A FRAMEWORK TO SUPPORT RESEARCH ON INFORMAL INFERENTIAL REASONING

ANDREW ZIEFFLER
University of Minnesota
zief0002@umn.edu

JOAN GARFIELD
University of Minnesota
jbg@umn.edu

ROBERT DELMAS
University of Minnesota
delma001@umn.edu

CHRIS READING
University of New England
creading@une.edu.au

ABSTRACT

Informal inferential reasoning is a relatively recent concept in the research literature. Several research studies have defined this type of cognitive process in slightly different ways. In this paper, a working definition of informal inferential reasoning based on an analysis of the key aspects of statistical inference, and on research from educational psychology, science education, and mathematics education is presented. Based on the literature reviewed and the working definition, suggestions are made for the types of tasks that can be used to study the nature and development of informal inferential reasoning. Suggestions for future research are offered along with implications for teaching.

Keywords: *Statistics education research; Inference; Informal reasoning; Introductory statistics course; Topic sequencing*

1. INTRODUCTION

> Statistics is concerned with the gathering, organization, and analysis of data and with inferences from data to the underlying reality. (Moore, 1990, p. 127)

Drawing inferences from data is part of everyday life and critically reviewing results of statistical inferences from research studies is an important goal for most students who enroll in an introductory statistics course. Formal methods of statistical inference lead to drawing conclusions about populations or processes based on sample data. David Moore (2004) describes statistical inference as moving beyond the data in hand to draw conclusions about some wider universe, taking into account that variation is everywhere and the conclusions are therefore uncertain. Garfield and Ben-Zvi (2008) define statistical inference further by differentiating two important themes in statistical inference,

Statistics Education Research Journal, 7(2), 40-58, http://www.stat.auckland.ac.nz/serj

parameter estimation and hypothesis testing, and two kinds of inference questions, generalization (from samples) and comparison and determination of cause (from randomized comparative experiments). In general terms, the first theme is concerned with generalizing from a small sample to a larger population, whereas the second involves determining whether a pattern in the data can be attributed to a real effect.

For several decades, psychologists and education researchers have studied and documented the difficulties people have making inferences about uncertain outcomes (see Kahneman, Slovic, & Tversky, 1982). Falk and Greenbaum (1995) and others (e.g., Sotos, Vanhoof, Van den Noortgate, & Onghena, 2007) have described and classified difficulties students have in understanding and interpreting tests of significance and p-values and related concepts in statistical inference. Researchers have pointed to various reasons for these difficulties including: the logic of statistical inference (e.g., Cohen, 1994; Nickerson, 2000; Thompson, Saldanha, & Liu, 2007), students' intolerance for ambiguity (Carver, 2006), and students' inability to recognize the underlying structure of a problem (e.g., Quilici & Mayer, 2002). Other research has suggested that students' incomplete understanding of foundational concepts such as distribution (e.g., Bakker & Gravemeijer, 2004), variation (e.g., Cobb, McClain, & Gravemeijer, 2003), sampling (e.g., Saldanha & Thompson, 2002, 2006; Watson, 2004), and sampling distributions (e.g., delMas, Garfield, & Chance, 1999; Lipson, 2003) may also play a role in these difficulties.

Given the importance of understanding and reasoning about statistical inference, and the consistent difficulties students have with this type of reasoning, there have been attempts to expose students to situations that allow them to use informal methods of making statistical inferences (e.g., comparing two groups based on boxplots of sample data). Several papers have been presented and published in the past few years that describe 'informal statistical inference' and 'informal inferential reasoning' (e.g., Pfannkuch, 2005). However, it is not yet clear exactly what these two terms mean. Therefore, it is the intent of this paper to analyze the meaning of Informal Inferential Reasoning (IIR) by reviewing the literature related to this topic, and to provide both a working definition as well as a framework for designing tasks that can be used to study students' reasoning about statistical inference.

Cognitive frameworks have been useful in studying and describing students' statistical reasoning (see Jones, Langrall, Mooney, & Thornton, 2005). These models offer benchmarks for assessing students' reasoning and are useful for informing the development of assessment tasks, guiding teachers' instructional decision-making, and developing tasks to use in research programs. The main focus of this paper is to propose a preliminary framework that, although not a developmental model, can be used to identify and develop tasks that can be used to study IIR. The two main questions addressed in the paper are

1. What are the components of a framework needed to support research on informal inferential reasoning?
2. What types of tasks are suggested by this framework for the study of informal inferential reasoning and its development?

2. WHAT ARE THE COMPONENTS OF A FRAMEWORK NEEDED TO SUPPORT RESEARCH ON INFORMAL INFERENTIAL REASONING?

The previous section described the nature of statistical inference and the way concepts and procedures involved in statistical inference are often introduced in introductory statistics courses. In an attempt to understand the component parts of informal inferential reasoning (IIR), we look to the literature to identify some foundational areas of research. Because IIR uses the word "informal" it seemed useful to explore research in two possibly related areas of research in psychology and education: studies of informal knowledge and studies of informal reasoning. The research in both of these areas provides a foundation for understanding and defining IIR. This section begins with a brief review of the research in the areas of informal knowledge and then informal reasoning. This is followed by a review of the use of the terms "informal inference" and "informal inferential reasoning" by statistics educators and statistics education researchers in recent papers. The section concludes with a working definition of IIR based on the literature reviewed in these three areas.

2.1. INFORMAL KNOWLEDGE

There is much research on the nature of informal knowledge, particularly in the field of mathematics education. Informal knowledge is viewed as either a type of *everyday real world knowledge* that students bring to their classes based on out-of-school experiences, or a less formalized knowledge of topics resulting from prior formal instruction. Informal knowledge can be viewed as the integration of both of these and it is in this sense that the term is used in this paper. This view suggests that it is important to study and consider the role of informal knowledge in the formal study of a particular topic and is in line with constructivist views of learning, namely that informal knowledge is a starting point for the development of formal understanding.

Informal knowledge is also discussed in the literature on how experts use their informal knowledge in reasoning while solving problems, and in studies that compare the reasoning of experts and novices. Smith, diSessa and Rochelle (1993/1994) found that experts appear to differ from novices primarily in the extent of their experience with problems in a particular context. Although experts seemed to see a coherent picture of a problem they were solving, novices were more likely to approach each problem in a similar group from a different starting point, rather than seeing how the problems were related. Experts also tended to draw on their more extensive experiences and knowledge to use problem-solving strategies that related to the underlying structures of the problems more often than novices. For a more detailed synthesis of the research on experts versus novices see Bransford, Brown, and Cocking (2000).

When informal knowledge is incorrect, it is often regarded as a misconception (see Confrey, 1990). Smith et al. (1993/1994) argued that it is not beneficial to view students' "misconceptions" as wrong, because there are aspects of students' informal knowledge that are similar to experts' knowledge, but have been used incorrectly. Novice reasoning tends to have the same basic structure as expert reasoning, but novice reasoning often appears more concrete and less abstract due to novices' more limited experience. Instead of focusing only on the development of formal knowledge, Smith et al. suggest that instructors carefully design lessons that build and develop students' informal knowledge in order to lead them toward formal understanding of a particular topic.

Along the same lines, Gravemeijer and Doorman (1999) argue that it is important to have students build on their informal knowledge to reinvent formal concepts and representations and at the same time expand their common sense understanding of real world phenomena. This approach acknowledges, rather than discounts (e.g., by labeling erroneous use of knowledge as misconceptions), the informal knowledge that students bring to the classroom.

An important question emerges: How can students' informal knowledge best be utilized in formal instruction? Some researchers point to the role of interactive activities where students work and discuss together what they are learning. It has been suggested and demonstrated that social interaction that requires negotiation of meaning, under the direction of shared social norms for communication helps support the transformation of informal knowledge to culturally shared formal understanding (Cobb & McClain, 2004; Cobb, Yackel, & Wood, 1992; Mack, 1995).

Another way of developing students' informal knowledge is to specifically evolve this type of knowledge through activities that motivate and "set the stage" for formal instruction at a later time (see for example, Papert & Harel, 1991). Schwarz, Sears, and Chang (2007) have found positive results in their attempts to explicitly develop and utilize students' prior knowledge as they learn specific statistical concepts. Garfield, delMas, and Chance (2007) have also found some success in developing college students' formal ideas of variability from informal ideas.

In summary, the literature reviewed suggests that

1. Informal knowledge can consist of different types of understanding that students bring to a new learning task, and may combine knowledge based on real world experience with knowledge gained from previous instruction (Gravemeijer & Doorman, 1999; Smith et al., 1993/1994).
2. Informal knowledge may be an important starting point on which to build formal knowledge, and should be considered in designing curricula (Gravemeijer & Doorman, 1999; Smith et al., 1993/1994).
3. Instruction may be designed to help students construct specific types of informal knowledge that is needed for eventual instruction involving formal knowledge of a particular concept (Garfield, delMas, & Chance, 2007; Schwarz, Sears, & Chang, 2007).
4. Activity-based learning that requires social interaction and the negotiation of meaning can facilitate the development of informal knowledge (Cobb & McClain, 2004; Cobb, Yackel, & Wood, 1992; Mack, 1995).

This review of the nature of informal knowledge suggests that developing students' informal knowledge related to statistical inference may ease their transition to understanding formal ideas of inference.

2.2. INFORMAL REASONING

Informal reasoning (sometimes referred to as informal logic) has been defined by cognitive psychologists as the type of reasoning that occurs in non-deductive situations, such as decision making, that is employed in everyday life (Voss, Perkins, & Segal, 1991). Perkins, Farady, and Bushey (1991) characterize informal reasoning as "a process of *situation modeling*" in which a person builds a model of the situation in question by "articulating the dimensions and factors involved … and invok[ing] a variety of common sense, causal, and intentional principles both to construct and to weigh the plausibility of alternative scenarios" (p. 85).

Formally, there is very little agreement in the literature as to what is meant by 'informal reasoning'. This may be due to the reliance of informal reasoning on context or subject matter (Perkins, 1985b; Walton, 1989). There are, however, two commonalities across a majority of the papers reviewed. First, informal reasoning is most often viewed through the lens of argumentation theory (e.g., Kuhn, 1991; Means & Voss, 1996; Sadler, 2004; Sadler & Zeidler, 2004). Secondly, informal reasoning is often contrasted to formal reasoning or logic (e.g., Evans, Newstead, & Byrne, 1993; Miller-Jones, 1991; Pfannkuch, 2006; Schoenfeld, 1991).

Researchers tend to study informal reasoning through dialogical argumentation–the expression or means by which researchers gain access to informal reasoning (e.g., Driver, Newton, & Osborne, 2000; van Eemeren et al., 1996). The assessment of informal reasoning through argumentation, which draws heavily from Toulmin's (1958) model of argumentation, has led to many findings. For example, Perkins et al. (1991) have found through a series of studies that informal reasoning, despite claims to the contrary, "is marred by incompleteness and bias is the norm rather than the exception" (p. 90). Sadler and Zeidler (2004) caution that "while it is valid to assert that strong argumentation reveals strong informal reasoning, the opposite claim, weak argumentation denotes weak informal reasoning, is not necessarily the case … naïve arguments might be the result of either insufficient informal reasoning or poorly articulated, but proficient informal reasoning" (p. 73).

In summary, the literature reviewed suggests that

1. The quality of informal reasoning does not necessarily improve with increased content knowledge (e.g., Kuhn, 1991; Means & Voss, 1996; Perkins, 1985b; Perkins et al., 1991);
2. Informal reasoning is unlikely to improve with maturation, education, or life experience (e.g., Klahr, Fay, & Dunbar, 1993; Kuhn, Garcia-Mila, Zohar, & Andersen, 1995; Perkins, 1985b; Perkins et al., 1991; Schauble, 1990, 1996; Schauble & Glaser, 1990; Voss, Blais, Means, Greene, & Ahwesh, 1986);
3. Motivation, or interest in the problem context, has little impact on informal reasoning quality (e.g., Perkins, 1989; Perkins et al., 1991);
4. General intelligence influences people's informal reasoning, but people selectively use that intelligence to build their own case rather than to explore an issue more fully (e.g., Perkins, 1985a; Perkins, 1989; Perkins et al., 1991);
5. Informal reasoning is a matter of "know-how" and can be improved through instruction (e.g., Nickerson, Perkins, & Smith, 1985; Perkins, Bushey, & Farady, 1986; Perkins et al., 1991; Schoenfeld, 1982; Schoenfeld & Herrmann, 1982).

Informal reasoning seems to be an important part of IIR because of the role of evidence and argumentation in making statistical predictions and decisions.

2.3. DEFINING INFORMAL INFERENTIAL REASONING

As mentioned earlier, IIR is a relatively recent concept in the research literature and various definitions have been presented. Rubin, Hammerman, and Konold (2006) define IIR as reasoning that involves the related ideas of properties of aggregates (e.g., signal and noise, and types of variability), sample size, and control for bias. Pfannkuch (2006) defines IIR as the ability to interconnect ideas of distribution, sampling, and center, within an empirical reasoning cycle (Wild & Pfannkuch, 1999). Bakker, Derry, and Konold (2006) suggest a theoretical framework of inference that broadens the meaning of statistical inference to allow more informal ways of reasoning and to include human judgment based on contextual knowledge. One statistician has described informal inference as "going beyond the data at hand" and "seeking to eliminate or quantify chance as an explanation for the observed data" through a reasoned argument that employs no formal method, technique, or calculation (Rossman, 2007). Ben-Zvi (2006) compares inferential reasoning to argumentation, and emphasizes the need for this type of reasoning to include data-based evidence.

These different definitions of IIR share many things in common. In an attempt to combine these perspectives, we present a working definition of informal inferential reasoning as *the way in which students use their informal statistical knowledge to make arguments to support inferences about unknown populations based on observed samples.* In addition, building on their own experiences and understanding from teaching formal ideas and methods of statistical inference to college students, the authors see informal inferential reasoning as a process that includes

- Reasoning about possible characteristics of a population (e.g., shape, center) based on a sample of data;
- Reasoning about possible differences between two populations based on observed differences between two samples of data (i.e., are differences due to an effect as opposed to just due to chance?); and
- Reasoning about whether or not a particular sample of data (and summary statistic) is likely (or surprising) given a particular expectation or claim.

This is in contrast to formal statistical inferential reasoning, which may include significance tests and/or confidence intervals. For example, the type of formal reasoning about a one-sample test of significance that we hope to help students eventually develop requires an understanding of the interconnections between

- An underlying theory or hypothesis that is to be tested;
- A sample of data that can be examined; and
- A distribution of a statistic for all possible samples under the assumption that the theory or hypothesis is true.

This integration involves comparing the observed sample statistic to the distribution of statistics for all possible samples to see how unlikely the occurrence is (i.e., how far out in either of the tails it falls). The farther out in one of the tails, the less plausible it is that the observed results are due to chance and, therefore, the more convincing that there is a true difference or effect. This formal reasoning also includes the understanding of a p-value as an indicator of how likely or surprising a sample result, or a result more extreme, is under a certain hypothesis, and the action of rejecting this hypothesis if the p-value is small enough.

In summary, the IIR Framework has the following three components:

1. Making judgments, claims, or predictions about populations based on samples, but not using formal statistical procedures and methods (e.g., p-value, *t* tests);
2. Drawing on, utilizing, and integrating prior knowledge (e.g., formal knowledge about foundational concepts, such as distribution or average; informal knowledge about inference such as recognition that a sample may be surprising given a particular claim; use of statistical language), to the extent that this knowledge is available; and
3. Articulating evidence-based arguments for judgments, claims, or predictions about populations based on samples.

Note that this definition refers to IIR as a process for making inferences that does not utilize the formal methods of statistical inference described earlier and that may or may not include use of formal statistical concepts or language.

2.4. WHY STUDY INFORMAL INFERENTIAL REASONING?

Given the importance of statistical inferential reasoning, and given the difficulties most people have with this type of reasoning, a better pedagogical approach to this topic is needed. One possible cause for students' difficulty with formal statistical inferential reasoning is that they lack both experience with stochastic events that form the underpinnings of statistical inference (Pfannkuch, 2005), and experience of reasoning about these events. Statistics educators and statistics education researchers have recently been exploring the idea that if students begin to develop the informal ideas of inference (as defined above) early in a course or curriculum, they may be better able to learn and reason about formal methods of statistical inference. For example, early on in a course, students could engage in discussions of when, compared to some agreed upon expectation, a sample is surprising. This discussion could be revisited during different activities that introduce different samples and distributions. Over time, this may develop the prior statistical knowledge needed to understand the idea of a p-value when it is introduced later in a course.

There is also a belief that, if students become familiar with reasoning about inference in an informal manner, such as making speculations about what might be true in a population or populations, based on samples of data, that the method of doing this formally may be more accessible. Finally, because statistical inference integrates many important ideas in statistics—such as data representation, measures of center and variation, the normal distribution, and sampling—introducing informal inference early and revisiting the topic throughout a single course or curriculum across grades could provide students with multiple opportunities to build the conceptual framework needed to support inferential reasoning. These suggestions for ways in which students reason informally about statistical inference, as well as possible methods for developing that reasoning, are currently untested conjectures that need to be studied.

Now that a working definition of IIR and a rationale for studying IIR have been provided, many questions emerge. For example, how can researchers investigate and describe the nature of IIR in students? Or, what are ways to challenge students to reveal their informal inferential reasoning? The three components of the IIR framework can be used to help answer these questions because they support the development of tasks to examine students' intuitive IIR as well as their developing reasoning.

3. WHAT TYPES OF TASKS CAN BE USED TO STUDY INFORMAL INFERENTIAL REASONING AND ITS DEVELOPMENT?

The research literature contains examples of two different approaches that researchers have used to study IIR. One approach focuses on the nature of this reasoning or naïve methods of reasoning about inference given problems and statistical information. An objective of this type of study is often to examine how students reason about or make inferences given a particular problem without having encountered formal methods of statistical inference via instruction. A second approach is the examination of the development of IIR as students experience curricula (e.g., a course or unit of instruction) designed to build reasoning. The objective of this type of study is often to see how the nature of students' inferential reasoning changes as they are provided with resources, tools, and curriculum. Both of these approaches need well-designed tasks that allow researchers to capture and evaluate students' IIR.

Reading (2007) suggested that tasks used in a study of students' informal inferential reasoning would not only need to examine how students integrate the components of the IIR framework listed in Section 2.3, but also capture ideas of statistical inference such as generalizing to an appropriate population beyond a collected sample, basing inferences on evidence, choosing between competing models (i.e., hypotheses), expressing a degree of uncertainty in making an inference, and making connections between the results and problem context. Furthermore, the research literature on informal reasoning and informal knowledge would suggest that tasks should be designed to elicit multiple arguments from students, as well as separate novice reasoning from expert reasoning.

The framework provided in the working definition in this paper suggests the design of tasks that challenge students to

1. Make judgments, claims, or predictions about a population based on samples, but not using formal statistical procedures and methods (e.g., p-value, *t* tests);

2. Draw on, utilize, and integrate prior knowledge (formal and informal) to the extent that this knowledge is available; and
3. Articulate evidence-based arguments for judgments, claims, and predictions about populations based on samples.

Three general types of task have been used in research studies that meet these criteria. They may be categorized as tasks that ask students to

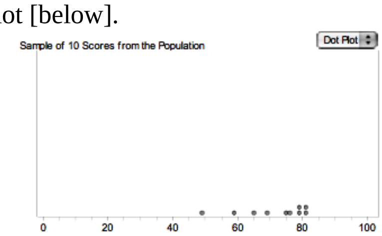
1. Estimate and draw a graph of a population based on a sample;
2. Compare two or more samples of data to infer whether there is a real difference between the populations from which they were sampled, and
3. Judge which of two competing models or statements is more likely to be true.

Examples of tasks for each of three categories are provided in the following sections.

3.1. ESTIMATE AND DRAW A GRAPH OF A POPULATION BASED ON A SAMPLE

Very few examples appear in the literature where students have been asked to speculate about graphical characteristics of a population based on a sample of data. Bakker (2004) referred to this type of prediction as “growing a sample” and used the following task in a teaching experiment with eighth grade students in the Netherlands. Students were asked to predict a graph of weights for a class of 27 eighth grade students and then graphs for three classes together, which had a total of 67 students, based on small random samples of student weights. After being shown the computer-simulated data sets for one class of 27 students and all three classes together, they were asked to describe the differences between their two graphs and then to compare these to the real graphs of weight data. In the last part of the activity, students were asked to create graphs for the population of all students in their city that were no longer sets of points but were continuous distributions of data. This multistage activity ended in an IIR activity that had students make a conjecture about an unknown population. Based on Bakker’s activity, the following task (see Figure 1) was created and used by Zieffler, delMas, Garfield, and Gould (2007) to reveal students’ IIR in an introductory college statistics course.

Imagine the test scores for a group of college students in a very large lecture class on psychology (*n*=1000 students). The test scores for a random sample of ten students from this class are shown in the dot plot [below].



- Now, consider a random sample of 25 students drawn from the same class. Try to imagine what THAT graph might look like. Use the graphing area to sketch a dot plot of the 25 scores that you might expect to see for a random sample of 25 students. Explain your reasoning.
- Next, think about the entire class of 1000 students that these students are sampled from. What would you expect the distribution for the entire population of all 1000 students test scores to look like? Draw an outline of the distribution and explain your reasoning.

Figure 1. Predicting characteristics of a population from a sample task (Zieffler et al., 2007)

3.2. COMPARE TWO OR MORE SAMPLES OF DATA TO INFER WHETHER THERE IS A REAL DIFFERENCE BETWEEN THE POPULATIONS

There are many examples in the literature of tasks that ask students to compare two or more groups of data, although not all of them ask the students to reason beyond the samples to the populations from which they have been selected. For example, Watson and Moritz (1999) used tasks in which students in grades 3 through to 9 had to compare two data sets to help them begin to make inferences about group differences. Although these tasks did not look beyond the data sets to larger populations, they set the stage for such inferences, providing a foundation for statistical inference. One such task is in Figure 2.

Two schools are comparing some classes to see which is better at quick recall of 9 maths facts. In each part of this question you will be asked to compare different classes. First consider two classes, the Blue class and the Red class. The scores for the two classes are shown on the two charts below. Each box is one person's test, and the number inside is their score. In the *Blue* class, 4 people scored 2 correct and 2 people scored 3. In the *Red* class, 3 people scored 6 correct and 3 people scored 7.

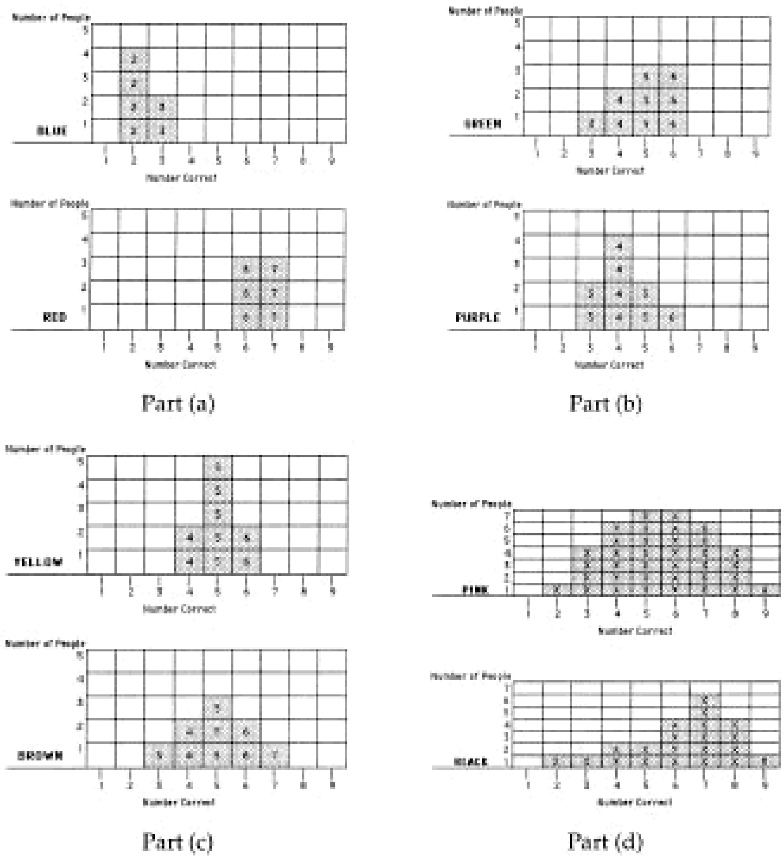


Figure 2. Comparing groups task (Watson & Moritz, 1999)

In contrast, Pfannkuch (2005, 2006) used a task that asked students to compare sets of data for daily temperatures for two different cities in New Zealand, and challenged the students to make some informal inferences beyond the sample data. According to Pfannkuch (2005), "students were required to pose a question (e.g., Which city has the higher maximum temperatures in summer?), analyze the data, draw a conclusion, justify the conclusion with three supporting statements, and evaluate the statistical process” (p. 272). All students constructed a pair of boxplots to make the comparison (see Figure 3).

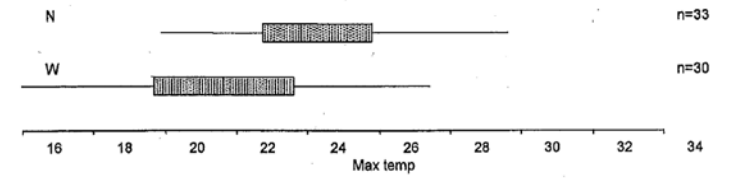
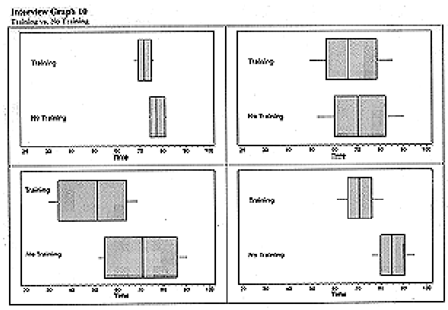


Figure 3. Boxplots of temperature data from comparing groups task (Pfannkuch, 2005)

Suppose that there is a special summer camp for track athletes. There is one group of 100 athletes that run a particular race, and they are all pretty similar in their height, weight, and strength. They are randomly assigned to one of two groups. One group gets an additional weight-training program. The other group gets the regular training program without weights. All the students from both groups run the race and their times are recorded, so that the data could be used to compare the effectiveness of the two training programs.

- Describe what you would expect to see in a comparison of two graphs if the difference between the two groups of athletes is really not due to the training program.
- Describe what you would expect to see in a comparison of two graphs if the difference between the two groups of athletes is really due to the training program.

Presented below are some possible graphs that show boxplots for different scenarios, where the running times are compared for the students in the two different training programs (one with weight training and one with no weight training). Examine each pair of graphs and think about whether or not the sample data would lead you to believe that the difference in running times is caused by these two different training programs. (Assume that everything else was the same for the students and this was a true, well-designed experiment.)



- Which set of boxplots show the MOST convincing evidence that the weight-training program was more effective in DECREASING athlete’s running times? Explain.
- Which set of boxplots shows the LEAST convincing evidence that the weight-training program was more effective? Explain.
- Rank the four pairs of graphs on how convincing they are in making an argument that the weight-training program was more effective in decreasing athletes’ times (from the least convincing to the most convincing evidence). Explain your reasoning.
- For the pair of graphs that provide the most convincing evidence, would you be willing to generalize the effects of the training programs to all similar athletes on track teams, based on these samples? Why or why not?

Figure 4. Comparing groups task (Zieffler et al., 2007)

A third type of task has students compare multiple pairs of sample data and has the students judge which pair provides the most compelling evidence to support a true difference in population means. For example, Zieffler et al. (2007) used a task (Figure 4) to have students make conjectures about whether or not a special athletic training program was effective, based on two samples of data. This task required students to look beyond the samples of data to compare two populations, and to choose which pair of samples gave the most compelling evidence to support a claim that there was a real difference.

3.3. JUDGE WHICH OF TWO COMPETING MODELS OR STATEMENTS IS MORE LIKELY TO BE TRUE

A review of the literature found three different styles of task that have been used to challenge students to choose between two competing models or claims, based on sample data. One style uses data generated by a probability device, and uses proportions or percentages to summarize the sample data. For example, Stohl and Tarr (2002) and Tarr, Stohl Lee, and Rider (2006) used the *Schoolopoly* problem (Figure 5) that asked sixth-grade students to judge whether a die is fair or not based on observed tosses via a computer simulation. Students were asked to provide what they "consider 'compelling evidence' in formulating and evaluating arguments based on data" (Tarr, Stohl Lee, & Rider, 2006, p. 1).

Schoolopoly: Is the die fair or biased?

Background

Suppose your school is planning to create a board game modeled on the classic game of *Monopoly*. The game is to be called *Schoolopoly* and, like *Monopoly*, will be played with dice. Because many copies of the game expect to be sold, companies are competing for the contract to supply dice for *Schoolopoly*. Some companies have been accused of making poor quality dice and these are to be avoided since players must believe the dice they are using are actually "fair." Each company has provided dice for analysis and you will be assigned one company to investigate:

Luckytown Dice Company	*Dice, Dice, Baby!*
Dice R' Us	*Pips and Dots*
High Rollers, Inc.	*Slice 'n' Dice*

Your Assignment

Working with a partner, investigate whether the dice sent to you by the company are *fair* or *biased*. That is, collect data to infer whether all six outcomes are equally likely and answer the following questions:

1. Do you believe the dice you tested are fair or biased? Would you recommend that dice be purchased from the company you investigated?
2. What *compelling evidence* do you have that the dice you tested are fair or unfair?
3. Use your data to estimate the probability of each outcome, 1-6, of the dice you tested.

Collect data about the dice supplied to you. Note that each single trial represents the outcome of one roll of a "new" virtual die provided by the company.

Copy any graphs and screen shots you want to use as evidence and print them for your poster. Give a presentation pointing out the highlights of your group's poster.

Figure 5. Competing models task: Schoolopoly (Stohl & Tarr, 2002)

A second style of competing models task was used by Rubin, Hammerman, and Konold (2006). They had teachers determine whether a particular change in process either occurred or did not occur during a specified interval of time. This task is described in Figure 6.

> The Mus-Brush Company produces mushroom brushes, using a large machine whose output is on average 215 brushes every two minutes *if it is working normally*. If the electricity to the machine is interrupted, even for a brief time, it will slow down such that the output of the machine will be 10% lower on average. The Mus-Brush Company was robbed last night; in forcing the door open, the thief disrupted the electricity and the machine became less productive from that time on. There is a prime suspect who has an alibi between midnight and 3AM (he was seen at a bar), so the police have a special interest in determining if the break-in occurred before midnight or after 3, since the suspect has no alibi for that time interval. We have data on Mus-Brush production every two minutes from 8PM until 5AM. Our job is to decide whether there is enough evidence to argue that the break-in occurred between 12 and 3, thus getting the suspect off the hook.

Figure 6. Competing models task: Mus-Brush Company (Rubin et al., 2006)

In contrast to these first two styles of tasks, Zieffler et al. (2007) used a population of quantitative data as a basis for the null model (see Figure 7) and asked students to make decisions about whether a certain educational outcome observed in a sample of data (change in mean test score) was due to chance or not. This was part of a multipart task that asked students to first imagine different possible samples for the population and sketch a hypothetical graph of the distribution of these samples.

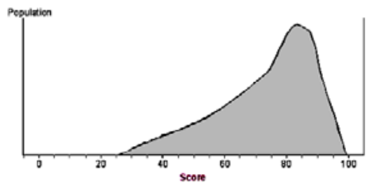
> Shown below is a graph of scores from many sections of students who have taken this course and exam. For this population, the average score is 74.
>
> 
>
>
> A random sample of 50 students in the class this year, given the exact same exam, had a mean exam score for of 78.
>
> 1. Do you think that the teacher can say that this year's students did better on average than what would be expected? Explain.
> 2. Do you think this higher sample average score could just be due to chance?

Figure 7. Competing models task modified from Zieffler et al. (2007)

3.4. ANALYSIS OF THE THREE TYPES OF TASKS

Table 1 illustrates how each of the three types of tasks shown incorporates the essential components of IIR that have been described in the research literature (see Section 2.3).

These tasks can be used in an interview, or on a written assessment to capture students' informal inferential reasoning, or can be embedded in classroom activities designed to promote the development of IIR. Each task may be used to reveal the extent to which students have integrated their prior knowledge about foundational concepts. They challenge students to make judgments and predictions about a population without the use of formal statistical methodology. Lastly, by having students explain their reasoning, usually more than once during the task, they elicit the articulation of students' arguments and justifications for their predictions and judgments.

Table 1. Specification of how each type of task incorporates the three components of IIR

TYPE OF TASK	IIR COMPONENT		
	Make judgments or predictions	Use or integrate prior knowledge	Articulate evidence-based arguments
Estimate and draw a population graph	Predict characteristics of a population (shape, center, spread) that are represented in a student-constructed graph	Bring in intuitive or previously learned knowledge and language to predict the characteristics of a population (e.g., idea of shape, words like skewed)	Requires an explanation of how the characteristics of the population graph were chosen
Compare two samples of data	Judge whether there is a difference between two populations; based on similarities or differences in samples of data.	Bring in intuitive or previously learned knowledge and language to compare two samples of data (e.g., between- and within-groups variation)	Requires an explanation of why students determined whether or not there is a difference in the two populations
Judge between two competing models	Judge whether sample data provide more support for one model than another	Bring in intuitive or previously learned knowledge and language to judge between two competing models (e.g., sampling variability, chance variation)	Requires an explanation of why students chose one model over the competing model

These tasks (or parallel versions of the tasks) could be given to students at multiple times throughout a course or unit of instruction to examine how students' reasoning develops. This would allow instructors to examine how students use their informal knowledge and informal reasoning to draw conclusions and make inferences as they experience instruction related to informal or formal methods of statistical inference. This assessment could be done formatively which would allow for more opportunity for feedback to students, which in turn, would create more learning and research opportunities.

4. SUMMARY AND IMPLICATIONS FOR RESEARCH AND TEACHING

In this paper, we set out to answer two main questions regarding informal inferential reasoning. First, what are the components of a framework needed to support research on informal inferential reasoning? Drawing on the research literature, we proposed a working definition of IIR that comprises three components: (1) making judgments, claims, or predictions about populations based on samples, but not using formal statistical procedures and methods (e.g., p-value, *t* tests); (2) drawing on, utilizing, and integrating prior knowledge to the extent that this knowledge is available; and (3) articulating evidence-based arguments for the judgments, claims, and predictions about populations based on samples.

The second aim of this paper was to identify the types of tasks suggested by this framework for the study of informal inferential reasoning and its development? We have proposed two approaches that might be taken by researchers in studying IIR. We have also suggested the types of tasks that might be helpful in these types of studies, as well as provided concrete examples of such tasks.

The IIR framework (see Section 2.3) and suggested tasks (see Section 3) can be useful in further research studies in various ways. First, by referring to a common working definition in future studies, researchers may better be able to build on and connect their work to previous and subsequent studies. Second, using common tasks allows for comparisons across different instructional settings and groups of students. Third, these tasks have value in both teaching and research settings as they may be used in class activities to promote student reasoning and to challenge students to explain and articulate their understanding and rationale for making inferences, which could be studied to see how students' reasoning changes during the activity.

Researchers could also draw on the proposed IIR framework to help analyze students' responses to tasks designed to elicit IIR. Drawing on the first component of the framework, a researcher might ask whether the student made reasonable inferences about one or more populations based on one or more samples. For example, for the task in Figure 1 (predicting characteristics of a population from a sample), a reasonable inference might be that the center (mean) of the population is near the value that appears to be at the center of the sample, and the variability in the population is likely to be greater than the variation displayed in the sample.

Drawing on the second component of the framework, a researcher might ask how the student used and integrated informal knowledge (e.g., everyday knowledge of the problem context, prior knowledge about statistical concepts, real world knowledge and experience, and statistical language) in making inferences. Another question might be whether the use of problem context has impeded, or over-ridden, the use of data in making inferences. For example, using the same task, an integrated response might incorporate ideas of random sample as being representative of the population; ideas of distribution (e.g., shape, center, variation); and real world knowledge of the problem-context (e.g., test scores, college students' study habits). Another question of interest might be how heavily the student depends on prior knowledge of statistics (previously learned concepts) and how much the student depends on his or her knowledge of the world (or experience), a balance that may change over the course of instruction.

Drawing on the third component of the framework, a researcher might ask how the student has used evidence to support his/her arguments in making inferences, and also, how well the evidence used supported the inferences made. For example, using the same task, the response should include data-based explanation for why the student chose a particular population distribution (e.g., the population will likely have a mean near 71 because the sample had a mean of 71.3 and this sample was drawn randomly from the population so it should be representative).

Although the IIR framework proposed may be useful in studying the development of IIR (e.g., during an activity, or over a unit of instruction, an entire course, or even a curriculum) there is not yet a developmental model of how IIR develops from the earliest and most informal stage to transitioning to formal statistical reasoning. A variety of theories of developmental growth exist that could be used to underpin a "developmental" framework for IIR. One such cognitive model of learning is the Structure of Observed Learning Outcome (SOLO) based developmental framework (see Pegg, 2003). Based on the research presented at the Fifth International Research Forum on Statistical Reasoning, Thinking and Literacy (SRTL-5), Reading (2007) suggested that a two-cycle SOLO-based framework be considered for describing the cognitive growth students experience when reasoning about statistical inference. The first cycle would involve reasoning about underlying concepts, including naïve inference that is not chance related. The second cycle would involve using these underlying concepts in a more "formal" way that incorporates reasoning about chance events.

The authors note that a working definition for IIR is a definition in progress. They hope that over time others will contribute to refining and updating this definition as more information is gained about the nature and development of students' informal inferential reasoning. There is a need for more research to explore the role of foundational concepts, data sets and problem contexts, and technology tools in helping students to reason informally, and then formally, about statistical inference. There are many unanswered questions about the best sequence of ideas and activities and the role of these in making the transition from informal to formal methods of statistical inference.

ACKNOWLEDGEMENTS

We appreciate the support of a grant from the National Science Foundation (NSF CCLI – 0535912).

REFERENCES

Bakker, A. (2004). Reasoning about shape as a pattern in variability. *Statistics Education Research Journal*, *3*(2), 64-83.
[Online: http://www.stat.auckland.ac.nz/~iase/serj/SERJ3(2)_Bakker.pdf]

Bakker, A., Derry, J., & Konold, C. (2006). Technology to support diagrammatic reasoning about center and variation. In A. Rossman & B. Chance (Eds.), *Working cooperatively in statistics education: Proceedings of the Seventh International Conference on Teaching Statistics*, Salvador, Brazil. [CDROM]. Voorburg, The Netherlands: International Statistical Institute.
[Online: http://www.stat.auckland.ac.nz/~iase/publications/17/2D4_BAKK.pdf]

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

Ben-Zvi, D. (2006). Scaffolding students' informal inference and argumentation. In A. Rossman & B. Chance (Eds.), *Working cooperatively in statistics education: Proceedings of the Seventh International Conference on Teaching Statistics*, Salvador, Brazil. [CDROM]. Voorburg, The Netherlands: International Statistical Institute.
[Online: http://www.stat.auckland.ac.nz/~iase/publications/17/2D1_BENZ.pdf]

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

Carver, R. (2006, August). *Ambiguity intolerance: An impediment to inferential reasoning?* Paper presented at the Joint Statistics Meetings, Seattle, WA.

Cobb, P., & McClain, K. (2004). Principles of Instructional Design for Supporting the Development of Students' Statistical Reasoning. In D. Ben-Zvi and J. Garfield (Eds.), *The Challenge of Developing Statistical Literacy, Reasoning, and Thinking* (pp. 375-396). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Cobb, P., McClain, K., & Gravemeijer, K. P. E (2003). Learning about statistical covariation. *Cognition and Instruction, 21*, 1-78.

Cobb, P., Yackel, E., & Wood, T. (1992). A constructivist alternative to the representational view of mind in mathematics education. *Journal of Research in Mathematics Education, 23*(1), 23-33.

Cohen, J. (1994). The earth is round (*p* < .05). *American Psychologist*, *49*(12), 997-1003.

Confrey, J. (1990). A review of the research on students' conceptions in mathematics, science, and programming. *Review of Research in Education, 16*, 3-56.

delMas, R., Garfield, J., & Chance, B. (1999). A model of classroom research in action: Developing simulation activities to improve students' statistical reasoning. *Journal of Statistics Education, 7*(3).
[Online: http://www.amstat.org/publications/jse/secure/v7n3/delmas.cfm]

Driver, R., Newton, P., & Osborne, J. (2000), Establishing the Norms of Scientific Argumentation in Classrooms, *Science Education, 84*, 287-312.

Evans, J. St. B. T., Newstead, S. E., & Byrne, R. M. J. (1993). *Human reasoning: The psychology of deduction*. Hove (UK): Lawrence Erlbaum Associates Ltd.

Falk, R., & Greenbaum, C. W. (1995). Significance tests die hard: The amazing persistence of a probabilistic misconception. *Theory and Psychology, 5,* 75-98.

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

Garfield, J., delMas, R., & Chance, B. (2007). Using students' informal notions of variability to develop an understanding of formal measures of variability. In M. Lovett and P. Shah (Eds.), *Thinking with Data (Proceedings of the 33rd Carnegie Symposium on Cognition)* (pp. 117-147). New York: Erlbaum.

Gravemeijer, K., & Doorman, K. (1999). Context problems in realistic mathematics education: A calculus course example. *Educational Studies in Mathematics, 39*(1/3), 111-129.

Jones, G. A., Langrall, C. E., Mooney, E. S., & Thornton, C. A. (2005). Models of development in statistical reasoning. In D. Ben-Zvi and J. Garfield (Eds.), *The Challenge of Developing Statistical Literacy, Reasoning, and Thinking* (pp. 97-117). Dordrecht, The Netherlands: Kluwer Academic Publishers.

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

Klahr, D., Fay, A., & Dunbar, K. (1993). Heuristics for scientific experimentation: A developmental study. *Cognitive Psychology, 25*, 111-146.

Kuhn, D. (1991). *The skills of argument*. Cambridge, UK: Cambridge University Press.

Kuhn, D., Garcia-Mila, M., Zohar, A., & Andersen, C. (1995). Strategies of knowledge acquisition. *Monographs of the Society for Research in Child Development, 60*(4, Serial No. 245).

Lipson, K. (2003). The role of the sampling distribution in understanding statistical inference. *Mathematics Education Research Journal, 15*(3), 270-287.

Mack, N. K. (1995). Confounding whole-number and fraction concepts when building on informal knowledge. *Journal of Research in Mathematics Education, 26*(5), 422-441.

Means, M. L., & Voss, J. F. (1996). Who reasons well? Two studies of informal reasoning among children of different grade, ability, and knowledge levels. *Cognition and Instruction, 14*(2), 139-178.

Miller-Jones, D. (1991). Informal reasoning in inner-city children. In J. F. Voss, D. N. Perkins, and J. Segal (Eds.), *Informal reasoning and education* (pp. 107-130). Hillsdale, NJ: Lawrence Erlbaum Associates.

Moore, D.S. (1990). Uncertainty. In L.A. Steen (Ed.), *On the shoulders of giants* (pp. 95-173). Washington, DC: National Academy Press.

Moore, D.S. (2004). *The basic practice of statistics* (3rd ed.). New York: W. H. Freeman.

Nickerson, R. S. (2000). Null hypothesis significance testing: A review of an old and continuing controversy. *Psychological Methods, 5*(2), 241-301.

Nickerson, R., Perkins, D. N., & Smith, E. (1985). *The teaching of thinking*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Papert, S., & Harel, I. (Eds.) (1991). *Constructionism*. Norwood, NJ: Ablex Publishing.

Pegg, J. (2003). Assessment in mathematics: A developmental approach. In J. Royer (Ed.), *Mathematical Cognition* (pp. 227-259). Greenwich, CT: Information Age Publishing.

Perkins, D. N. (1985a). Postprimary education has little impact on informal reasoning. *Journal of Educational Psychology, 77*(5), 562-571.

Perkins, D. N. (1985b). Reasoning as imagination. *Interchange, 16*(1), 14-26.

Perkins, D. N. (1989). Reasoning as it is and could be. In D. Topping, D. Crowell, & V. Kobayashi (Eds.), *Thinking: The third international conference* (pp. 175-194). Hillsdale, NJ: Lawrence Erlbaum Associates.

Perkins, D. N., Bushey, B., & Farady, M. (1986). *Learning to reason* (Final report for grant no. NIE-G-83_0028). Cambridge, MA: Harvard Graduate School of Education.

Perkins, D. N., Farady, M., & Bushey, B. (1991). Everyday reasoning and the roots of intelligence. In J. F. Voss, D. N. Perkins, and J. Segal (Eds.), *Informal reasoning and education* (pp. 83-105). Hillsdale, NJ: Lawrence Erlbaum Associates.

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

Pfannkuch, M. (2006). Informal inferential reasoning. In A. Rossman & B. Chance (Eds.), *Working cooperatively in statistics education: Proceedings of the Seventh International Conference on Teaching Statistics*, Salvador, Brazil. [CDROM]. Voorburg, The Netherlands: International Statistical Institute.
[Online: http://www.stat.auckland.ac.nz/~iase/publications/17/6A2_PFAN.pdf]

Quilici, J. L., & Mayer, R. E. (2002). Teaching students to recognize structural similarities between statistics word problems. *Applied Cognitive Psychology, 16*, 325-342.

Reading, C. (2007, August). *Cognitive development of reasoning about inference*. Discussant reaction presented at the Fifth International Research Forum on Statistical Reasoning, Thinking and Literacy (SRTL-5), University of Warwick, UK.

Rossman, A. (2007, August). *A statistician's view on the concept of inferential reasoning*. Paper presented at the Fifth International Research Forum on Statistical Reasoning, Thinking and Literacy (SRTL-5), University of Warwick, UK.

Rubin, A., Hammerman, J., & Konold, C. (2006). Exploring informal inference with interactive visualization software. In A. Rossman & B. Chance (Eds.), *Working cooperatively in statistics education: Proceedings of the Seventh International Conference on Teaching Statistics*, Salvador, Brazil. [CDROM]. Voorburg, The Netherlands: International Statistical Institute.
[Online: http://www.stat.auckland.ac.nz/~iase/publications/17/2D3_RUBI.pdf]

Sadler, T. D. (2004). Informal reasoning regarding socioscientific issues: A critical review of research. *Journal of Research in Science Teaching, 41*(5), 513-536.

Sadler, T. D., & Zeidler, D. L. (2004). The significance of content knowledge for informal reasoning regarding socioscientific issues: Applying genetics knowledge to genetic engineering issues. *Science Education, 89*, 71-93.

Saldanha, L. A., & Thompson, P. W. (2002). Conceptions of sample and their relationship to statistical inference. *Educational Studies in Mathematics, 51*(3), 257-270.

Saldanha, L. A., & Thompson, P. W. (2006). Investigating statistical unusualness in the context of resampling: Students exploring connections between sampling distributions

and statistical inference. In A. Rossman & B. Chance (Eds.), *Working cooperatively in statistics education: Proceedings of the Seventh International Conference on Teaching Statistics*, Salvador, Brazil. [CDROM]. Voorburg, The Netherlands: International Statistical Institute.
[Online: http://www.stat.auckland.ac.nz/~iase/publications/17/6A3_SALD.pdf]

Schauble, L. (1990). Belief revision in children: The role of prior knowledge and strategies for generating evidence. *Journal of Experimental Child Psychology, 49,* 31-57.

Schauble, L. (1996). The development of scientific reasoning in knowledge-rich contexts. *Developmental Psychology*, *32,* 102-119.

Schauble, L., & Glaser, R. (1990). Scientific thinking in children and adults. In D. Kuhn (Ed.), *Developmental perspectives on teaching and learning thinking skills* (pp. 9-27). Basel, Switzerland: Karger.

Schoenfeld, A. H. (1982). Measures of problem-solving performance and of problem-solving instruction. *Journal for Research in Mathematics Education*, *13*(1), 31-49.

Schoenfeld, A. H. (1991). On mathematics as sense-making: An informal attack on the unfortunate divorce of formal and informal mathematics. In J. F. Voss, D. N. Perkins, and J. Segal (Eds.), *Informal reasoning and education* (pp. vii-xvii). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Schoenfeld, A. H., & Herrmann, D. J. (1982). Problem perception and knowledge structure in expert and novice mathematical problem solvers. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8,* 484-494.

Schwarz, D. L., Sears, D., & Chang, J. (2007). Reconsidering prior knowledge. In M. Lovett and P. Shah (Eds.), *Proceedings of the 33rd Carnegie Symposium on Cognition: Thinking with Data*. Mahweh, NJ: Erlbaum.

Smith, J. P., diSessa, A. A., & Roshelle, J. (1993/1994). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *The Journal of the Learning Sciences, 3*(2), 115-163.

Sotos, A. E. C., Vanhoof, S., Van den Noortgate, W., & Onghena, P. (2007). Students' misconceptions of statistical inference: A review of the empirical evidence from research on statistics education. *Educational Research Review, 2*, 98-113.

Stohl, H. & Tarr, J. E. (2002). Developing notions of inference using probability simulation tools. *Journal of Mathematical Behavior, 21,* 319-337.

Tarr, J. E., Stohl Lee, H., & Rider, R. (2006). When data and chance collide: Drawing inferences from simulation data. In G. Burrill (Ed.), *Thinking and reasoning with data and chance: Sixty-eighth NCTM yearbook* (pp. 139-150). Reston, VA: National Council of Teachers of Mathematics.

Thompson, P., Saldanha, L., & Liu, Y. (2004, April). *Why statistical inference is hard to understand*. Paper presented at the Annual Meeting of the American Educational Research Association, San Diego.

Toulmin, S. (1958). *The uses of argument*. Cambridge, UK: Cambridge University Press.

van Eemeren, F. H., Grootendorst, R., Henkemans, F. S., Blair, J. A., Johnson, R. H., Krabbe, E. C. W., et al. (1996). *Fundamentals of argumentation theory: A handbook of historical backgrounds and contemporary developments*. Mahweh, NJ: Erlbaum.

Voss, J. F., Blais, J., Means, M. L., Greene, T. R., & Ahwesh, E. (1986). Informal reasoning and subject matter knowledge in the solving of economics problems by naïve and novice individuals. *Cognition and Instruction, 3*(3), 269-302.

Voss, J. F., Perkins, D. N., & Segal, J. W. (1991). Preface. In J. F. Voss, D. N. Perkins, and J. Segal (Eds.), *Informal reasoning and education* (pp. vii-xvii). Hillsdale, NJ: Lawrence Erlbaum Associates.

Walton, D. N. (1989). *Informal logic: A handbook for critical argumentation*. Cambridge, UK: Cambridge University Press.

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

Watson, J. M., & Moritz, J. B. (1999). The beginnings of statistical inference: Comparing two data sets. *Educational Studies in Mathematics, 37*, 145-168.

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

Zieffler, A., delMas, R., Garfield, J., & Gould, R. (2007, August). *Studying the development of college students' reasoning about statistical inference.* Paper presented at the Fifth International Research Forum on Statistical Reasoning, Thinking and Literacy (SRTL-5), University of Warwick, UK.

ANDREW ZIEFFLER
Educational Psychology
Room 250 EdSciB
4101
56 E River Road
Minneapolis, MN 55455