Journal of Statistics Education

ISSN: (Print) 1069-1898 (Online) Journal homepage: https://www.tandfonline.com/loi/ujse20

The Challenge of Developing Statistical Reasoning

Joan Garfield

To cite this article: Joan Garfield (2002) The Challenge of Developing Statistical Reasoning, Journal of Statistics Education, 10:3, , DOI: 10.1080/10691898.2002.11910676

To link to this article: https://doi.org/10.1080/10691898.2002.11910676

Published online: 01 Dec 2017.

The Challenge of Developing Statistical Reasoning

Joan Garfield
University of Minnesota

Journal of Statistics Education Volume 10, Number 3 (2002),
ww2.amstat.org/publications/jse/v10n3/garfield.html

Key Words: Assessment; Statistical reasoning.

Abstract

This paper defines statistical reasoning and reviews research on this topic. Types of correct and incorrect reasoning are summarized, and statistical reasoning about sampling distributions is examined in more detail. A model of statistical reasoning is presented, and suggestions are offered for assessing statistical reasoning. The paper concludes with implications for teaching students in ways that will facilitate the development of their statistical reasoning.

1. What is Statistical Reasoning?

Statistical reasoning may be defined as the way people reason with statistical ideas and make sense of statistical information (Garfield and Gal 1999). This involves making interpretations based on sets of data, graphical representations, and statistical summaries. Much of statistical reasoning combines ideas about data and chance, which leads to making inferences and interpreting statistical results. Underlying this reasoning is a conceptual understanding of important ideas, such as distribution, center, spread, association, uncertainty, randomness, and sampling.

Statistical reasoning is a topic of interest to many types of people, including:

- Psychologists, who study how people make judgments and decisions involving statistical information (often using incorrect intuitions or misconceptions),
- Doctors and others in the medical profession, who need to understand and interpret risks, chances of different medical outcomes, and test results,
- Journalists and science writers, who are interested in how to best explain and critique statistical

information in the media,

- Political analysts, who are interested in studying and interpreting polls and elections, and
- Statistics teachers, who want to teach students not only a set of skills and concepts but also how to reason about data and chance.

The expression "statistical reasoning" is widely used and appears in many different contexts. A Web search using the phrase "Statistical Reasoning" produced a list of almost three thousand Web pages that contain the words "statistical" and "reasoning." This list revealed the following categories of Web pages:

- Advertisements for statistics textbooks (that have "statistical reasoning" in their titles or promotional materials),
- Materials from Colleges' or instructors' Web pages for statistics courses offered in a variety of different disciplines (such as mathematics, statistics, psychology, education, engineering, physical therapy, and the health sciences), and
- Presentations, grant proposals, and papers that include discussions of statistical reasoning.

The Web search also produced the home page of the electronic *Journal of Applied Statistical Reasoning* and a book (not a statistics textbook) devoted to the topic of improving statistical reasoning (Sedlmeier 1999).

A quick scan of the materials on the Web that describe courses or textbooks suggests that people are using the term "statistical reasoning" to represent the desired outcomes of a statistics course, and that this expression is used interchangeably with "statistical thinking." There were no clear definitions offered regarding what statistical reasoning means and there did not appear to be clear connections between what was in a course or textbook and the development of particular reasoning skills. For example, some courses were traditional statistics courses with a focus on computations and no use of computing packages. Some courses were more focused on concepts and big ideas, while other courses combined concepts, computation, and computing. Overall, there did not appear to be a consensus in the broad statistics community as to what statistical reasoning means and how to develop this type of reasoning in statistics courses. The next section reviews some of the research literature on statistical reasoning and the ways this term has been used in research studies.

2. Statistical Reasoning in the Research Literature

Chervaney, Collier, Fienberg, Johnson, and Neter (1977) and Chervaney, Benson, and Iyer (1980) defined statistical reasoning as what a student is able to do with statistical content (recalling, recognizing, and discriminating among statistical concepts) and the skills that students demonstrate in using statistical concepts in specific problem solving steps. They viewed statistical reasoning as a three-step process:

- Comprehension (seeing a particular problem as similar to a class of problems),
- Planning and execution (applying appropriate methods to solve the problem), and
- Evaluation and interpretation (interpreting the outcome as it relates to the original problem).

The authors proposed a systems approach for teaching and assessing statistical reasoning based on this model. However, there is no published literature to describe or support the use of this model.

In their unique volume on teaching statistics, Hawkins, Jolliffe, and Glickman (1992) discuss statistical reasoning and thinking together, summarizing some of the research in this area, but not distinguishing between the two types of processes. These authors acknowledge that little is known about the statistical reasoning process or about how students actually learn statistical ideas.

Nisbett (1993) edited a collection of research studies conducted by himself and many of his colleagues on statistical reasoning and ways to train people to follow "rules for reasoning." Early in this collection he offers a set of generalizations based on the research in this area which are summarized below:

1. People have intuitive rules that apply to some statistical problems in everyday life.

2. There are individual differences in the degree to which people understand these rules and apply them to concrete problems.

3. Instruction in statistics changes the way people view the world.

4. The rule systems people use are at the "formal operations" level of abstraction.

5. People may be able to apply statistical rules in one setting (for example, to random generating devices such as dice) but rarely or never to similar problems that involve social content.

6. Training people to apply rules by coding a domain in terms of the rule can improve their use of these rules across domains.

Nisbett concludes that people can improve their statistical reasoning if they learn statistical rules, that some people will understand these rules better than others, that everyone's use of statistical rules can by improved by direct instruction, that this instruction can be abstract, and that people should be taught how to "decode the world" to make it easier for them to apply these rules. He admits that instructors can do a much better job of teaching statistical rules than they currently do.

Research by this author on assessing statistical reasoning (see Garfield 1998a, 1998b), revealed that students can often do well in a statistics course, earning good grades on homework, exams, and projects, yet still perform poorly on a measure of statistical reasoning such as the Statistical Reasoning Assessment (Garfield 1998b). These results suggest that statistics instructors do not specifically teach students how to use and apply types of reasoning. Instead, most instructors tend to teach concepts and procedures, provide students opportunities to work with data and software, and hope that reasoning will develop as a result. However, it appears that reasoning does not actually develop in this way. Current research (see delMas, Garfield and Chance 1999) is focused on exploring and describing the development (and assessment) of statistical reasoning skill, particularly in the area of statistical inference.

Sedlmeier (1999) claims that statistical reasoning is rarely taught and when it is taught (that is, training people to use specific rules such as those described by Nisbett and colleagues), it is rarely successful. He discusses "everyday statistical reasoning," summarizes the research on training to improve statistical reasoning, and presents some training programs of his own designed to teach people to correctly use specific types of reasoning (involving conditional probabilities, samples, and Bayesian inference, for examples).

Lovett (2001) provides a detailed review of the research on statistical reasoning, which she views as falling into one of three approaches: theoretical studies (in the 1970's), empirical studies (in the 1980's), and classroom-based studies (in the 1990's). She describes her own work at Carnegie Mellon University, which is aimed at both understanding and improving students' statistical reasoning, and which integrates all three approaches. She suggests a model for a learning environment to help students develop correct statistical reasoning which will be evaluated in future research studies.

To summarize the research studies on statistical reasoning described above, it appears that inquiry in this area is still evolving. There is no clear consensus about how to help students develop statistical reasoning or how to determine the level and correctness of their reasoning. Perhaps with more classroom-based studies that examine particular types of reasoning, the prerequisite knowledge and skills for each type of reasoning, and the impact of different instructional activities on reasoning, researchers may be better able to understand the process of how correct statistical reasoning develops.

3. Types of Correct and Incorrect Statistical Reasoning

There is an abundance of research on incorrect statistical reasoning, indicating that statistical ideas are often misunderstood and misused by students and professionals alike. Psychologists (such as Kahneman, Slovic, and Tversky 1982) and educators (such as Garfield and Ahlgren 1988) have collected convincing information that shows how people often fail to use the methods learned in statistics courses when interpreting or making decisions involving statistical information. This body of research indicates 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. Some of the types of errors and misconceptions are described below:

Misconceptions involving averages: Averages are viewed as the most common number (the value that occurs more often than the others). People often believe that to find an average one must always add up all the numbers and divide by the number of data values (regardless of outliers). A mean is viewed as the same thing as a median, and there is a belief that one should always compare groups by focusing exclusively on the difference in their averages.

The outcome orientation: An intuitive model of probability that leads students to make yes or no decisions about single events rather than looking at the series of events (Konold 1989). For example, a weather forecaster predicts the chance of rain to be 70% for 10 days. On 7 of those 10 days it actually rained. How good were his forecasts? Many students will say that the forecaster did not do a good job, because it should have rained on all days on which he gave a 70% chance of rain. These students appear to focus on outcomes of single events rather than being able to look at series of events. To students with an outcome orientation, a 70% chance of rain means that it should rain. Similarly, a forecast of 30% rain would mean it will not rain.

Good samples have to represent a high percentage of the population: Most people have strong intuitions about random sampling and most of these intuitions are wrong (Kahneman, et al. 1982). While the Law of Large Numbers guarantees that large samples will be representative of the population from which they are sampled, students' intuitions tell them that it is the ratio of the sample size to the population that is more important to consider. Many believe that it does not matter how large a sample is or how well it was chosen, but that it must represent a large percentage of a population to be a good sample. Therefore, they may be skeptical about a sample that is very large, but represents a small percentage of the population. They do not realize that well-chosen samples do a good job of representing a population, even if the ratio sample size to population size is small.

The "law of small numbers": People believe that samples should resemble the populations from which they are sampled. Many people also believe that any two random samples, regardless of how small they are, will be more similar to each other and to the population than sampling theory would predict. This misconception has led even experienced researchers to use small samples for making inferences and generalizations about populations (Kahneman, et al. 1982).

The representativeness misconception: People estimate the likelihood of a sample based on how closely it resembles the population. Therefore, a particular sequence of *n* tosses of a fair coin that has an approximately even mix of heads and tails is judged more likely than a sequence with more heads and fewer tails. For example, the result HTHHTT is judged as a more likely sequence of 6 tosses of a fair coin than HTHHHH (Kahneman, et al. 1982). Another example of this misconception is the Gambler's Fallacy, which is found in people who believe that after a long series of heads when tossing coins, a tail is more likely to occur on the next toss than another head. This is a fallacy, because if the coin is fair, then the probability of getting a head or a tail on the next toss is equally likely.

The Equiprobability bias: Different outcomes of an experiment tend to be viewed as equally likely. For example, if there are different numbers of science majors and business majors in a class, some students may view as equally likely the outcomes of selecting a science major or selecting a business major when one students is randomly drawn from the class list. Another example is when students are asked to compare the chances of getting different outcomes of three dice rolls, students tend to judge as equally likely the chance of rolling three fives and the chance of obtaining exactly one five. However, the probability of rolling one five is higher than the probability of obtaining three fives, because there are several ways to roll one five, and only one way to roll three fives (Lecoutre 1992).

In contrast to the work of psychologists, who have focused on errors in reasoning and misconceptions, educational researchers have focused on the types of correct reasoning they would like students to develop. Some have designed instructional materials or activities to develop reasoning skills such as the ones described below:

Reasoning about data: Recognizing or categorizing data as quantitative or qualitative, discrete or continuous; and knowing why the type of data leads to a particular type of table, graph, or statistical measure.

Reasoning about representations of data: Understanding the way in which a plot is meant to represent a sample, understanding how graphs may be modified to better represent a data set; being able to see beyond random artifacts in a distribution to recognize general characteristics such as shape, center and spread.

Reasoning about statistical measures: Understanding why measures of center, spread, and position tell different things about a data set; knowing which are best to use under different conditions, and why they do or do not represent a data set; knowing why using summaries for predictions will be more accurate for large samples than for small samples; knowing why a good summary of data includes a measure of center as well as a measure of spread and why summaries of center and spread can be useful for comparing data sets.

Reasoning about uncertainty: Correctly using ideas of randomness, chance, and likelihood to make judgments about uncertain events, knowing why not all outcomes are equally likely, knowing when and why the likelihood of different events may be determined using different methods (such as a probability tree diagram, a simulation using coins, or a computer program).

Reasoning about samples: Knowing how samples are related to a population and what may be inferred from a sample, knowing why a well chosen sample will more accurately represent a population and why there are ways of choosing a sample that make it unrepresentative of the population; knowing to be skeptical of inferences made using small or biased samples.

Reasoning about association: Knowing how to judge and interpret a relationship between two variables, knowing how to examine and interpret a two-way table or scatterplot when considering a bivariate relationship, knowing why a strong correlation between two variables does not mean that one causes the other.

The next section gives an example of one of these areas by describing a series of educational research studies investigating students' reasoning about samples and sampling.

4. The Case of Reasoning about Samples

Over the last several years, I have collaborated with Bob delMas and Beth Chance on research involving students' reasoning about samples and sampling (see Chance, delMas and Garfield, in press). This ongoing research has examined classroom activities that make effective use of computer simulations to shape students' statistical reasoning about foundational topics such as probability, variability, samples, sampling distributions, and the normal distribution. Most of the research has involved studying how students develop correct reasoning as they use carefully designed activities that involve a computer simulation program, Sampling Sim.

This ongoing, collaborative research has led to an effective method of promoting statistical reasoning. This method guides students in using simulation software to make predictions, simulate results, and use these results to test and evaluate our predictions (delMas, et al. 1999). Supported by a grant from the National Science Foundation, we developed software and related instructional materials and investigated their effect on students' reasoning. These materials and software tools may be accessed at the project Web site: www.gen.umn.edu/faculty_staff/delmas/stat_tools.

In a recent study, we conducted interviews with students to gain a more in-depth understanding of students' statistical reasoning about variability, samples, and sampling distributions. The students had completed a graduate-level introductory statistics course in the College of Education and Human Development at the University of Minnesota. These students were majoring in areas such as nursing, social work, and education. Students were interviewed only once. The interviews, which lasted from 45 to 60 minutes, guided the participants through several open-ended questions about variability and sampling. Students were also shown how to use an interactive activity with the Sampling Sim software program.

The interviews were videotaped, transcribed, and viewed many times in an attempt to determine students' initial understanding of how sampling distributions behave and how feedback from the computer simulation program helped students develop an integrated reasoning of key concepts. The researchers identified stages that the students went through as they progressed from faulty to correct reasoning about sampling distributions during their interview (Chance, et al., in press). This led to a proposed framework that describes the development of students' statistical reasoning about sampling distributions (see Table 1). The framework is an extension of one developed by Graham Jones and colleagues to capture the statistical thinking of middle school students (Jones, Langrall, Thornton, and Mogill 1997; Tarr and Jones 1997; Jones, Thornton, Langrall, Putt, and Perry 1998). Current research is now being conducted on the validation and possible extension of this framework to other areas of statistical reasoning and to students at the secondary and postsecondary level (delMas, Garfield, and Chance 2001).

Table 1: Model of Statistical Reasoning

Model of Statistical Reasoning: Sampling Distributions
Level 1. Idiosyncratic reasoning The student knows some words and symbols related to sampling distributions, uses them without fully understanding them, often incorrectly, and may scramble these words with unrelated information.
Level 2. Verbal reasoning The student has a verbal understanding of sampling distributions and the Central Limit Theorem, but can not apply this knowledge to actual behavior. For example, the student can select a correct definition or state the implications of the Central Limit Theorem as it describes sampling distributions, but does not understand how the key concepts such as variability, average, and shape are integrated.
Level 3. Transitional reasoning The student is able to correctly identify one or two dimensions of the sampling process without fully integrating these dimensions. For example, the student only understands the relationship between the sampling distribution shape and the population shape, the fact that large samples lead to more normal looking sampling distributions, or that larger sample size leads to a narrower sampling distribution (decreased variability among sample means).
Level 4. Procedural reasoning The student is able to correctly identify the dimensions of the sampling process but does not fully integrate them nor understand the process that generates sampling distributions. For example, the student can correctly predict which sampling distribution corresponds to the given parameters, but cannot explain the process and does not have confidence in his or her predictions.
Level 5. Integrated process reasoning The student has a complete understanding of the process of sampling and sampling distributions and is able to coordinate the rules (Central Limit Theorem) and behavior of the sampling process. The student can explain the process in her or his own words and makes correct predictions with confidence.

Drawing on current research, theories of learning, and extensive experience as statistics teachers, we propose that in order for students to fully understand and reason about sampling distributions at the highest level, they need to experience a variety of activities: text or verbal explanations, concrete activities involving sampling from finite populations, and interactions with simulated populations and sampling distributions when the parameters are varied. This model contradicts some researchers in psychology who argued for teaching specific training rules. An example of how to help students develop statistical reasoning about samples and sampling is described below:

1. The students read assigned materials about samples before class meets, and begin class with a general discussion about the reading.

2. An in-class, hands-on activity is used where students take samples from a finite population, varying size of samples and type of sampling method used. For example, the "Random Rectangles Activity" or "Cents and the Central Limit Theorem" from *Activity Based Statistics* (Scheaffer,

Gnanadesikan, Watkins, and Witmer 1996) may be used. After students gather the data, the results are pooled and used to compare the effects of changing sample size and the different results caused by sampling techniques such as random or judgmental samples.

3. Students use simulations on a computer (such as the "Sampling Sim" software) to study the effects of changing parameters on samples and sampling distributions. An effective model is one where students make predictions, test them out with simulated data, and then evaluate the correctness of their predictions (see delMas, Garfield, and Chance 1998).

4. There is a class discussion about what students learned from the simulation activity, highlighting the key outcomes observed when drawing repeated samples while changing sample size and population parameters.

5. Students are assessed to determine their ability to reason about samples and sampling distributions. An example of a post-test to use in assessing student reasoning appears at the end of this paper.

Although the model described in Table 1 was focused on reasoning about samples and sampling distributions, it may be applied to other types of statistical reasoning, as shown in Table 2. This model also has implications for assessment of students' levels of statistical reasoning.

Table 2: General Model of Statistical Reasoning

Model of Statistical Reasoning
Level 1. Idiosyncratic reasoning The student knows some statistical words and symbols, uses them without fully understanding them, often incorrectly, and may scramble them with unrelated information. For example, students have learned the terms mean, median, and standard deviation as summary measures, but use them incorrectly (for example, comparing the mean to the standard deviation, or making judgments about a good mean or standard deviation).
Level 2. Verbal reasoning The student has a verbal understanding of some concepts, but cannot apply this to actual behavior. For example, the student can select or provide a correct definition but doesn't fully understand the concepts (for example, why the mean is greater than the median in positively skewed distributions).
Level 3. Transitional reasoning The student is able to correctly identify one or two dimensions of a statistical process without fully integrating these dimensions, such as, that a larger sample size leads to a narrower confidence interval, that a smaller standard error leads to a narrower confidence interval.
Level 4. Procedural reasoning The student is able to correctly identify the dimensions of a statistical concept or process but does not fully integrate them or understand the process. For example, the student knows that correlation does not imply causation but cannot fully explain why.
Level 5. Integrated process reasoning The student has a complete understanding of a statistical process, coordinates the rules and behavior. The student can explain the process in his or her own

words with confidence. For example, a student can explain what a 95% confidence interval means in terms of the process of repeatedly sampling from a population.

5. Assessing Statistical Reasoning

Most assessment instruments used in research studies of statistical reasoning and understanding consist of items presented to students or adults individually as part of clinical interviews or in small groups. In contrast, traditional paper-and-pencil assessment instruments often focus on computational skills or problem solving rather than on reasoning and understanding. Questions and task formats that culminate in simple “right or wrong” answers do not adequately reflect the nature of students’ thinking and problem solving, and therefore provide only limited information about students’ statistical reasoning processes and their ability to construct or interpret statistical arguments (Gal and Garfield 1997).

Garfield and Chance (2000) offer some suggestions for classroom assessment techniques to evaluate students’ statistical reasoning. These include:

Case studies or authentic tasks: detailed problems based on a real context that reveal students’ strategies and interpretations as they solve the problem.

Concept maps: visual representations of connections between concepts that students may either complete or construct on their own.

Critiques of statistical ideas or issues in the news: short written reports to reveal how well students reason about information provided in a news article, including comments on missing information as well as conclusions and interpretations offered in the article.

Minute papers: brief, anonymously written remarks provided by students that may include explanations of what they have learned, comparisons of concepts or techniques, etc.

Enhanced multiple-choice items: items that require students to match concepts or questions with appropriate explanations, may be used to capture students’ reasoning and measure conceptual understanding

The Statistical Reasoning Assessment: a multiple-choice test designed to assess students’ correct and incorrect reasoning for a sample of statistical concepts (Garfield 1998a, 1998b).

Informal methods may also be used during class activities, such as asking students to provide written or verbal interpretations of data, explanations of concepts, or matching different types of representations (for example, matching boxplots to histograms, or matching graphs to statistics, as shown in Scheaffer et al. 1996). These methods can help inform an instructor about the level of students’ statistical reasoning about particular concepts or procedures, which may be quite different from students’ ability to compute and carry out procedures. The contrast in these two types of assessments can be quite startling as well as enlightening, and may suggest that additional instructional activities are needed.

6. Summary: Implications for Teachers

Although the term “statistical reasoning” is often used in different ways, it appears to be universally accepted as a goal for students in statistics classes. It has been shown that statistical reasoning used in everyday life as well as in classes is often incorrect, due to the different intuitions and rules that people use when evaluating statistical information.

While most instructors would like their students to develop statistical reasoning, research shows that it is not enough to instruct students about the correct rules and concepts in order for them to develop an integrated understanding to guide their reasoning. It may be tempting to conclude that if students have been well taught and have performed well on exams, that they are able to reason correctly about statistical information. However, unless their reasoning is carefully examined, especially in applied contexts, these students may only be at the early stages of reasoning and not have an integrated understanding needed to make correct judgments and interpretations.

Activities specifically designed to help develop students’ statistical reasoning should be carefully integrated into statistics courses. For example, activities such as having students match verbal descriptions to plots of data (see Rossman and Chance 2001) can help develop reasoning about data and distribution, and activities that challenge students to consider what makes a standard deviation larger or smaller (see delMas 2001) can help students develop reasoning about variability. Similarly, activities that guide students through visual simulations of sampling distributions, varying sample size and population parameters, can help develop students' reasoning about sampling distributions. In order to better understand how well students are developing and using statistical reasoning, statistics teachers may include assessments that measure students’ reasoning so that they are not misled in their evaluation of student learning (that is, high scores on computational procedures).

Recent research on developmental models of statistical reasoning may help instructors better understand the process of developing correct statistical reasoning and guide them in developing instructional methods and assessments. More research is needed, especially classroom research conducted in a variety of settings, to help determine how instructional methods and materials may best be used to help students develop correct statistical reasoning.

References

Chance, B., delMas, R., and Garfield, J. (in press) "Developing reasoning about sampling distributions," in *The Challenge of Developing Statistical Literacy, Reasoning and Thinking*, eds. D. Ben-Zvi and J. Garfield, Dordrecht, The Netherlands: Kluwer Publishers.

Chervaney, N., Benson, P.G., and Iyer, R. (1980), "The planning stage in statistical reasoning," *The American Statistician*, 34, 222-226.

Chervaney, N., Collier, R. Fienberg, S., Johnson, P, and Neter, J. (1977), "A framework for the development of measurement instruments for evaluating the introductory statistics course," *The American Statistician*, 31, 17-23.

delMas, R. (2001), "What makes the standard deviation larger or smaller?" *Statistics Teaching and Resource Library* [Online]. (www.starlibrary.net)

delMas, R., Garfield, J., and Chance, B. (1998), "Assessing the effects of a computer microworld on statistical reasoning," in *Proceedings of the Fifth International Conference on Teaching Statistics*, ed. L.

Pereira-Mendoza, Voorburg, The Netherlands: International Statistical Institute, 1083-1090.

----- (1999), "A Model of Classroom Research in Action: Developing Simulation Activities to Improve Students' Statistical Reasoning," *Journal of Statistics Education* [Online], 7(3). (ww2.amstat.org/publications/jse/secure/v7n3/delmas.cfm)

----- (2001), "Validating types of reasoning about sampling distributions," Presentation at the Second International Research Forum on Statistical Reasoning, Thinking and Literacy, Armidale, Australia.

Gal, I., and Garfield, J. (eds.) (1997), *The Assessment Challenge in Statistics Education*, Amsterdam: IOS Press and the International Statistical Institute.

Garfield, J. (1998a), "Challenges in Assessing Statistical Reasoning," Paper presented at the American Educational Research Association annual meeting, San Diego, CA.

----- (1998b), "The Statistical Reasoning Assessment: Development and Validation of a Research Tool," in *Proceedings of the Fifth International Conference on Teaching Statistics*, ed. L. Pereira-Mendoza, Voorburg, The Netherlands: International Statistical Institute, 781-786.

Garfield, J., and Ahlgren, A. (1988), "Difficulties in learning basic concepts in statistics: implications for research," *Journal for Research in Mathematics Education*, 19, 44-63.

Garfield, J., and Chance, B. (2000), "Assessment in Statistics Education: Issues and Challenges," *Mathematics Thinking and Learning*, 2, 99-125.

Garfield, J., and Gal, I. (1999), "Teaching and Assessing Statistical Reasoning," in *Developing Mathematical Reasoning in Grades K-12*, ed. L. Stiff, Reston, VA: National Council Teachers of Mathematics, 207-219.

Hawkins, A., Jolliffe, F., and Glickman, L. (1992), *Teaching Statistical Concepts*, London: Longman Publishers.

Jones, G. A., Langrall, C. W., Thornton, C. A., and Mogill, A. T. (1997), "A framework for assessing and nurturing young children's thinking in probability," *Educational Studies in Mathematics*, 32, 101-125.

Jones, G. A., Thornton, C. A., Langrall, C. W., Putt, I. J., and Perry, B. (1998), "A framework for assessing and fostering students' statistical thinking," Manuscript submitted for publication.

Kahneman, D., Slovic, P. and Tversky, A. (1982), *Judgement Under Uncertainty: Heuristics and Biases*, Cambridge, UK: Cambridge University Press.

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

Lecoutre, M. P. (1992), "Cognitive models and problem spaces in 'purely random' situations," *Educational Studies in Mathematics*, 23, 557-568.

Lovett, M. (2001), "A collaborative convergence on studying reasoning processes: A case study in statistics," *Cognition and Instruction: Twenty-Five Years of Progress*, eds. D. Klahr and S. Carver, Mahwah, NJ: Lawrence Erlbaum, 347-384.

Nisbett, R. (1993), *Rules for Reasoning*, Mahwah, NJ: Lawrence Erlbaum.

Rossman, A. J., and Chance, B. L. (2001), *Workshop Statistics: Discovery with Data*, New York: Springer Verlag.

Scheaffer, R., Gnanadesikan, M., Watkins, A., and Witmer, J. (1996), *Activity-Based Statistics: Instructor Resources*, New York: Springer Verlag.

Sedlmeier, P. (1999), *Improving Statistical Reasoning: Theoretical Models and Practical Implication*, Mahwah, NJ: Lawrence Erlbaum.

Tarr, J. E., and Jones, G. A. (1997), "A framework for assessing middle school students' thinking in conditional probability and independence," *Mathematics Education Research Journal*, 9, 39-59.

Joan Garfield
Department of Educational Psychology
University of Minnesota
Minneapolis, MN 55455
USA
jbg@umn.edu

Volume 10 (2002) | Archive | Index | Data Archive | Information Service | Editorial Board | Guidelines for Authors | Guidelines for Data Contributors | Home Page | Contact JSE | ASA Publications