



HAL
open science

A Bayesian inspired approach to reasoning about uncertainty: 'How confident are you?'

Sibel Kazak

► **To cite this version:**

Sibel Kazak. A Bayesian inspired approach to reasoning about uncertainty: 'How confident are you?'. CERME 9 - Ninth Congress of the European Society for Research in Mathematics Education, Charles University in Prague, Faculty of Education; ERME, Feb 2015, Prague, Czech Republic. pp.700-706. hal-01287091

HAL Id: hal-01287091

<https://hal.science/hal-01287091>

Submitted on 11 Mar 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Bayesian inspired approach to reasoning about uncertainty: 'How confident are you?'

Sibel Kazak

Pamukkale University, Denizli, Turkey, skazak@pau.edu.tr

This paper focuses on a task design that is aimed at eliciting young students' reasoning about uncertainty as it relates to their personal degree of confidence through a Bayesian inspired informal inferential reasoning about chance games. This Bayesian inspired approach is described and discussed based on preliminary analyses of data from a teaching experiment in a designed-based research study. With this approach, beginning by a hypothesis (or prediction) about the fairness of a game and revising it based on new information appear to come natural to the students. The change in strength of their personal level of confidence seems to vary by the conflicting results obtained by playing the game, the size of the data collected, and the multiple computer simulations conducted using TinkerPlots software.

Keywords: Uncertainty, probability, inferential reasoning, task design.

INTRODUCTION

Making decisions and inferences based on data is part of everyday life. In statistics, statistical inference deals with "drawing conclusions about populations or processes based on sample data" (Zieffler, Garfield, delMas, & Reading, 2008, p. 40) by means of certain techniques, such as confidence intervals and hypothesis testing. As argued by Engel and Erickson (2013) it is yet hard to grasp the logic behind statistical inference for students. In order to develop the foundation for the formal ideas involved in statistical inference early on, informal statistical inference has become the focus of the statistics education research (Rubin, Hammerman, & Konold, 2006; Makar & Rubin, 2009). As a relatively recent concept, informal statistical inference refers to a way of making conclusions about a population or process from which the data come by using statistical processes (Pfannkuch, 2006). In addition, the underlying reasoning process leading

to informal statistical inference is called informal inferential reasoning (Makar, Bakker, & Ben-Zvi, 2011).

In recent research studies, there has been a great emphasis particularly on informal inferential reasoning to help students develop deeper understanding of statistical concepts, ideas, and processes (Ben-Zvi, Aridor, Makar, & Bakker, 2012; Jacob & Doerr, 2013; Makar, 2014; Paparistodemou & Meletiou-Mavrotheris, 2008). One of these ideas at the heart of informal statistical inference is uncertainty because drawing conclusions beyond the data about a wider population requires an articulation of uncertainty (Makar & Rubin, 2009) and probabilistic justifications (Rossman, 2008). When it comes to teaching probability as part of statistics, a current debated topic is to introduce Bayesian thinking at school level because of its applications in realistic situations and its closeness to how people actually reason about uncertainty (Chernoff, 2014; Nilsson, Blomberg, & Ärleback, 2014; Martignon & Erickson, 2014). Therefore more research into supporting young students' development of Bayesian thinking is needed.

The aim of this paper is to describe a Bayesian inspired approach in a task design for eliciting students' reasoning about uncertainty as it relates to personal degree of confidence. First, I outline the ideas behind this approach for task design, which is followed by a brief description of the study and the task. Then I focus on how these ideas are explored during the task as a way to articulate students' reasoning about uncertainty with some preliminary analyses of students' reasoning.

THEORETICAL BACKGROUND

Informal statistical inference and inferential reasoning

In an informal statistical inference, a common underlying reasoning process involves “assessing the strength of evidence against a claim” (Rossman, 2008, p. 7) based on observed data. This type of informal inferential reasoning known as Fisherian approach can be described as follows (Rossman, 2008): formulating an initial hypothesis (or null hypothesis); evaluating that if the hypothesis were true, observed data would have been very unlikely (i.e., intuitively computing a p-value); and rejecting the initial hypothesis based on the very small p-value. According to Rossman, students yet do not appear to use this common reasoning naturally when making statistical inference.

Alternatively, another approach to statistical inference involves a Bayesian perspective based on the subjectivist interpretation of probability. A main distinction between Bayesian inference and Fisherian approach is drawing conclusions based on a subjective or personal assessment of uncertainty. So the reasoning process from a Bayesian approach includes starting with a priori probabilities associated with a hypothesis based on a personal belief and updating these probabilities in the light of new information or data (Rossman, 2008). Albert (2002) argues that the Bayesian reasoning is more intuitive than the Fisherian perspective in statistical inference and better reflects the commonsense thinking about uncertainty in everyday life.

Reasoning about uncertainty

Both in school mathematics curricula and research on students' understandings of uncertain events, games and experiments involving chance devices, such as coin, dice, and spinners, are quite widely used. Games of chance provide a rich context for children to explore random situations, to notice the unpredictability of outcomes and to see the need for probability (Cañizares, Batanero, Serrano, & Ortiz, 2003) while making decisions under uncertainty. The notion of fairness in game situations has also been recognized as a motivating and productive area of inquiry for students investigating probability and uncertainty (Pratt, 2000; Watson & Moritz, 2003; Stohl & Tarr, 2002). For example students can build on their intuitive ideas of fairness to evaluate whether each player has an equal chance of winning or whether each possible

outcome is equally likely in different games involving coins and dice.

These studies on fairness in chance games mainly focus on the concept of probability from classical and frequentist approaches. There is a lack of research on young students' conceptions of subjective probability, which is closely related to the Bayesian reasoning mentioned above. One of the earlier studies both relevant to games of chance and subjective perspective to probability (Huber & Huber, 1987) suggests that young children are able to use personal knowledge or beliefs when comparing the likelihoods of chance events in the contexts of sports and gambling. It is also noted that children's subjective probability evaluations tend to be more stable in the gambling context because the objective probabilities could be assessed also through the sections in the spinner device used in the task (Huber & Huber, 1987).

OVERVIEW OF TASK DESIGN

The purpose of the task design was to elicit students' reasoning about uncertainty in the context of informal statistical inference about the fairness of chance games. A previous study using an earlier version of this task suggested that students' reasoning about uncertainty was inherent both in the chance games and in personal degree of confidence (Kazak, Fujita, & Wegerif, 2013). The task then was revised to closely look at students' reasoning about uncertainty as it relates to personal degree of confidence, which is the focus of this paper. To build on students' intuitive and informal inferential reasoning, a Bayesian inspired approach was adopted to design the Matching Tokens Game task in which students were asked to make probability assessments and state their level of confidence in judging the fairness of a game. This approach enabled students to articulate uncertainty while evaluating the fairness of different chance games by making an initial hypothesis and expressing their confidence in the likelihood of a particular game actually being fair or not, and then by revising both their hypothesis and level of confidence with new information obtained from the data through physical experiments and computer simulations. The underlying process is usually viewed consistent with people's way of developing intuitions based on learning from their experiences and revising their beliefs as new information is acquired (Falk & Konold, 1992).

Study background

The task design is part of a design-based research study. As described by Cobb, Confrey, diSessa, Lehrer, and Shauble (2002), this study involved an iterative design through planning, testing, and revising conjectures about students' learning and ways of supporting their learning of domain specific content, i.e. probability and statistics. The research cycle involved designing instructional materials and a learning environment that supports the desired learning goals, conducting teaching sessions, and retrospective analysis. Three iterations in local schools in Exeter, UK were conducted as part of the larger research study. In these teaching experiments, students worked in groups of two or three on a joint activity. Each group was given worksheets and videotaped while working around a computer. The task described here was tested and revised based on the earlier iterations. The empirical data used to explain a Bayesian inspired approach in this paper are from the last iteration with eleven 10–11-year-old students: group A (Justin, Owen, Matt), group B (Taylor and Sam), group C (Meg, Julie, and Jaily), and group D (Maya, Eleanor, and Jena).

Task and tools

The Matching Tokens Games involve randomly drawing a token from each bag shown in Figure 1. For example, in Game 1 one bag has 3 red tokens and 1 blue token and the other has 1 red token and 3 blue tokens. To play the game, a token will be randomly drawn from each bag. If both tokens are the same color, students win. If they are different (mixed) color, teacher wins. The question posed to students before playing the game is whether the game is fair or not. These four games

were chosen and sequenced based on the students' responses on the previous two iterations of the task. Students investigated each game one after another in the given order.

Through adapting a Bayesian inspired approach this game context was introduced to students in a specific structure involving three phases and ten questions. As seen in Figure 2, in the prediction phase students began by formulating a hypothesis about the fairness of Game 1 based on their personal knowledge/belief. Then on a scale (0–10) they evaluated how confident they were about the un/fairness of the game initially based on the explanation they were asked to give in question 1. In the game-playing phase, students working in groups collected as many data as they wanted by playing the game with the given bags and recording their results and used the results to update their initial hypothesis as well as their level of confidence if needed. In the modeling phase, they built a computer model to simulate the game results and to collect more data, and again revised their previous hypothesis and level of confidence in the light of new information.

TinkerPlots software (Konold & Miller, 2011) was used as a modeling tool in this study. The Sampler tool in *TinkerPlots* allowed students to build their own chance models using a variety of devices (i.e., mixer, spinner, bars, stacks, curve, counter) that can be filled with different elements to sample from. It also enabled students to collect outcomes and carry out a large number of trials quite quickly. For instance, to build a model of Game 1 in *TinkerPlots* (see Figure 3) group B used two connected mixer devices, one with

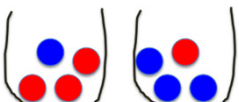
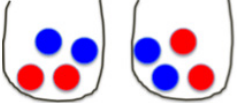
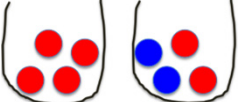
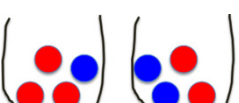
<p>Game 1</p> <p>Bag one: 3 red tokens, 1 blue token Bag two: 1 red token, 3 blue tokens</p>	
<p>Game 2</p> <p>Bag one: 2 red tokens, 2 blue tokens Bag two: 2 red tokens, 2 blue tokens</p>	
<p>Game 3</p> <p>Bag one: 4 red tokens Bag two: 2 red tokens, 2 blue tokens</p>	
<p>Game 4</p> <p>Bag one: 3 red tokens, 1 blue token Bag two: 2 red tokens, 2 blue tokens</p>	

Figure 1: Games given on each worksheet

GAME 1: Names:

PREDICT	1) Discuss and agree in your group: We think the game is FAIR / UNFAIR because ...		2) On the scale below, mark the number that best represents how confident you are that the game is FAIR / UNFAIR. Before Playing the Game ← 0 1 2 3 4 5 6 7 8 9 10 → Not at all Confident Sort of Confident Totally Confident		
	PLAY W/ BAGS	3) How many times did you play the game? _____	Number of times 'same' wins	Number of times 'mixed' wins	5) Mark the number that best represents how confident you are that the game is FAIR / UNFAIR. After Playing the Game ← 0 1 2 3 4 5 6 7 8 9 10 → Not at all Confident Sort of Confident Totally Confident
	MODEL & PLAY	7) How many times did you play the game in TinkerPlots? _____	% of 'same' wins	% of 'mixed' wins	9) Mark the number that best represents how confident you are that the game is FAIR / UNFAIR. After Modelling the Game in TinkerPlots ← 0 1 2 3 4 5 6 7 8 9 10 → Not at all Confident Sort of Confident Totally Confident
	4) Write the total number of winnings for each outcome in table on the right.			6) What would make you more certain?	
	8) Write the percent of winnings for each outcome in the table on the right.			10) What would make you more certain?	

Figure 2: Student worksheet for Game 1

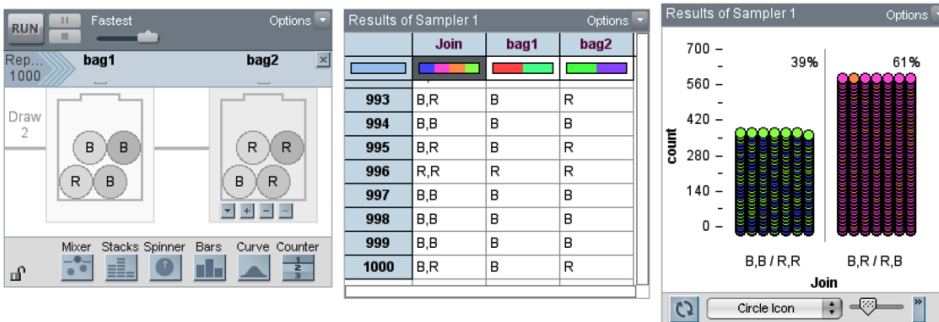


Figure 3: Computer model of Game 1

three red (R) and one blue (B) balls and the other with one red and three blue balls, representing the number of red and blue tokens in each bag. Repeat number is set to 1000. After randomly being selected from each mixer, the outcomes of each trial are displayed in the results table for 1000 repetitions. In the graph, the individual outcomes, 'B,B' and 'R,R' then 'B,R' and 'R,B', are combined into a bin by dragging one into the other to display the percentage of "the same color" and "the mixed color" outcomes respectively.

CHARACTERISTICS OF A BAYESIAN INSPIRED APPROACH TO REASONING ABOUT UNCERTAINTY

In order to develop students' reasoning about uncertainty through informal inferential reasoning, the task design involved a Bayesian inspired approach to informal inference in the context of chance games. Each game investigation starts by having students state their initial hypothesis about the fairness of the game with an explanation and then rate their level of confidence about the un/fairness of the game on a scale from 0 to 10. In order to test their initial predictions students play the game to collect some data and record their results on the worksheet (questions 3 and 4 in Figure 2). Judging the fairness of these games,

except Game 2, were found to be highly counter-intuitive based on the students' initial responses in the previous iterations of the task (Kazak, Wegerif, & Fujita, 2014). So, students are expected to update their initial hypothesis and level of confidence based on new evidence from the game results. Subsequently students build a model of the game in *TinkerPlots* to collect more data to reevaluate their previous hypothesis and confidence level through computer simulations. Next I describe the phases of the student investigations characterized by the adopted Bayesian inspired approach with some analyses of students' responses on the worksheets.

Forming a hypothesis with a level of confidence

Based on their current understanding of probability students began by stating their initial hypothesis about fairness of each game and explanation for it. These students had previous experience with simple events, but the ones described in the games required an understanding of combined outcomes. Thus they relied primarily on their intuitive ideas, which generally led them to an incorrect judgement. The explanations given for their predictions indicated that they seemed to focus either on the symmetry in the total number of red and blue tokens in the bags with an additive reasoning in Game 1 or on the likelihood of simple events in each bag, i.e., the chance of getting a red token from a bag in Game 3 and Game 4. Unlike these counter-intuitive games (1, 3 and 4), Game 2 was consistent with the students' intuitions. Although their predictions about Game 2 were correct, their reasoning was problematic. They expected that the symmetry in the combined bags (equal number of red and blue tokens) would generalize to the combined outcomes.

Starting with students' initial predictions about the fairness of the games was an essential part of the Bayesian inspired approach used in the task. When dealing with uncertainty, we often draw upon a variety of evidence, but particularly personal knowledge or belief or past experience in the absence of empirical results or theoretical knowledge. From the Bayesian perspective this is the basis for subjective probability. Since these beliefs can change based on new evidence, it is important to assess personal degree of confidence in the initial hypothesis or prediction and look at how it changes over the course of gathering new relevant information. Each group's responses on the worksheets showed that students

were likely to be more confident about their initial hypothesis in Game 2 that was more intuitive. The initial confidence level in Game 2 was on average 9.4 (out of 10) for all groups while the average level of confidence in Game 1, Game 3, and Game 4 was 8.9, 8.4, 8.8 respectively. These values can be argued as indications for how much students are willing to rely on their personal beliefs or knowledge without additional evidence. Since the fairness judgment for the games 1, 3, and 4 were less intuitive to them, none of the groups correctly identified whether the game was fair or not initially.

Using information based on experiment data to update level of confidence

To test their initial predictions, each group played the game as many times as they wanted. The number of times that the games were played, i.e. the number of trials, tended to be small and varied among groups, from 5 to 30. Students mostly did not incline to revise their initial beliefs based on the game results unless they believed that they had contradicting results. For example, group A changed their first predictions in Game 1 (based on "same wins=11", "mixed wins=19") and Game 3 (based on "same wins=2", "mixed wins=3"), and group B modified theirs in Game 1 (based on "same wins=5", "mixed wins=15") and Game 4 (based on "same wins=5", "mixed wins=5") after playing the game. These changes in predictions involved switching to 'unfair' in Game 1 and to 'fair' in Game 3 and Game 4. The changes in the level of confidence seemed to vary depending on how much they were convinced by the actual game results. While the level of confidence dropped for Game 3 and Game 4 by -4.5 and -1 points respectively, there was no adjustment for Game 1 in both groups.

Using information based on simulated data to update level of confidence

To further test their initial or current updated hypotheses, each group built a model of the games using *TinkerPlots* to quickly collect more data. Naturally the number of trials in computer simulations became higher, ranging from 100 to 100000. However, students mostly seemed to find 100 trials large enough to base their decisions. The simulated data results generally seemed to help students update their current hypotheses. When asked what would make them more certain in questions 6 and 10 (Figure 2), some groups tended to suggest conducting 'more tests', i.e., running the simulation again with the same number

of trial. Indeed group A and group C carried out a few more simulations with the same number of trials using *TinkerPlots* and recorded the results. At the end of the task, the groups arrived at the correct judgment about the fairness in the majority of the games after running multiple tests using computer simulations. Moreover, there was a positive increase in the level of confidence overall during the simulations. While the average change in confidence level in the case of switching to a right judgment was +0.6, this was doubled (i.e., +1.2) when students already had the right judgment from the game results. This result can be an indication of the effectiveness of additional evidence to support current hypothesis and personal belief.

CONCLUSIONS

The main premise of a Bayesian inspired approach to reasoning about uncertainty is that students' informal inferential reasoning about the fairness of the chance games is closely associated with their personal degree of confidence. Starting with a hypothesis about whether the game is fair or not and revising it based on new information come natural to students. The task design described in this paper is intended to help this process be more systematic by scaffolding students' reasoning step by step in each of the three phases (Figure 2). This way students do not only change their predictions about the fairness of the game based on data they collect but also update the level of their confidence which is linked to their certainty level.

The preliminary analysis offered in this paper suggested that students' first 'intuition-based' hypotheses about the fairness of the games mainly led them to wrong judgments initially particularly in the counter-intuitive games. However, both their conjectures and level of confidence on them tended to improve as they collected more data. More specifically, conducting a large number of trials and multiple tests using *TinkerPlots* modeling tool allowed students to change their beliefs about the fairness of the games and the strength of their confidence. Moreover, obtaining 'surprising' results in the game playing phase generally helped students revise their initial predictions while their level of confidence seemed to decrease to some extent. Students however tended to show more confidence in their personal belief about the fairness of the game when it was confirmed by 100 or more

data collected during the modeling and simulation phase of the task.

Overall, the study shows the importance of integrating the personal beliefs, level of confidence, and empirical results from experiments and simulations in reasoning about uncertainty when making inferences about the fairness of chance games. The approach used in this task suggests a way to bring Bayesian thinking to the school level and makes it accessible to young students as well. Further research on this approach using clinical interviews with more individual groups might be useful to deepen our understanding of how students' experiences based on game results and simulation results affect their beliefs about the fairness of the games, personal degree of confidence, and strategies in the long run.

ACKNOWLEDGEMENT

This research was supported by a Marie Curie Intra European Fellowship within the 7th European Community Framework Programme. The author thanks the STATSTALK project members, Rupert Wegerif and Taro Fujita.

REFERENCES

- Albert, J. (2002). Teaching Introductory Statistics from a Bayesian Perspective. In B. Phillips (Ed.), *Proceedings of the 6th International Conference on Teaching Statistics*. Cape Town, South Africa: International Statistical Institute.
- Ben-Zvi, D., Aridor, K., Makar, K., & Bakker, A. (2012). Students' emergent articulations of uncertainty while making informal statistical inferences. *ZDM: The International Journal on Mathematics Education*, 44, 913–925.
- Cañizares, M. J., Batanero, C., Serrano, I., & Ortiz, J. J. (2003). *Children's Understanding of Fair Games*. Paper presented in Working Group 5: The Third Congress of European Research in Mathematics Education (CERME3). See http://www.dm.unipi.it/~didattica/CERME3/proceedings/Groups/TG5/TG5_canizares_cerme3.pdf.
- Chernoff, E. J. (2014). Will the real Bayesian probability please stand up!? In K. Makar, B. de Sausa, & R. Gould (Eds.), *Proceedings of the 9th International Conference on Teaching Statistics*. Flagstaff, AZ, USA. Voorburg, The Netherlands: International Statistical Institute.
- Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Shauble, L. (2003). Design experiments in educational research. *Educational Researcher*, 32, 9–13.

- Engel, J., & Erickson, T. (2013). *Informal Inferential Reasoning: a Computer-based Training Environment*. Paper presented in the 59th ISI World Statistics Congress, Hong Kong.
- Falk, R., & Konold, C. (1992). The psychology of learning probability. In F. S. Gordon & S. P. Gordon (Eds.), *Statistics for the twenty-first century* (pp. 151–164). Washington, DC: Mathematical Association of America.
- Huber, B. L., & Huber, O. (1987). Development of the concept of comparative subjective probability. *Journal of Experimental Child Psychology*, 44(3), 304–316.
- Jacob, B., & Doerr, H. (2013). Students' informal inferential reasoning when working with the sampling distribution. In B. Ubuz, C. Haser, & M.A. Mariotti (Eds.), *Proceedings of CERME8* (pp. 829–839). Ankara: Middle East Technical University.
- Kazak, S., Fujita, T., & Wegerif, R. (2013). 'How confident are you?' Supporting Young Students' Reasoning about Uncertainty in Chance Games Through Students' Talk and Computer Simulations. In *The Eight International Research Forum on Statistical Reasoning, Thinking, and Literacy (SRTL-8)*. University of Minnesota: MN, USA.
- Kazak, S., Wegerif, R., & Fujita, T. (2014). Supporting students' probabilistic reasoning through the use of technology and dialogic talk. In S. Pope (Ed.), *Proceedings of the 8th British Congress of Mathematics Education* (pp. 215–222). Nottingham, UK.
- Konold, C., & Miller, C. D. (2011). *TinkerPlots2.0: Dynamic data exploration*. Emeryville, CA: Key Curriculum.
- Makar, K., Bakker, A., & Ben-Zvi, D. (2011). The reasoning behind informal statistical inference. *Mathematical Thinking and Learning*, 13, 152–173.
- Makar, K., & Rubin, A. (2009). A framework for thinking about informal statistical inference. *Statistics Education Research Journal*, 8, 82–105.
- Makar, K. (2014). Young children's explorations of average through informal inferential reasoning. *Educational Studies in Mathematics*, 86, 61–78.
- Martignon, L., & Erickson, T. (2014). Proto-Bayesian reasoning of children in fourth class. In K. Makar, B. de Sausa, & R. Gould (Eds.), *Proceedings of the 9th International Conference on Teaching Statistics*. Flagstaff, AZ, USA. Voorburg, The Netherlands: International Statistical Institute.
- Nilsson, P., Blomberg, P., & Årlebäck, J. B. (2014). Exploring realistic Bayesian modeling situations. In K. Makar, B. de Sausa, & R. Gould (Eds.), *Proceedings of the 9th International Conference on Teaching Statistics*. Flagstaff, AZ, USA. Voorburg, The Netherlands: International Statistical Institute.
- Papariotodemou, E., & Meletiou-Mavrotheris, M. (2008). Developing young students' informal inference skills in data analysis. *Statistics Education Research Journal*, 7, 83–106.
- 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, Bahia, Brazil. Voorburg, The Netherlands: International Statistical Institute.
- Pratt, D. (2000). Making sense of the total of two dice. *Journal of Research in Mathematics Education*, 31, 602–625.
- Rossman, A. J. (2008). Reasoning about informal statistical inference: one statistician's view. *Statistical Education Research Journal*, 7(2), 5–19.
- 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. Voorburg, The Netherlands: International Statistical Institute.
- Stohl, H., & Tarr, J. E. (2002). Developing notions of inference with probability simulation tools. *Journal of Mathematical Behavior*, 21(3), 319–337.
- Watson, J. M., & Moritz, J. B. (2003). A Longitudinal Study of Students' Beliefs and Strategies for Making Judgments. *Journal for Research in Mathematics Education*, 34, 270–304.
- Zieffler, A., Garfield, J., delMas, R., & Reading, C. (2008). A framework to support research on informal inferential reasoning. *Statistics Education Research Journal*, 7, 40–58.