WATCHING HANDS SHUFFLE DATA IMPROVES SUBSEQUENT UNDERSTANDING OF R-BASED SIMULATIONS OF RANDOMNESS

Icy (Yunyi) Zhang, Mary C. Tucker, and James W. Stigler
University of California, Los Angeles
yunyi9847@g.ucla.edu

Simulation can be a powerful tool for teaching and learning statistics. Yet, students often struggle to understand the processes that underlie computer-based simulation activities. Recently, researchers have tried pairing computer-based simulation with more concrete, hands-on experiences to make simulation-processes more evident, but only in settings where students perform the hands-on simulation themselves. In this research, we ask whether manual simulation is necessary. Participants who were introduced to the shuffle function by watching a video of a hands-on activity before watching a video of live-coding performed better in a posttest than participants who were introduced to the shuffle function by watching an instructor write and run some R code and explain what that code did. This study provides pedagogical and practical insights into the use of hands-on demonstrations to complement computer simulation in remote teaching.

INTRODUCTION

The ability to make informal statistical inferences (i.e., to reason about different potential characteristics of the population based on the sample data, to deduce potential data generating processes based on differences observed from the sample data, and to evaluate whether a particular data generating process is likely based on simulated results is critical to statistical thinking and understanding (Zieffler et al., 2008). Engaging in informal statistical inference, especially in the early stages of learning, can improve students' knowledge representation and make statistics more accessible (van Dijke-Droogers et al., 2020). Acknowledging this need, educators have argued for a shift from a focus on formal statistical inference towards a "focus on the core logic of inference", using repeated simulation to facilitate students' thinking about random data generating processes and to prompt questioning about whether sample data could be generated by randomness alone. However, these recommendations have yet to be broadly implemented in introductory statistics courses. Part of the problem is that abstract statistical concepts like randomness and (simulated) sampling distributions are notoriously difficult for introductory students. Because randomness is complex in nature, most textbooks skirt the topic to avoid overburdening students and eliciting additional confusion (Liu, & Thompson, 2002).

Fortunately, recent advances in computing technology offer new ways of making complex concepts like randomness more accessible and concrete for students (Chance, & Rossman, 2006; Hodgson, Burke, 2000; Mills, 2002). One of these technological advances is simulation (Mills, 2002). Unlike traditional lecture-based instruction in which students are told about a concept, computer-based simulation activities allow students to experience abstract phenomena directly and thus construct a deeper, more flexible understanding of the concept and when it applies. Further, because simulation-based tools often present outcomes visually, they can help students see what data might be generated if an experiment were repeated multiple times (Hancock, & Rummerfield, 2020).

For example, students often struggle to grasp the concept that the sampling distribution of means is normally distributed, even for non-normal population distributions. Instructors often tell students the properties of sampling distributions and then represent those properties using mathematical equations. However, to develop a deeper, more flexible understanding of sampling distributions, students may need opportunities to see and experience the *process* of creating a sampling distribution themselves. With computer simulations, students can draw many random samples from non-normal distributions, calculate the mean of each sample, and create histograms to visualize concretely what the sampling distribution of means looks like (Hancock, & Rummerfield, 2020; Mills, 2002).

Simulation techniques have been used to help students understand a variety of statistical concepts in addition to sampling distributions, including the concepts of randomness (Chance, & Rossman, 2006; Zhang, & Maas, 2019) and statistical power (Garfield and del Mas, 1994). Although these and other studies provide some evidence that simulation can benefit learning, our review of the relevant literature suggests mixed evidence overall for the effectiveness of simulation as an instructional tool (delMas et al., 1999; Chance, et al., 2004; Lane 2015). Though some studies show benefits of simulation, others have shown that the use of computer simulations provides only limited benefit to students, and can in some cases impede learning by exacerbating students' misunderstandings or increasing their level of confusion (Watkins et al., 2014).

One possible explanation for why simulation has been ineffective in the past is that students lack the prerequisite knowledge and experiences to understand what the simulation is actually doing. To combat this issue, researchers have tried pairing computer-based simulation with more concrete, hands-on experiences (e.g., shuffling cards) to make simulation processes more understandable for students. Such hands-on experiences are thought to more actively engage students in understanding the processes that go on in the background of many pre-made statistical simulation packages (Pfaff & Weinberg, 2009). In one of the few experimental studies to date, Hancock & Rummerfield (2020) found a small but significant effect in which students learned more about the concept of sampling distributions when instruction with simulation applets was preceded by a hands-on activity. Clearly, more research is needed to replicate and understand the mechanism behind such effects.

In the current research, we ask whether it is necessary for students to perform the hands-on simulation themselves, or whether they might similarly benefit from watching, on video, another person perform such activities. The extensive literature on embodied cognition and gesture suggests that students may, in fact, learn as much from observing others perform hands-on activities as they would from performing those same activities themselves. For example, studies have shown that simply observing instructors' gestures improves learning of mathematics and statistics in both virtual and live classroom settings (Cook et al., 2013; Rueckert, et al., 2017; Son et al., 2018).

Our focus was on helping students understand the shuffle function in R as a means of simulating a random data generating process. We randomly assigned students to one of two conditions: a *hands-on* condition, in which students were introduced to the shuffle function by watching a video of a hands-on demonstration that involved shuffling cards; and a control (*live-coding*) condition in which students were introduced to the shuffle function by watching a video in which an instructor talked aloud and provided explanations as they wrote and executed R code. Based on evidence from the gesture and embodied cognition literature, we hypothesized that students in the hands-on condition would learn more than students in the live-coding condition.

METHOD

Participants

Thirty-three students from the University of California, Los Angeles participated in the study. They were specifically recruited because they had taken introductory psychological statistics during the preceding quarter and thus had a common set of background experiences relevant to the study. Students' statistics instructors from the prior quarter invited students to participate by email. Students were told that their participation would help the textbook authors to improve the book for future classmates. Those who chose to participate were given a five-dollar gift card after completing the study.

DESIGN & PROCEDURE

The study was conducted as a Qualtrics survey (https://www.qualtrics.com). On clicking the survey link, students were randomly assigned into one of the two conditions: *hands-on* (n = 18) or *live-coding* (n = 15).

Both versions of the survey were structured in the same way. It started with a question in which students rated their attitudes toward programming in R, followed by two free response questions designed to elicit their existing knowledge of the shuffle function in R. Then, participants in both conditions watched a series of two videos about the shuffle function and the concept of randomness. The second video was identical across the two conditions. It was a live-coding segment in which a narrator talked through a sequence of R commands and their interpretation. The first video was matched on content across the two groups, but one showed hands shuffling data (hands-on), similar to what might happen in an in-class hands-on exercise, while the other showed a screen recording of R being entered and run in a Jupyter notebook (similar to the live-coding video that followed).

After each video, participants rated how difficult the video was to comprehend. After watching both videos, participants answered 22 posttest questions assessing their understanding of the video, and some questions designed to get at how well they liked the exercise.

MATERIALS

The two videos shown first (one hands-on, the other live-coding) were essentially identical in content. Both videos explained how the shuffle function works. In the live-coding condition, the narrator's voice was recorded to explain what was happening in a video in which the screen showed code being typed and run in a Jupyter notebook (Figure 1). The hands-on video used an almost identical audio track, but instead of showing code being entered and run it showed a person's hands cutting a data table into pieces, and then re-arranging those pieces, simulating what would happen to a data frame when running the shuffle function. The only differences in narration across the two videos was in the language used to describe shuffling, be it physically with pieces of paper or with code using R (Kluyver et al., 2016).

The hands-on version of the video was recorded by placing a camera so as to look down from above on the hand movements of the instructor. The live-coding video was made by screen recording the instructor typing and running code in a Jupyter notebook. The second live-coding video (common across the two conditions) was similar in format to the first live-coding video.

The second live-coding video involved applying concepts learned in the first video to a larger dataset adapted from a real experiment. The dataset (called the laptop dataset) involved one independent variable (whether students viewed a laptop screen during class) and two dependent variables (students' final grades and students' self-rated level of distraction). In the second live-coding video, the instructor used the shuffle function in R to explore whether there was an effect of condition on these two outcome measures.

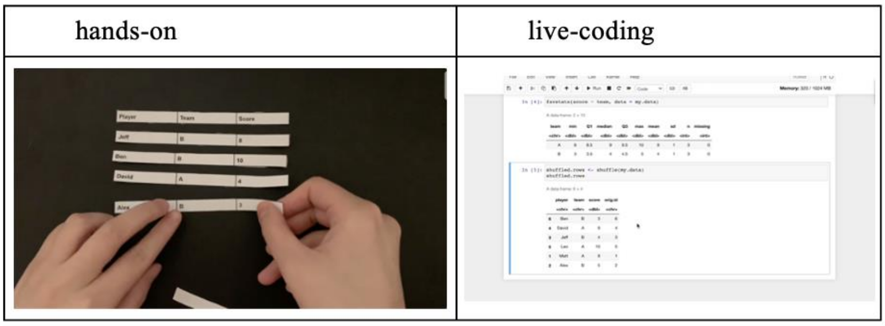

Figure 1. Screen grabs from the hands-on video (left) and the live-coding video (right)

MEASURES

Pre survey & Pretest

The pre survey asked students to rate, on a scale of 0 to 10, how well they understood the shuffle function.

The pretest contained two open response questions: “What does the shuffle() function do?” and “When would you use the shuffle() function?”. The purpose of the pretest was to make sure, given the small sample size, that the two experimental groups did differ in their understanding of the shuffle function prior to watching the videos.

Posttest & Post survey

The posttest contained 22 questions designed to assess students’ understanding of the shuffle function and the concept of randomness. It also included transfer questions that asked students to engage with statistical inferences. For example, it gave students one shuffled and one non-shuffled faceted histogram and asked students to reason whether there could be a difference between the two conditions. Each question was given a maximum of one point, with possible scores ranging from 0 to 22. The post survey asked students again to rate, on a scale of 0 to 10, how well they understood the shuffle function. It also asked students if they liked this way of learning R and whether they learned a lot from the activity.

RESULTS

An independent sample t test with condition as the independent variable and pretest score as the dependent variable did not find a significant difference between student’s prior understanding of the shuffle function across the two conditions ($t(31) = .17$, $p = .864$, $d = .06$).

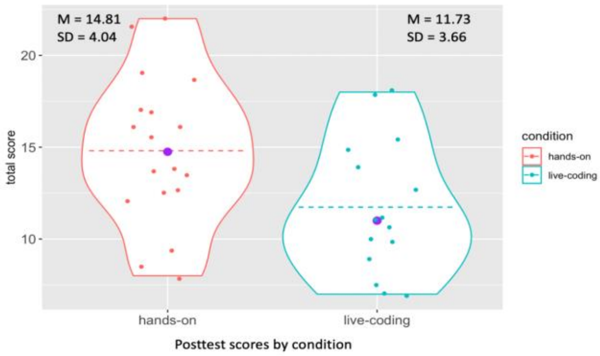

Figure 2. Violin plot showing posttest scores by condition. Dashed lines show the mean of each group. Purple dots show the median.

Figure 2 shows participants’ overall posttest scores by condition. Descriptively, participants in the hands-on condition performed better on average than participants in the control condition. An independent sample t test with condition as the independent variable and posttest score as the dependent variable found a significant difference between students in two conditions ($t(31) = 2.27$, $p = .031$, $d = .79$). When controlling for pretest performance by keeping it at constant in a regression with both pretest and condition as predictors, an independent sample t test yielded similar results: there was a significant difference between the two conditions on their posttest performance ($t(31) = 2.23$, $p = .033$).

Next, we examined participants’ self-rated understanding of the shuffle function before and after watching the videos. The modal student did not report any change in their understanding (with pre-rating minus post-rating equal to 0, see Figure 3). Though there did appear to be a slight advantage in self-rated learning for the hands-on group, the difference between conditions was not statistically significant ($t(31) = 1.30$, $p = .204$, $d = .46$).

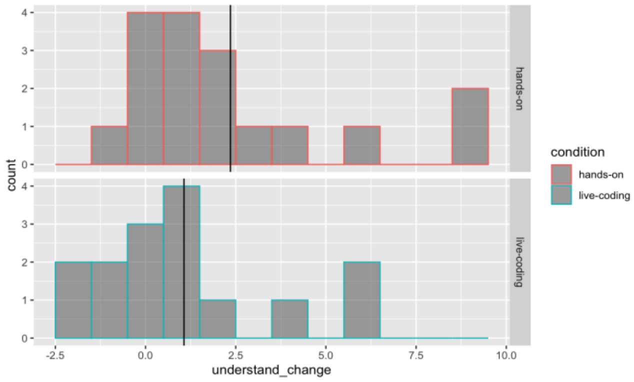

Figure 3. Histogram showing participants' change of self-rated understanding of the shuffle function by condition. Black line shows the mean.

DISCUSSION

In this study, we found preliminary evidence that preceding live-coding video with a video of showing a hands-on simulation of the shuffle function could improve students' understanding of the shuffle function and the concept of randomness. The study is, to our knowledge, the first to test experimentally whether students themselves must engage in hands-on experiences in order for those experiences to positively impact learning. It is also one of the first to consider coherently an educational context that is purely online: in this study, both the hands-on simulation and the live-coding simulation were delivered through instructional videos.

In terms of mechanism, the study suggests that it may not be the actual manipulation of materials that causes the effect of hands-on experiences, but instead that activation of embodied representations brought on by watching the video of hands shuffling pieces of paper. The embodied cognition literature would say that watching the movement of the hands is enough to activate another pathway for students to take in and process information, besides the already active pathway of language processing.

Practically, given the growing interest in using statistical programming languages like R as pedagogical tools, the findings of this study provides important and encouraging insights into the use of hands-on demonstrations to complement computer simulation in remote teaching. Physically cutting and shuffling pieces of paper can be environmentally demanding and time consuming for both in-person and online classes. The demand of physically performing those tasks may even add extraneous cognitive load and divert students off the intended path of understanding the underlying idea behind hands-on simulations. This study shows promising evidence that students themselves may not need to perform those hands-on simulations for the hands-on component to be effective. Instructors can thus make more efficient use of class time by editing and fastforwarding the instructional videos to only show students the relevant parts of the hands-on simulations.

Nevertheless, it is important to keep in mind that this study is still exploratory and is limited by its small sample size. Future studies should try to replicate the study with a larger sample size in order to draw more casual conclusions, and also compare participating in real hands-on activities with a condition in which the activities are viewed on video.

REFERENCES

Cook, S. W., Duffy, R. G., & Fenn, K. M. (2013). Consolidation and transfer of learning after observing hand gesture. *Child development*, *84*(6), 1863-1871.

Chance, B., delMas, R. C., and Garfield, J. (2004). "Reasoning About Sampling Distributions," in *The Challenge of Developing Statistical Literacy, Reasoning and Thinking*, eds. D. Ben-Zvi and J.Garfield, Dordrecht: Springer, pp. 295–323.

Chance, B., & Rossman, A. (2006, July). Using simulation to teach and learn statistics. In Proceedings of the Seventh International Conference on Teaching Statistics (pp. 1-6). Voorburg, The Netherlands: International Statistical Institute.

delMas, R., Garfield, J., and Chance, B. L. (1999). "A Model of Classroom Research in Action: Developing Simulation Activities to Improve Students' Statistical Reasoning," *Journal of Statistics Education*, 7.

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

Hancock, S. A., & Rummerfield, W. (2020). Simulation methods for teaching sampling distributions: Should hands-on activities precede the computer?. *Journal of Statistics Education*, *28*(1), 9-17.

Hodgson, T., & Burke, M. (2000). On simulation and the teaching of statistics. Teaching Statistics, *22*(3), 91-96.

Kluyver, T., Ragan-Kelley, B., Pérez, F., Granger, B. E., Bussonnier, M., Frederic, J., ... & et al. (2016). Jupyter Notebooks-a publishing format for reproducible computational workflows (Vol. 2016, pp. 87-90).

Lane, D. M. (2015), "Simulations of the Sampling Distribution of the Mean Do Not Necessarily Mislead and Can Facilitate Learning," Journal of Statistics Education, 23. Available at https://doi.org/10.1080/10691898.2015.11889738.

Liu, Y., & Thompson, P. (2002, October). Randomness: Rethinking the foundation of probability. In *Proceedings of the twenty fourth annual meeting of the North American chapter of the International Group for the Psychology of Mathematics Education*. Athens, GA (Vol. 3, pp. 1331-1334).

Mills, D. M. (2002). Using computer simulation methods to teach statistics: A review of the literature. Journal of Statistics Education, *10*(1).

Mills, J. D. (2004). Learning Abstract Statistics Concepts Using Simulation. Educational Research Quarterly, *28*(4), 18-33.

Pfaff, T. J., and Weinberg, A. (2009). "Do Hands-on Activities Increase Student Understanding?: A Case Study," *Journal of Statistics Education*, 17. Available at https://doi.org/10.1080/10691898.2009.11889536

Rueckert, L., Church, R. B., Avila, A., & Trejo, T. (2017). Gesture enhances learning of a complex statistical concept. Cognitive research: principles and implications, *2*(1), 1-6.

Son, J. Y., Ramos, P., DeWolf, M., Loftus, W., & Stigler, J. W. (2018). Exploring the practicing-connections hypothesis: Using gesture to support coordination of ideas in understanding a complex statistical concept. Cognitive research: principles and implications, *3*(1), 1-13. http://dx.doi.org/10.1186/s41235-017-0085-0

van Dijke-Droogers, M., Drijvers, P., & Bakker, A. (2020). Repeated sampling with a black box to make informal statistical inference accessible. *Mathematical Thinking and Learning*, *22*(2), 116-138.

Watkins, A. E., Bargagliotti, A., and Franklin, C. (2014). "Simulation of the Sampling Distribution of the Mean Can Mislead," *Journal of Statistics Education*, 22. Available at https://doi.org/10.1080/10691898.2014.11889716.

Zhang, X., & Maas, Z. (2019). Using R as a Simulation Tool in Teaching Introductory Statistics. *International Electronic Journal of Mathematics Education*, *14*(3), 599-610.

Zieffler, A., Garfield, J., Delmas, R., & Reading, C. (2008). A framework to support research on informal inferential reasoning. *Statistics Education Research Journal*, *7*(2).