# On Simulation and the Teaching of Statistics

**KEYWORDS:**
*Teaching;*
*Simulation;*
*Misconceptions.*

*Ted Hodgson and Maurice Burke*
Department of Mathematical Sciences, Montana State University, Montana, USA.
e-mail: hodgson@mathfs.math.montana.edu

**Summary**
The use of simulation as an instructional tool can promote a deep conceptual understanding of statistics *and* lead to misunderstandings. Teachers need to be aware of the misconceptions that can arise as a result of simulation and carefully structure classroom activities so as to derive the benefits of this powerful instructional tool.

## ◆ INTRODUCTION ◆

The power of simulation as a problem-solving tool is well known. With the current emphasis on projects, activities and hands-on learning, however, simulation plays an additional role in the statistics classroom. Namely, simulation serves as a *context* for teaching statistical concepts.

As an example, consider Hodgson and Borkowski's (1998) use of simulation and an unknown population of cards to illustrate the differences between two common sampling techniques. In the activity, students repeatedly collect simple and stratified random samples of size 4, calculate the corresponding sample means, construct histograms of the results and use their histograms to estimate the population mean. The population, which is listed in table 1, consists of 20 red-numbered and 20 black-numbered cards. Moreover, the colour of each card is indicative of its value: red cards correspond to 'small' numerical values, whereas the value of each black card is 'large'. Of course, with such a small population, there is no real need to sample. However, the only information that the students are provided with about the population is that there are equal numbers of red and black cards; they are not told about the numerical values on the cards, or about the size of the population. The population itself is contained in a paper sack. Thus, it is only through sampling that students can learn about the population.

| X | Colour | Frequency |
|---|---|---|
| 6 | red | 4 |
| 7 | red | 4 |
| 8 | red | 4 |
| 9 | red | 4 |
| 10 | red | 4 |
| 26 | black | 4 |
| 27 | black | 4 |
| 28 | black | 4 |
| 29 | black | 4 |
| 30 | black | 4 |

**Table 1.** The underlying population for the sampling simulation

Although the stated task focuses on the mean of the population, the primary instructional objectives are the sampling techniques. Specifically, each sampling technique generates a distinctive histogram and it is the analysis of these histograms that leads to a better understanding of the sampling techniques. With simple random samples, for example, the ratio of red to black cards varies. Some samples contain all red cards, some are all black, some contain a combination of the two. As a result, simple random samples yield widely varying estimates of the population mean, as is seen in figure 1. On the other hand, the ratio of red to black cards in each of the stratified samples mirrors that of the population. Simple random samples are drawn from each stratum, but the ‘appropriate’ balance of red and black cards in the stratified sample promises to yield a more precise estimate of the population mean, as shown in figure 2. Thus, the activity provides students with first-hand knowledge of the benefits that accompany stratified sampling.

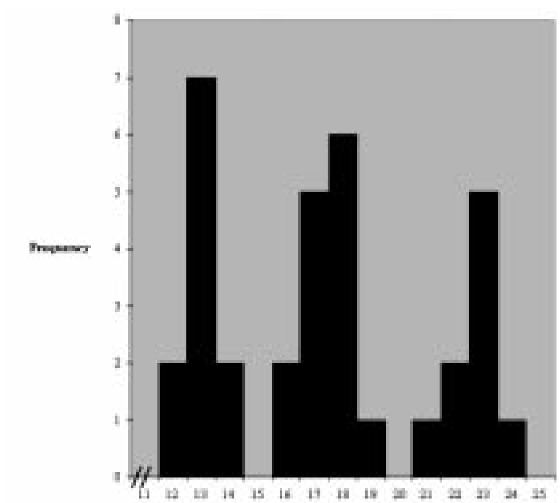

**Fig 1.** A frequency histogram of students’ ($N = 34$) estimates of the population mean, using simple random samples of size 4

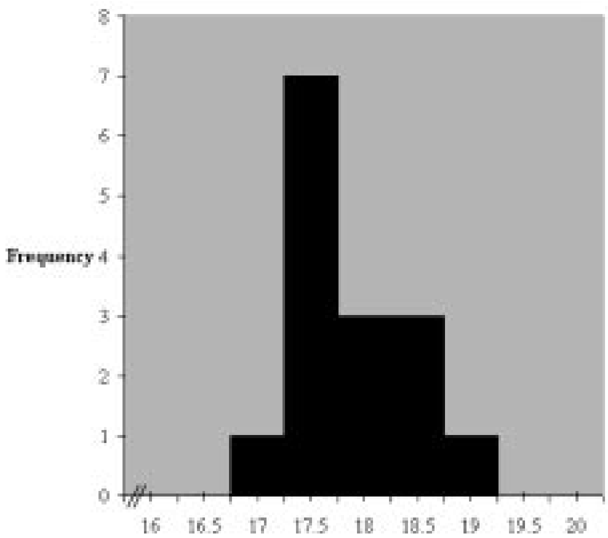


**Fig 2.** A frequency histogram of students’ ($N = 15$) estimates of the population mean, using stratified random samples of size 4

Despite its widespread use as an instructional tool, simulation has received relatively little attention in the research literature. In this article, therefore, we seek to promote discussion about this important instructional technique. To accomplish this objective, we first examine the rationale underlying its use. Why is it used and what are the expected benefits? Secondly, we briefly review the results of several studies, including two of our own, that assess the impact of simulation on students’ understanding of statistics. Lastly, drawing upon the existing literature and the results of our own preliminary investigations, we offer several recommendations regarding the effective use of simulation in the statistics classroom.

## ◆ THE PROMISE OF SIMULATION ◆

In his review of research on the teaching and learning of statistics, Shaughnessy (1992) claims that students emerge from traditional statistics courses with little more than mechanistic understanding. The use of simulation, on the other hand, promises to promote a conceptual understanding of statistics. For example, compare Hodgson and Borkowski’s (1998) development of sampling techniques with a more traditional approach to the topic. In the traditional presentation, students might read the definition of stratified random samples and the procedures by which they are collected. Subsequently, students would be given data that are subdivided into strata and asked to compute the sample mean. Occasionally, problems arise, such as how to allow for strata that differ in size. However, the primary focus of the activity is procedural – students learn *how* to compute stratified samples.

In contrast, Hodgson and Borkowski's activity encourages students to consider the conceptual underpinnings of stratified sampling. In particular, although the activity focuses on the mechanics of stratified sampling, students also consider the underlying *why* and *what* questions. Why use stratified sampling? What advantages does it offer? Additionally, the hands-on nature of the activity promises to lead to a deeper and more personalized level of understanding. Students are not simply told about the differences between the two sampling techniques. Rather, their understanding emerges from their actions – collecting samples, calculating the sample means, constructing histograms and comparing the results generated by each sampling technique.

In general, the instructional use of simulation promises to provide students with deeper conceptual understandings. Without widespread research on students' learning of statistics, however, we can question whether or not this promise is actually delivered. Fortunately, beliefs about the effectiveness of simulation-based instruction are supported by recent research. In a series of experiments with university students, for example, Garfield and del Mas (1989, 1994) report that computer-based simulation activities positively affect students' understanding of complex concepts, such as statistical power. Similarly, several recent studies report positive effects on students' conceptual understanding when traditional statistics curricula are restructured to include simulation and an active exploration of statistical concepts (Gnanadesikan *et al.* 1997; Romero *et al.* 1995; Sullivan 1995).

## ◆ THE PITFALLS OF SIMULATION ◆

Despite its intuitive appeal and emerging empirical support, the use of simulation provides no guarantee that students will form appropriate conceptions of statistical ideas. In their analysis of students' understanding of the Law of Large Numbers, for example, Well *et al.* (1990) discovered that many students do not understand the effect of sample size on variability, even after considerable experience with computer simulations. Likewise, Oursland (1997) reports that hands-on activities, especially those in which students use concrete materials to conduct experiments and collect data, do not always facilitate students' conceptual development and may actually hinder their efforts to learn and use statistics. According to Oursland, the random error inherent in hands-on measurement often misdirects students' attention and masks the underlying content objectives.

In our own research (see e.g. Hodgson 1996), we have uncovered similar pitfalls. In fact, we have found that not only does simulation provide no immunity against misconceptions, it can actually *contribute* to their formation. For example, in order to introduce the concepts of sampling distributions and the Central Limit Theorem, the students in one of our recent introductory statistics classes drew simple random samples of size 4 (with replacement) from a jar containing numbered slips of paper and calculated the sample means. The data were then pooled and students built histograms of the overall results, described the shape of the resulting distribution and used their data to draw inferences about the true mean of the population.

Subsequently, the students used an interactive TurboPascal program, constructed by one of the authors, to simulate the sampling activity. Specifically, the program allowed students to select an underlying population, the sample size and the number of samples to collect. It then collected samples of the indicated size, computed the sample means and graphed the relative frequency distribution of sample means, as shown in figure 3. Additionally, the program graphed the *expected* relative frequency distribution of values drawn from the underlying population. Of course, the results of individual trials will vary. Overall, however, the program allowed students to compare the centre and variability of the sampling distribution with that of the underlying population. Furthermore, since the underlying population remains constant, students were able to vary the sample size and observe its effect on variability. In theory, an understanding of the Central Limit Theorem would emerge as students observed the outcomes generated by large sample sizes and a large number of repetitions.

Following the completion of the activity and prior to any formal discussion of their observations, the students were administered an open-ended questionnaire that sought to uncover their understanding of sampling distributions and the

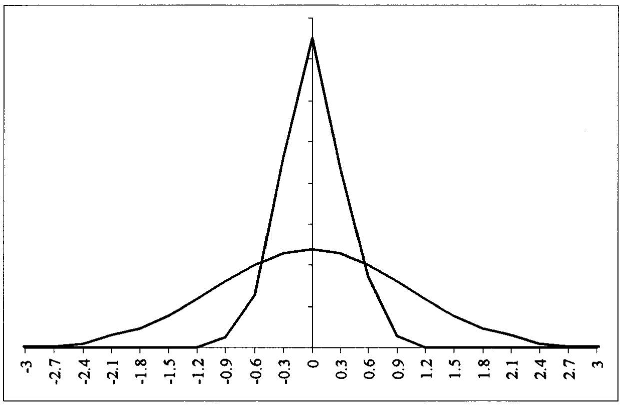


**Fig 3.** The *expected* relative frequency of a sample drawn from the underlying population distribution (standard Normal) and relative frequency distribution of 500 sample means for samples of size 9

Central Limit Theorem. Although students' responses indicated that they had indeed acquired some understanding of these statistical concepts, the activity also produced some unintended results. Specifically, one-third of the 18 respondents expressed the belief that one *must* draw multiple samples in order to make valid statistical inferences. Moreover, follow-up interviews indicated that these beliefs were directly attributable to the classroom activity, as the following excerpt indicates:

> What I learned about sampling distributions is that the sample mean is good at estimating the true mean, especially if the size of the sample is large. What our experiments showed, though, is that you've got to be careful. We can get good information about the population with sample means, *but only if we can take lots of samples.* (emphasis added)

In retrospect, we had hoped that students would emerge from the activity with greater confidence in sampling techniques (e.g. the sample mean represents an appropriate and relatively precise estimate of the population mean), and some understanding of the relationship between large sample sizes, increases in precision and the emergence of Normality. Students' responses did indicate progress toward these primary objectives, but they also acquired beliefs about repeated sampling and its role in statistical decision making. Namely, students acquired the belief that one *must* sample repeatedly to make valid decisions, which is erroneous.

## ◆ DISCUSSION ◆

Intuitively, the use of simulation can lead to the development of deeper conceptual understandings of statistics, and emerging research seems to support this claim. On the other hand, there is also evidence that hands-on instruction provides no guarantees about the formation of appropriate conceptions and may even promote the formation of misunderstandings. The effectiveness of instructional activities that employ simulation, therefore, does not seem to be an inherent property of the activities themselves, but depends upon the actions of teachers and students. Thus, the natural question to ask is 'How does one use simulation effectively to teach statistics?'

Although we do not presume to have the definitive answer to this question (if, in fact, one exists), we do believe that our preliminary research, and the findings of others, provide some clues. In particular, the discovery that simulation-based instruction may contribute to students' misunderstandings suggests that the effectiveness of these activities is intricately tied to students' sense-making efforts. In theory, students attend to the salient features of the activities and develop rich conceptions of the underlying statistical ideas. Our research suggests, however, that the 'salient' features of simulation-based activities may not be readily apparent to students. As a result, students attend to salient *and* non-salient features of the activities. Moreover, the data suggest that the focus on non-salient features of the activities is one source of statistical misconceptions.

In our sampling distribution activity, for example, students were to conduct simulations and develop generalizations on the basis of the results. From the students' perspective, however, repeated sampling masked the true intent of the activity. In particular, the students appeared to view the purpose of gathering large numbers of samples to be simply a real-world strategy for finding a population parameter. In so doing, they missed the real purpose of the simulation, which was to develop an understanding that, if certain conditions are met (e.g. the sample is sufficiently large), then there is a high probability that random samples will yield sample means that are 'near' the true mean of the population.

## ◆ RECOMMENDATIONS ◆

As teachers, it is important to remember that our perceptions of all activities, including those that employ simulation, differ from those of our students. Specifically, we interpret classroom instruction with the help of sophisticated patterning, problem-solving and generalization skills. Moreover, our familiarity with statistics allows us to focus on the concepts embedded in each activity. On the other hand, students cannot draw upon an extensive reservoir of skills and concepts. Without these tools, students naturally focus on the salient and non-salient features of the activities. For simulation to be an effective teaching tool, therefore, students must be able to discern the salient from the non-salient features.

Of course, the most obvious remedy is to provide students with an appropriate repertoire of skills and conceptual understanding. Constructivist theory clearly indicates, however, that we cannot simply provide them with appropriate skills and knowledge (Noddings, 1990). In fact, helping students to grasp the conceptual underpinnings of statistics is one of the goals of simulation-based instruction. Alternatively, teachers can enhance the effectiveness of simulation-based instruction by 'amplifying' the underlying statistical concepts and helping students to sift through the 'noise' associated with the simulation. To accomplish this objective, we recommend the use of a four-step instructional model. It should be noted, however, that these steps are not new to education. In fact, most are age-old, time-tested techniques. Nonetheless, we believe that the combination of these steps enhances the effectiveness of simulation-based instruction.

First, the use of *pre-organizers* can help students discern the salient from the non-salient features of an activity. As an example, prior to our sampling distribution activity, students' attention should have been directed toward the essentials of the activity, such as the relative proximity of the sample and population mean, the relationship between sample size and variability, and the overall shape of the sampling distribution. In so doing, students may have been able to avoid mistaken beliefs about the role of repeated sampling.

Second, *informal assessment* should be an integral and ongoing component of the exploratory process. Circulate about the room during an activity and assess *what* students are learning. Communicate with your students. Ask questions. Have students explain their reasoning. Have them identify the important statistical points, as they see them. Through informal assessment, teachers can identity emerging misconceptions and make the necessary instructional adjustments.

Third, the fact that students can acquire misconceptions highlights the importance of *debriefing*. Following the completion of an activity, have students share their beliefs and understandings. As they do so, amplify appropriate beliefs, build upon partial ones, and contrast appropriate beliefs with inappropriate ones. More often than not, students' misconceptions have a logical basis and arise from simple misunderstandings or inappropriate chains of inference. If students look back at their activity – and at the roots of their understandings – inappropriate beliefs can be revealed and corrected.

Finally, *follow-up* exercises can deepen students' understanding of the target statistical concepts and improve retention. Following the completion of Hodgson and Borkowski's sampling activity, for example, the authors encourage teachers to vary the numbers on the cards so that there is *no* relationship between a card's value and its colour. Subsequently, students collect simple and stratified random samples and examine the results. In this case, the two sampling techniques lead to similar estimates of the population mean. Thus, the extension allows students to develop deeper understandings of the appropriate uses of each technique.

## ◆ CONCLUSIONS ◆

Although this article identifies potential weaknesses in the use of simulation to teach statistics, it is not intended as an indictment of simulation or other student-centered instruction. The mathematics and statistics education literature is replete with evidence of the effectiveness of constructivist teaching. In statistics, instruction that incorporates simulation promises to help students acquire a conceptual, and not merely mechanical, understanding of the subject. Research suggests, however, that simulation-based activities provide no guarantee that students will acquire the intended conceptual understandings. If used in conjunction with appropriate introductory, informal assessment, debriefing and follow-up activities, we have found simulation to be an effective vehicle for teaching and learning statistics.

**References**

Garfield, J.B. and del Mas, R. (1989). Reasoning about chance events: assessing and changing students' conception of probability. In: C. Maher, G. Goldin and R.B. Davis (eds), *Proceedings of the Eleventh Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*, pp. 189–95. Rutgers, NJ: Rutgers University Press.

Garfield, J.B. and del Mas, R. (1994). Students' formal and informal understanding of statistical power. Paper presented at the Fourth International Conference on Teaching Statistics, Marrakesh, Morocco.

Gnanadesikan, M., Scheaffer, R.L., Watkins, A.E. and Witmer, J.A. (1997). An activity-based statistics course. *Journal of Statistics Education*, **5**(1).

Hodgson, T. (1996). The effects of hands-on activities on students' understanding of selected statistical concepts. In: E. Jakubowski, D. Watkins and H. Biske (eds), *Proceedings of the Eighteenth Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*, pp. 241–6. Columbus, OH: ERIC Clearinghouse for Science, Mathematics, and Environmental Education.

Hodgson, T. and Borkowski, J.J. (1998). Why stratify? *Teaching Statistics*, **20**(1), 68–71.

Noddings, N. (1990). Constructivism in mathematics education. In: R.B. Davis, C. Maher and N. Noddings (eds), *Constructivist Views on the Teaching and Learning of Mathematics* (*Journal for Research in Mathematics Education Monograph Number 4*), pp. 7–18. Reston, VA: National Council of the Teachers of Mathematics.

Oursland, M. (1997). *Comparing the Cognitive Differences Resulting from Modeling Instruction: Using Computer Microworld and Physical Instruction to model Real-World Problems.* Unpublished doctoral dissertation, Montana State University, Bozeman, MT.

Romero, R., Ferrer, A., Capilla, C., Zunica, L., Balasch, S., Serra, V. and Alcover, R. (1995). Teaching statistics to engineers: an innovative pedagogical experience. *Journal of Statistics Education,* **3**(1).

Shaughnessy, J.M. (1992). Research in probability and statistics: reflections and directions. In: D. Grouws (ed.) *Handbook of Research on Mathematics Teaching and Learning*, pp. 465–94. New York: MacMillan Publishing.

Sullivan, M.M. (1995). Development of conceptual understanding in descriptive statistics. In: D.T. Owens, M.K. Reed and G.M. Millsaps (eds), *Proceedings of the Seventeenth Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*, p. 338. Columbus, OH: ERIC Clearinghouse for Science, Mathematics, and Environmental Education.

Well, A.D., Pollatsek, A. and Boyce, S.J. (1990). Understanding the effects of sample size on the mean. *Organizational Behavior and Human Decision Processes*, **47**, 289–312.

# NEWS & NOTES

There is only one news item this time, concerning Larry Lesser who is one of the regular contributors to the journal. Larry has an article 'Sum of Songs: Making Mathematics Less Monotone!' in the May 2000 issue of Mathematics Teacher that shows how songs can be used in mathematics classes. 'Teaching Statistics' readers may find the first series of examples particularly interesting – they illustrate connections between probability/statistics content and contemporary lyrics in various genres. The article is also available on the internet at http://www.nctm.org/mt/2000/05/songs.html.

GERALD GOODALL, Editor

# LOOK AHEAD

Forthcoming articles include

- A new least squares regression model
- The toothless bathing beauty and the $t$ test
- Simulation using the TI-83