Title
Utilizing Dynamic and Embodied Visualization to Facilitate Understanding of Normal Probability Distributions

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 43(43)

ISSN
1069-7977

Authors
Zhang, Icy(Yunyi)
Son, Ji
Blank, Idan
et al.

Publication Date
2021

Peer reviewed

Utilizing Dynamic and Embodied Visualization to Facilitate Understanding of Normal Probability Distributions

Icy(Yunyi) Zhang (yunyi9847@g.ucla.edu)
Department of Psychology, University of California, Los Angeles
760 Westwood Plaza, Los Angeles, CA 90024 USA

Ji Y. Son (jiyunson@gmail.com)
Department of Psychology, California State University, Los Angeles
5151 State University Dr., Los Angeles, CA 90032-8227 USA

Idan A. Blank* (iblank@psych.ucla.edu)
Department of Psychology, University of California, Los Angeles
760 Westwood Plaza, Los Angeles, CA 90024 USA

James W. Stigler* (stigler@ucla.edu)
Department of Psychology, University of California, Los Angeles
760 Westwood Plaza, Los Angeles, CA 90024 USA

Abstract

Teachers often use drawings of the normal distribution to support explanations of related statistical concepts, assuming that the normal curve provides a common language for such discussions. However, we find that students may not understand the basic features of the normal curve. In Study 1, we showed that students who already have studied the normal distribution in a college-level class do not understand basic concepts associated with it. Then, in Study 2, we investigated whether a brief instructional, narrated video could improve students' understanding of the normal probability distribution. Specifically, we compared three instructional formats: static slides, a video recording of a hand physically drawing those plots, and a screen recording of the hand-drawing. Despite the brevity of the intervention, we found significant improvements in students' understanding of the normal probability distribution and related probability concepts. The findings are discussed in relation to the dynamic representation and embodied cognition literature.

Keywords: dynamic visualization, embodied representation, statistics education, learning media

In college-level introductory statistics classes, understanding the concept "distribution", and particularly the concept "probability distribution", is an important step for students as they transition from data analysis to statistical reasoning and inference (Batanero et al., 2004). Even beyond the technical importance of probability distributions, teachers often use visual representations of such distributions—especially normal distributions—as a way to explain other, more complex, concepts, such as the *p*-value that results from a statistical hypothesis test (Ainsworth, 2008; Batanero et al., 2004). Yet students, and especially novice learners, often have difficulty connecting fundamental concepts to the visualizations we use to explain them (Rau, 2017). In essence, teachers may be asking students to make sense of novel concepts via visualizations they do not understand (Airey & Linder, 2009).

Perhaps because teachers themselves have a solid grasp of probability distributions—including, in particular, normal distributions—they may assume that students understand, at least, what it means to interpret the normal distribution as a probability distribution. For example, when teachers shade in some part of a normal curve to represent the probability represented by that area, they assume that students already understand that the total area under the curve equals 1. However, such assumptions might often remain untested in the classroom, and they may not be valid (Rau, 2017; Airey & Linder, 2009).

To more closely examine these issues, the studies reported here had two objectives. First, we explored what students do and do not understand about normal distributions and the visualizations we use to represent them (Study 1). To this end, we surveyed a small sample of students nearing the end of a college-level introductory statistics class that had explicitly taught them about normal distributions. To foreshadow our results, we found evidence that many students, even after instruction, had only a weak understanding of what normal distributions were, with many misconceptions and confusions. This finding raises troubling questions about how strongly teachers should rely on normal distributions as a means for communicating about other statistical concepts (Chance et al., 2004).

Second, based on these findings, we attempted to teach students how to interpret common representations of normal probability distributions (Study 2). Specifically, we created instructional interventions in the form of brief, narrated videos that could be added as supplementary materials to ongoing statistics classes (we chose such interventions, in part, because our research was carried out during the COVID-19 pandemic, when classes were offered remotely). We aimed to provide a brief intervention that, nonetheless, would have lasting effects that could, in theory, support students in better interpreting future explanations of statistical concepts couched in terms of the normal curve. Moreover, we compared several, closely matched forms of interventions, and tested their tested their effects on both immediate and delayed post-tests.

In designing the instructional videos we drew heavily on two areas of study in Cognitive Science. First, we studied ”dynamic visualizations”, which replace the static nature of many online instructional videos, as well as many slides used in classrooms, with animations where visual plots gradually appeared on the screen, one sub-component after another, as if they were being drawn (in the style of Khan Academy) (Castro-Alonso,et al., 2015). We tested whether such dynamic representations could improve students' comprehension over and above what they might get from static slides. Second, we studied ”embodied representations”, i.e., videos that could potentially engage students' perceptual-motor representations to scaffold their understanding (de Koning & Tabbers, 2011). Specifically, we tested whether augmenting dynamic animations with a human hand that physically draws plots would enhance their effectiveness.

Prima facie, dynamic visualizations such as real-time drawing or animations would appear superior to static representations because they can explicitly represent processes, not just the end results of processes (Castro-Alonso et al., 2014; Hegarty, 2004; Chandler, 2004; Mayer & Moreno, 2002). Moreover, they may reduce cognitive load by ”distributing” new information across time (for a review, see Ainsworth, 2008) and, relatedly, better match the computational demands of learning (Tversky et al., 2002). Moreover, they may be more motivating than still images (Rieber, 1991). However, the advantages of dynamic representations are not clear cut. Although dynamic representations do tend to yield better implicit learning and learning attitudes (Lowe, 1999; Wright et al., 1999), they do not necessarily produce better conceptual learning outcomes (Hegarty, 2004; Tversky et al., 2002). One reason for this may be that dynamic representations, despite being engaging, might tax working memory due to their transient nature (Hegarty, 2004; Lowe, 2004; Chandler, 2004). When the video or the animation has advanced, it is gone - learners cannot access it anymore.

One way to potentially render dynamic visualizations more effective is by connecting them to the learner's physical experience in the world, via bodily actions such as gesture (de Koning & Tabbers, 2011). In other words, ”embodied” cognitive processing could aid students in reaping the benefits of dynamic visualizations. Whereas most of the literature of embodied cognition focuses on learning procedures and motor tasks, a growing body of work demonstrates that embodiment could support the development of conceptual understanding and higher-order skills such as problem solving and reading comprehension (Glenberg et al. 2008; Thomas & Lleras 2009; Zhang et al., 2021). Importantly, the facilitative effect of bodily movements during learning has been demonstrated even when the learners themselves are not the source of the bodily action but are simply observing others' bodily movements in a video (de Koning & Tabbers, 2013; Son et al., 2018; Glenberg et al. 2008; Thomas & Lleras 2009).

Based on these considerations, our instructional narrated video of a human hand drawing visual figures combined both dynamic and embodied qualities. Drawing as a learning tool has been studied independently of dynamic visualization and embodied cognition, and has been found to have certain advantages in its own right. Whether generated by the learner, or observed as it is generated by someone else, drawing can give learners more time to notice more (and more subtle) details in new representations, allowing them to think about how different elements of the drawing are connected in a process that unfolds over time (Landin, 2011).

Below, we first evaluate what students know about normal distributions and, then, test whether our intervention improves their understanding of this concept.

Study 1

Method

Participants Participants were 39 undergraduate students taking a 10-week, introductory statistics course at the University of California, Los Angeles. Due to the COVID-19 pandemic, the entire course was taught remotely (online). Students participated in the study for extra credit toward their final course grade and did not get any other form of compensation.

Design & Procedure Students were emailed an invitation to participate in the study near the end of the course. By that point, students had already been taught the basic concepts pertaining to normal distributions, and had used the standard normal distribution to calculate *p*-values for simple statistical tests. Students who chose to participate clicked a link to complete a Qualtrics survey (https://www.qualtrics.com). Following the survey, students were asked to (1) rate the difficulty of the survey as a whole, on a scale from 0 to 10 (0 = not hard at all), and (2) estimate how well they thought they performed, on a scale from 0 to 10 (0 = merely guessing answers). Participants could not go back across questions.

Materials The survey consisted of 15 questions about probability; of these, four questions specifically examined students' understanding of probability under a normal curve (for details, see **Results**). For example, we asked whether, and why, the area under a normal curve equaled 1; and tested students' ability to use the symmetry property of the normal curve to compare probabilities across its two sides. Below, we do not report results on the other 11 questions, which addressed topics such as *p*-value interpretation and statistical power.

Results

On average, students rated the difficulty of the questions as 6.44 out of 10 (SD = 1.89), and estimated their own performance as 4.80 out of 10 (SD = 2.05). Below, we qualitatively describe students' responses to each of four questions testing their understanding of the area under a normal curve.

In the first question, students were presented with a normal curve whose entire area was shaded (Figure 1A) and were told: “This is a normal probability distribution.” They were then asked if they could estimate the probability represented by the shaded region under the curve, and elaborate on their answer in an open response question.

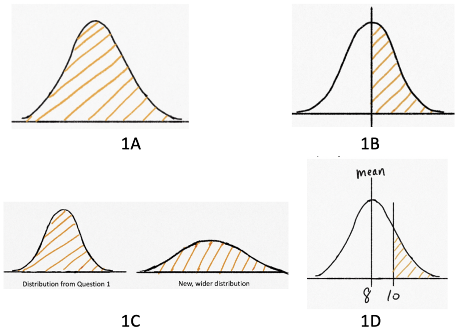

Figure 1: The figures presented to students in the survey

Out of 39 students, only 25 (64%) correctly answered that the total probability under the curve equals 1. Out of the 14 students who answered the question incorrectly, 10 students (25.6% of the total sample) said that the probability could not be estimated because there were no numbers on the *x*-axis. One student wrote: ”Since this is a bell shaped curve there is an equal amount of values before and after the median in the center of the curve. Given that there are no values on the graph I wouldn’t be able to estimate a specific number to represent the shaded region.”

The other 4 students either said that the area could be estimated but did not provide a specific number, or gave a wrong explanation. One student wrote that, “Yes you can, why, because the area under the curve, my estimate would be depending what it is asking and deal with *Z* scores”. The concept of a *Z* score was a frequently mentioned in students’ explanation, with many students thinking, erroneously, that a *Z* score was required to compute the total probability under the curve.

The second question presented students with another normal curve, with its peak (i.e., mean/median/mode) marked with a vertical line; only the area to the right of the line was shaded (Figure 1B). We asked students whether they could estimate the probability represented by the shaded area. In contrast to question 1, not all 39 students answers “Yes” (but two students did not provide a specific number). Most student—including the students who previously said they could not estimate the probability when the total area under the curve was shaded—seemed to understand the idea that half of the area represents a probability of 0.5. Given that knowing the probability of the entire area under the normal curve is, at least implicitly, logically required for calculating the area corresponding to half this distribution, it is surprising that the same students who said that numerical values were needed for estimating probability on the previous question now seemed to have no problem generating a probability. One of these students said, “The probability is 50% since half of the data points fall under the shaded region.”

The third question again presented students with the curve in Figure 1A but, this time, paired it with a different, wider, normal curve. The total area under each curve was shaded (Figure 1C). We asked students: “If we draw a normal distribution that is wider than the one in Question 1 (as shown below), how would the probability represented by the shaded part under the distribution change?” Only 14 students (35.9%) answered correctly that the probability would not change and provided a reasonable explanation of their answer. For example, one student answered: “The probability is still 100% because the whole distribution is shaded in”; another said: “It would not change at all. The area under the curve still represents the entire probability.”

The 24 students who answered Question 3 incorrectly made 3 main types of errors: (1) 11 students said that the probability would change if the distribution became wider. One of these students said: “The probability would change to encompass fewer Y values and more X values.” Another said: “The original distribution is normal and the wider distribution is not. The empirical rule only applies to normal distributions. So indicators of 68% or 2.5% would not exist.” (2) Seven students did not say whether the probability represented by the shaded region would change or not. For example, one student said only that “the peak is higher than the wider one.” (3) The remaining 7 students said that the probability would not change, but were not able to provide a sensible explanation of their answer. For example, one said: “ I don’t think it would change, making it wider would only help people clearly see the distinction between the *x*-axis, but I don’t think anything more.”

Finally, Question 4 tested whether students could use the symmetry property of the normal distribution to reason about probabilities. Therefore, it presented students with a normal curve having a mean of 8, and a vertical line marking a value of 10 on the *x*-axis. The area to the left of this line was shaded (Figure 1D). Students were told:

”Here’s a drawing of a normally shaped population distribution with a mean of 8. The probability of a randomly sampled data point being greater than 10 is 0.2. Based on this, what is the probability of a randomly sampled data point being greater than 6? Explain your answer.”

14 students (35.9%) correctly reasoned that the probability should be 0.8. The remaining 25 students, who provided incorrect probabilities, came up with a variety of explanations. As before, five students erroneously tried using the concept of *Z* scores or the ”empirical rule”, a shorthand to remember the percentage of values that fall within each standard deviation of the normal distribution, to explain their answers. One, for example, said:

”If the probability of a random data point being greater than 10 is .2, then 10 has a *Z* score of 2. This means that a change in value from 8 to 10 is measured in 2 standard deviations, so 6 to 8 is another 2 *Z* scores. So the probability of a randomly sampled data point being greater than 6 is about 98%, because it is represented by the area of the normal distribution above -2 standard deviations from the mean. ”

Another student said that the probability would be 0.6 "because there is a Z score of -2."

Discussion

The results of Study 1 showed that even students nearing the completion of an introductory statistics course at a highly selective university have, for the most part, only a shaky grasp of the concept of normal probability distributions. In particular, many students do not fully understand how the area under a normal curve can be used to represent probability, or that the total area under any probability distribution (regardless of its shape) would add up to 1. Students also are not generally able to infer probabilities based on the symmetric property of a normal probability distribution—at least in scenarios that are somewhat challenging—and often resort to concepts such as Z scores or the central limit theorem, inappropriately applying unnecessary or irrelevant concepts to the problem at hand.

Together, these results suggest that when teachers draw probability distribution curves on the board or display them on slides as a means of communicating statistical concepts, their students may not even interpret that basic features of such displays in the intended ways. For us, this raised the question of whether we could remedy students' misconceptions about probabilities under the normal curve through a brief instructional intervention, which we set out to do in Study 2.

Study 2

In Study 2 we set out to create a brief instructional intervention that, if successful, could provide students with the fundamental knowledge they would need to interpret the kinds of visual representations teachers commonly use to explain more advanced statistical concepts. The focus of the instruction was on the normal curve and its use as a probability distribution for modeling the distribution of a variable. Effects of the training were assessed on both immediate and delayed post-tests. The instruction was implemented in the form of a brief (16min) video, of which we created three versions. Students were randomly assigned to view one of the three versions:

For the first version of the video—the *Hand Drawing* condition (Figure 2, left)—we videotaped a human hand drawing on an iPad, and then recorded a voice narration of what was being represented in the drawing as it unfolded through time. We decided to use a narrated hand drawing because the resulting representation is both dynamic and embodied. Based on our reading of the literature, both of these features, especially when used together, might be expected to help students link the visual features of probability distributions to their conceptual meanings. Further, we chose this particular version of embodied representation because it is simple to implement: if this intervention works, then many instructors who create online materials in the style of Khan Academy (i.e., animated, dynamic drawings with no hand) could easily adjust their presentation to include their hand. In addition, compared to other options such as gesturing hands or a talking head, a hand is a less confounded manipulation of embodied representations, because it adds relatively little information that is not directly related to the statistical figures being drawn, such as communicative gestures or facial expressions.

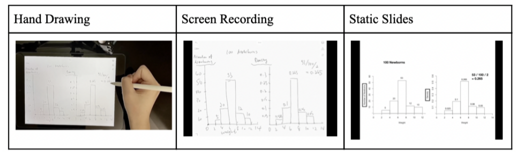

Figure 2: screenshot of instructional videos from the three experimental conditions in Study 2

For the second version—the *Screen Recording* condition (Figre 5, right)—we used the same audio track as in the *Hand Drawing* condition, but this time paired it with the iPad's screen recording of the drawing as it was being produced for the *Hand Drawing* condition. Thus, the only difference between these first two versions was in whether the hand could be seen doing the drawing or not, which allowed us to gauge the effect of embodiment over and above the effect of the dynamic representation without the hand.

Finally, the third version of the video—the *PowerPoint Condition* (Figure 2, right)—used the same audio track, but instead of showing the drawing unfold dynamically over time, it displayed a series of static PowerPoint slides. The content of the slides was designed to match the final output of the hand drawings in the other two conditions. Based on concerns raised in the previous literature, we tried to equate as much as possible the information contained across the three versions of the video.

In addition to an immediate assessment following the instructional video, we included a delayed assessment, which was administered several weeks after the initial video intervention. Measuring delayed effects is important for two reasons: (1) to tests whether learning lasts and can generalize beyond a single and controlled laboratory session (Halpern & Hakel, 2010; Stigler et al., 2019); and (2) because sometimes the impact of an intervention—on transfer in particular—shows up only in a delayed post-test (Adams et al., 2014; McLaren, Adams & Mayer, 2015) .

Two predictions can be derived from this design: first, if the *Screen Recording* group performs better on the post-test than the static *PowerPoint* group, that would mean that dynamic visualizations (here, drawing) aid learning over and above a presentation of the same content in static slides. Second, if the *Hand Drawing* group performs better on the post-test than the *Screen Recording* group (and the static *PowerPoint* group), it would demonstrate the unique value of adding an embodied element to dynamic visualizations. If we find support for only the second, but not first, prediction, it would suggest that dynamic visualizations on their own are not always sufficient to support learning about normal probability distributions.

Method

Participants Seventy-nine undergraduate students taking an introductory statistics course (Psychological Statistics) at University of California, Los Angeles during the summer session, participated in the study. Of these students, 71 took the delayed post-test. From among these, eight participants were excluded from the study based on predetermined exclusion criteria, which included (1) spending either less than 400s or more than 7200s on the survey; (2) reporting significant technical difficulties or disruptions while completing the survey (e.g. not having a quiet enough study environment for them to watch the instructional video); or (3) writing the same response for every free response question. Thus, following exclusion, we obtained a final sample of 63 undergraduates. The sample was ethnically diverse: 50.79% Asian, 4.76% Black or African American, 12.70% Hispanic or Latino, 23.81% White, and 7.93% multiracial or other.

Design & Procedure Similar to Study 1, students who wanted to participate in a Qualtrics survey for extra credit voluntarily clicked on the link, at which point the survey software randomly assigned them to one of three conditions: *Hand Drawing* (N=16),*Screen Recording* (N=23), or (static) *PowerPoint* (N=24).

All three conditions included an initial survey, followed by an instructional video, followed by an immediate and, later, a delayed post-test. In the initial survey, participants completed a pre-test with seven questions assessing their understanding of frequency histograms, density histograms, and probability distributions. Then, participants watched a 16min instructional video, which varied according to which condition they were assigned to. They then rated the pace of the video and how much of the video they felt they understood. Afterwards, they answered 15 post-test questions, three survey questions about their attitudes towards the video, and one screening question asking if anything went wrong during the experiment.

Approximately three weeks later, when students took their final course exam, they were informed that they could get additional extra credit by taking the delayed post-test. Students participated voluntarily in this activity. Importantly, the material covered in the course between the immediate and delayed post-tests did not vary across experimental groups. The material focused on using a modelling approach to explain variation. Only students who completed both post-test evaluations were included in our data analysis of the delayed post-test.

Materials Three versions of videos, each teaching the same set of basic concepts—"histogram", "normal distributions", and "probability distributions"—were designed. The *Hand Drawing* version of the video was made by using a video camera to record an instructor's hand drawing on a tablet according to a lesson plan for teaching the statistical concepts above (the video camera was external to the tablet). The *Screen Recording* version of the video was created using the tablet's internal screen-capture technology to record the drawings from the first video, without the actual hand, but with the same audio. The *PowerPoint* version of the video was contained static slides with recreated copies of the "end-state" of each screen in the first video, with the same audio.

Measures *Pre-test*. The pre-test contained seven questions designed to assess participants' existing knowledge of normal probability distributions. These same questions were also included on the post-test.

Immediate post-test. The immediate post-test contained 17 questions. These questions targeted conceptual understanding of the area and the probability under a normal probability distribution, its symmetry property, the features of a faceted histogram, the idea of using the normal distribution as a generative model of data, and the changes in probabilities as distributions become wider or narrower.

Delayed post-test. The delayed post-test contained 17 questions. Seven questions were duplicates of questions included on the immediate post-test, and the rest were new questions.

Three trained coders, blind to each participant's experimental condition, scored students' responses on both the immediate and delayed post-tests. Each question was randomly assigned to be coded by two coders. Disagreements in coding were discussed in a group meeting until a consensus was reached.

Results

Pre-test A one-way Analysis of Variance (ANOVA) with condition as the independent variable and pre-test score as the dependent variable found no significant difference between the three conditions ($F_{(2,60)}$ = .153, p = .858, η^2 = 0.005).

Immediate post-test Scores on the immediate post-test for each of the three conditions are shown in Figure 3. Descriptively, students in the *Hand Drawing* condition had higher scores than those in the other two conditions. A one-way ANOVA found a significant difference in post-test performance across the three groups ($F_{(2,60)}$ = 4.191, p = .020, η^2 = .12; Levene's test and normality checks were carried out and the data met the assumptions). Pairwise t-tests showed a significant difference between the *Hand Drawing* group and the (static) *PowerPoint* group ($t_{(38)}$ = 2.90, p = .005, Cohen's d = .92), but not between the *Hand Drawing* group and the *Screen Recording* group ($t_{(37)}$ = 1.69, p = .097), nor between the *PowerPoint* group and the *Screen Recording* group ($t_{(45)}$ = 1.32, p = .192). Pairwise, post-hoc comparisons were adjusted using the Bonferroni correction to 0.017 (0.05/3).

We computed the gains from questions that appeared both on pre and post test across three conditions by subtracting pretest score from a subset of immediate post-test score. *Hand Drawing* group had a mean gain of 2.03 (sd = 1.96). *Screen Recording* group had a mean gain of 1.04 (sd = 1.63). *PowerPoint* group had a mean gain of 0.85 (sd = 2.08). Critically, however, we did not include a control group that received no intervention.

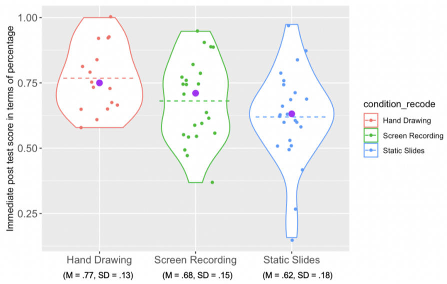

Figure 3: Violin plots showing immediate post-test scores by condition. Dashed lines are means. Purple dots are medians.

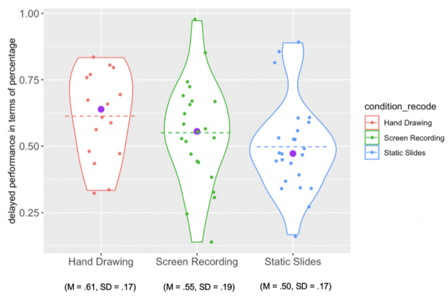

Figure 4: Violin plots showing delayed post-test scores by condition. Conventions are the same as in Figure 3.

To compare the effectiveness of the three interventions while controlling for potential differences across groups in pre-test performance, we conduced a one-way Analysis of Covariance (ANCOVA). The difference on immediate post-test performance between the three conditions remained significant ($F_{(2,59)}$ = 4.567, p = .014, η^2 = .14). Followup pairwise comparisons showed a significant difference between the *Hand Drawing* group and the *PowerPoint* group ($t_{(38)}$ = 3.07, p = .003) but not between the the *Hand Drawing* group and the *Screen Recording* group ($t_{(37)}$ = 1.79, p = .079), nor between the *PowerPoint* group and the *Screen Recording* group ($t_{(45)}$ = 1.40, p = .167).

Delayed post-test Scores on the immediate post-test for each of the three conditions are shown in Figure 4. Descriptively, the ordering of the three groups remained the same as during the immediate post-test. Nonetheless, a one-way ANOVA found no significant differences across conditions ($F_{(2,60)}$ = 1.978, p = .147, η^2 = .062). Still, pairwise t-tests showed a significant difference between the *Hand Drawing* group and the *PowerPoint* group ($t_{(38)}$ = 1.98, p = .052, d = .67), but not between the *Hand Drawing* group and the *Screen Recording* group ($t_{(37)}$ = 1.06, p = .293), nor between the *PowerPoint* group and the *Screen Recording* group ($t_{(45)}$ = 1.01, p = .316). When controlling for pre-test performance by keeping it constant, we again found a significant difference between the *Hand Drawing* group and the *PowerPoint* group ($t_{(37)}$ = 2.03, p = .046). However, these test results did not survive Bonferroni correction for multiple comparisons.

Discussion

In Study 2 we used a brief intervention to teach students concepts related to probability distributions. Specifically, we created a ”hand-drawing” video, utilizing both dynamic visualizations and embodied representations. We found that this 16-min video improved understanding compared to the ”static” PowerPoint intervention, which did not have these characteristics. The intervention that used dynamic visualizations without a visible hand produced outcomes in between the other two conditions, but did not differ significantly from either. This pattern was descriptively maintained three weeks after the intervention, although the differences were reduced.

This pattern of results provides evidence for the potency of combining dynamic visualizations with embodied representations. Moreover, the lack of significant difference between still PowerPoint slides and the animated (dynamic) Screen Recording is in alignment with the hypothesis that dynamic visualizations alone may overly tax limited working memory resources, making it difficult for learners to remember elements of the dynamic visualization for long enough to construct understanding. The embodied quality of the human hand in the Hand Drawing intervention may help to alleviate this problem by activating another pathway (i.e. the bodily movement of the instructor) for learners to encode and understand information, resulting in better learning outcomes. In addition, the findings are consistent with Mayer’s multi learning principles, in which embodiment such as drawing with an visible hand is hypothesized to help learning, especially when it guides specific cognitive processes (Mayer, 2014).

General Discussion

The findings altogether have two implications for statistics instructors: First, students may not have the fundamental understandings teachers assume they have when they use visualizations based on the normal curve to explain more advanced statistical concepts (e.g., p-values). But second, this gap in knowledge can be remedied with a brief intervention that could be delivered online, outside of class time, providing students with the basis to learn more advanced concepts.

Although we suspect that the most common instructional materials in current use are relatively static in nature (like the PowerPoint condition in our study), we have demonstrated that students can learn more from representations that are both dynamic and embodied. Further, these representations may be even more important in online instruction: whereas in-person instructors can, through their gestures, make static slides “come alive”, limitations of static slides are more difficult to overcome in the context of remote instruction. Fortunately, modern computer/web technologies put the creation of interventions of the sort we used within reach of most instructors, and make it possible to deliver these interventions outside of class, wherever students are.

A central question raised by our findings concerns the ”minimal” changes—in terms of dynamic visualizations and embodied representations—that instructors could make to their materials in order to improve student learning outcomes. For instance, some instructors introduce dynamic visualizations into their slides (without recording videos) by using animations and "step-by-step" reveals of information; The resulting instruction is perhaps somewhere between our screen-recording and PowerPoint conditions. In addition, instructors use gestures and pointing to draw attention to specific pointing of a slide, and may use embodied representations in non-manual modalities (e.g., repeating an important phrase / definition with a constant intonational / stress pattern). Future research could further the delineate ideal trade-offs between ease of implementation and student improvement.

References

Ainsworth, S. (2008). The educational value of multiple-representations when learning complex scientific concepts. in: Gilbert j.k., reiner m., nakhleh m. (eds) visualization: Theory and practice in science education. *Models and Modeling in Science Education*, *3*.

Airey, J., & Linder, C. (2009). A disciplinary discourse perspective on university science learning: Achieving fluency in a critical constellation of modes. *Journal of Research in Science Teaching*, *46*(1), 27-49.

Ayres, P., Marcus, N., Chan, C., & Qian, N. (2009). Learning hand manipulative tasks: When instructional animations are superior to equivalent static representations. *Computers in Human Behavior*, *25*, 348-353.

Batanero, C., Tauber, L. M., & Sánchez, V. (2004). Students' reasoning about the normal distribution. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 257–276). Dordrecht: Springer Netherlands. doi: 10.1007/1-4020-2278-6$_1$1

Castro-Alonso, J. C., Ayres, P., & Paas, F. (2014). Learning from observing hands in static and animated versions of non-manipulative tasks. *Learning and Instruction*, *34*, 11 - 21. doi: https://doi.org/10.1016/j.learninstruc.2014.07.005

Chandler, P. (2004). The crucial role of cognitive processes in the design of dynamic visualizations. *Learning and Instruction*, *14*, 353-357.

de Koning, B. B., & Tabbers, H. K. (2011). Facilitating understanding of movements in dynamic visualizations: an embodied perspective. *Educ Psychol Rev*, *23*(501–521). doi: https://doi.org/10.1007/s10648-011-9173-8

de Koning, B. B., & Tabbers, H. K. (2013). Gestures in instructional animations: A helping hand to understanding non-human movements? *Applied Cognitive Psychology*, *27*(5), 683-689. doi: https://doi.org/10.1002/acp.2937

delMas, R., Garfield, J., & Ooms, A. (2005). Using assessment items to study students' difficulty reading and interpreting graphical representations of distributions. In *Proceedings of the fourth international research forum on statistical reasoning, thinking and literacy.* Auckland, New Zealand: University of Auckland: Lawrence Erlbaum Associates.

Glenberg, A. M., Sato, M., Cattaneo, L., Riggio, L., Palumbo, D., & Buccino, G. (2008). Processing abstract language modulates motor system activity. *The Quarterly Journal of Experimental Psychology*, *61*(6), 905-919.

Hegarty, M. (2004). Dynamic visualizations and learning: getting to the difficult questions. *Learning and Instruction*, *14*(3), 343–351. doi: https://doi.org/10.1016/j.learninstruc.2004.06.007

Landin, J. (2011). Perceptual drawing as a learning tool in a college biology laboratory..

Lowe, R. (2004). Interrogation of a dynamic visualization during learning. *Learning and Instruction*, *14*(3), 257–274. doi: https://doi.org/10.1016/j.learninstruc.2004.06.003

Mayer, R. E. (2014). Principles based on social cues in multimedia learning: Personalization, voice, image, and embodiment principles. *The Cambridge handbook of multimedia learning*, *16*, 345-370.

McElhaney, K., Chang, H.-Y., Chiu, J., & Linn, M. (2014, 12). Evidence for effective uses of dynamic visualisations in science curriculum materials. *Studies in Science Education*, *51*, 49-85. doi: 10.1080/03057267.2014.984506

Rau, M. A. (2017). How do students learn to see concepts in visualizations? social learning mechanisms with physical and virtual representations. *Journal of Learning Analytics*, *4*(2), 240–263. doi: 10.18608/jla.2017.42.16

Rieber, L. P. (1991). Animation, incidental learning and continuing motivation. *Journal of Educational Psychology*, *83*, 318-328. doi: https://doi.org/10.1037/0022-0663.93.2.390

Son, J., 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. doi: http://dx.doi.org/10.1186/s41235-017-0085-0

Thomas, L. E., & Lleras, A. (2009). Swinging into thought: Directed movement guides insight in problem solving. *Psychonomic Bulletin Review*, *16*(719–723). doi: https://doi.org/10.3758/PBR.16.4.719

Tversky, B., Morrison, J. B., & Betrancourt, M. (2002). Animation: can it facilitate? *International Journal of Human-Computer Studies*, *57*(4), 247 - 262. doi: https://doi.org/10.1006/ijhc.2002.1017

Tversky, B., Morrison, J. B., & Bétrancourt, M. (2002). Animation: can it facilitate? *International Journal of Human-Computer Studies*, *57*(4), 247-262.

Wong, A., Marcus, N., Ayres, P., Smith, L., Cooper, G. A., Paas, F., & Sweller, J. (2009). Instructional animations can be superior to statics when learning human motor skills. *Computers in Human Behavior*, *25*(2), 339–347. doi: https://doi.org/10.1016/j.chb.2008.12.012

Wright, P., Milroy, R., & Lickorish, A. (1999). Static and animated graphics in learning from interactive texts. *European Journal of Psychology of Education*, *14*(2), 203-224.