsotl/v5n1/essays_about_sotl/PDFs/_vanderMerweWilkinson.pdf]

Vaughn, B. K. (2009). An empirical consideration of a balanced amalgamation of learning strategies in graduate introductory statistics classes. *Statistics Education Research Journal*, *8*(1), 106–130.
[Online: http://www.stat.auckland.ac.nz/~iase/serj/SERJ8%281%29_Vaughn.pdf]

Wild, C. (2006). The concept of distribution. *Statistics Education Research Journal*, *5*(2), 10–26.
[Online: http://www.stat.auckland.ac.nz/~iase/serj/SERJ5%282%29_Wild.pdf]

Zieffler, A., & Garfield, J. (2009). Modeling the growth of students' covariational reasoning during an introductory statistics course. *Statistics Education Research Journal*, *8*(1), 7–31.
[Online: http://www.stat.auckland.ac.nz/~iase/serj/SERJ8%281%29_Zieffler_Garfield.pdf]

Zieffler, A., Garfield, J., Alt, S., Dupuis, D., Holleque, K., & Chang, B. (2008). What does research suggest about the teaching and learning of introductory statistics at the college level? A review of the literature. *Journal of Statistics Education*, *16*(2).
[Online: http://www.amstat.org/publications/jse/v16n2/zieffler.pdf]

Zieffler, A., Garfield, J., delMas, R., & Reading, C. (2008). A framework to support research on informal inferential reasoning. *Statistics Education Research Journal*, *7*(2), 40–58.
[Online: http://www.stat.auckland.ac.nz/~iase/serj/SERJ7%282%29_Zieffler.pdf]

ANDREW ZIEFFLER
Educational Psychology
University of Minnesota
Room 250 EdSciB
56 E River Road
Minneapolis, MN 55455