RECOGNITION OF RANDOM PROCESSES FROM SIMULATED AUDITORY EXPERIENCES

Amy Renelle, Stephanie Budgett, and Rhys Jones
The University of Auckland, New Zealand
alin717@aucklanduni.ac.nz

Students frequently exhibit randomness misconceptions due to a multitude of reasons. The purpose of this paper is to explore whether auditory cues corresponding to a sequence of simulated events challenge students' intuitions of random processes. Results from this study indicate that randomness misconceptions were exhibited by the participants and, through running a simulation, their incorrect perceptions were then explored. From this, it is proposed that further research could investigate whether auditory cues are beneficial in challenging students' randomness intuitions in a classroom setting. The way in which the tool highlighted the students expectations of waiting times generated from a uniform distribution, and constant waiting times was unexpected outcome of this small pilot study and is yet to be fully explored.

INTRODUCTION

With increasing recognition of the importance of statistics across multiple disciplines, it has become clear that students require a more advanced statistical understanding of randomness than what is obtained through everyday observations (Gougis et al., 2017). Literature has documented randomness misconceptions seem to arise from disparities between these everyday observations and scientific teaching (e.g. Batanero, Arteaga, Serrano, & Ruiz, 2014; Gougis et al., 2017). To explore ways in which possible randomness misconceptions can be investigated, this study utilised technology to facilitate the use of multisensory elements and dynamic representations. With the aim of using technology and senses to help students connect concepts in new ways, this amalgamation between technology, multiple senses, and statistics is yet to be fully explored. Whilst the study also considered other research questions, this paper investigates whether or not the use of senses provides new avenues to explore and challenge students' misconceptions of randomness. Challenging student perceptions through auditory cues may provide a different way for connections between concepts to be made, helping students obtain a better understanding of essential statistical concepts.

BACKGROUND

Considering the prevalence of randomness misconceptions, Batenero et al. (2014) indicate that "[in] spite of being a basic idea in probability, randomness is not an easy concept. The term resists easy or precise definition…" (p. 347). Gougis et al. (2017), in discussing randomness misconceptions within science fields, emphasise the importance of students' understanding of randomness for clarity in a variety of different areas, including conducting investigations, describing observations, and interpreting results. However, Gougis et al. (2017) and Reimers, Donkin, and Le Pelley (2018) suggest that students often try to use an understanding of randomness obtained through their experiences of everyday events that cannot be used to adequately explain scientific or statistical examples. Students frequently appear to erroneously believe that random events or sequences must be unpredictable and have an irregular pattern, regardless of the number of observations (Batanero & Serrano, 1999). In their study, Reimers et al. (2018) also found that "…participants assume that the properties of long sequences of random outcomes—such as an equal proportion of heads and tails, and little internal structure— should also apply to short sequences" (p. 11). As randomness is a fundamental concept within statistics and other scientific fields, forming a basis for many other concepts including variation and independence, these misconceptions have further implications with students finding ideas related to randomness confusing (Batanero et al., 2014).

The use of technology in teaching statistics has many benefits, including facilitation of data investigation, consideration of relationships between variables and representations, as well as the ability to offer simulation experiences (Batanero, Chernoff, Engel, Lee, & Sánchez, 2016). Engel and Sedlmeier (2005) also state that simulation provides opportunities to challenge student misconceptions while Chance and Rossman (2006) indicate that "[technology], and simulation in particular, can be a very powerful tool in helping students learn statistics, particularly the ideas of long-run patterns and

randomness, in a concrete, interactive environment" (p. 1). Furthermore, digital technology allows for a combination of various senses to be utilised simultaneously, with the possibility of introducing relevant dynamic visualisations, corresponding sounds, connections visualised through colour, and participant interactivity (Holmes, 2009).

With advancing technology, it is becoming easier for students to explore concepts using a variety of senses. Shams and Seitz (2008) suggest that utilising multiple senses in the classroom is consistent with how we inherently learn and that information retrieval is enhanced when concepts are learnt through use of multiple senses. With the aid of digital technology, use of multiple senses within a classroom setting is more feasible and can be used to facilitate student learning (Shams & Seitz, 2008). Considering this from a statistical perspective, Mitchel and Weiss (2011) found that "...statistical learning may benefit from the availability of multisensory input" (p. 15) and suggest that "...learners are able to extract multiple statistical regularities simultaneously from audio-visual input" (p. 9). However, Cognitive Load Theory, where excess information overwhelms students at a detriment to their learning (Paas, Renkl, & Sweller, 2003), may highlight a negative impact from introducing a variety of senses. This paper will focus predominantly on observations regarding auditory cues produced from simulated random sequences and, as a possible limitation of using multisensory learning, Cognitive Load Theory will be considered in relation to the use of simulated sound sequences.

THE STUDY

The Main Study: Purpose and Method

The purpose of this exploratory study was to investigate the use of different representations of randomness and to determine if these experiences could be used to facilitate students' communication of their randomness perceptions. The small exploratory study comprised two first-year university students, who carried out tasks designed to explore their perceptions of randomness. The participants had previously completed a first-year probability course and were expected to have encountered similar ideas before. Using a 'think-aloud' protocol, qualitative data was generated throughout the task as the researchers queried the participants' actions and reasoning. The complete task was designed to aid participants in the articulation of their perceptions of randomness at different stages of investigation through use of a digital tool. The participants were invited to use this tool and follow through a storyboard progression that guided the investigation of several datasets in different modes.

The Scampy Tool: Features and Simulation Mode

The *Scampy Tool* was based on an original prototype tool developed at the University of Auckland by the Statistics Department to help students visualize randomness, and to make connections between the Poisson and exponential distributions (Budgett & Pfannkuch, 2018). This pilot study is the first to investigate student perspectives using this tool – further research using the tool is currently being conducted. The tool is comprised of three components: Data Mode, Video Mode, and Simulation Mode. Simulation Mode (*Figure 1*) will form the basis of the task concerned in this paper and, whilst interesting to consider, the other modes of the tool are beyond the scope of this paper. Key features of the tool include dynamically linked representations, participant interaction, colours, and sound effects. *Figure 1* shows data being generated in Simulation Mode. The timeline along the bottom of the image represents the times at which cars pass a particular point, denoted by the person-icon. The user specifies the distribution giving rise to the random observations that represent the time that is observed between cars passing this point, defined as the waiting time. In the example shown in *Figure 1*, the times between cars passing the person-icon are being generated by the uniform distribution with an average of two cars per second and the simulation will finish once 100 observations have been generated.
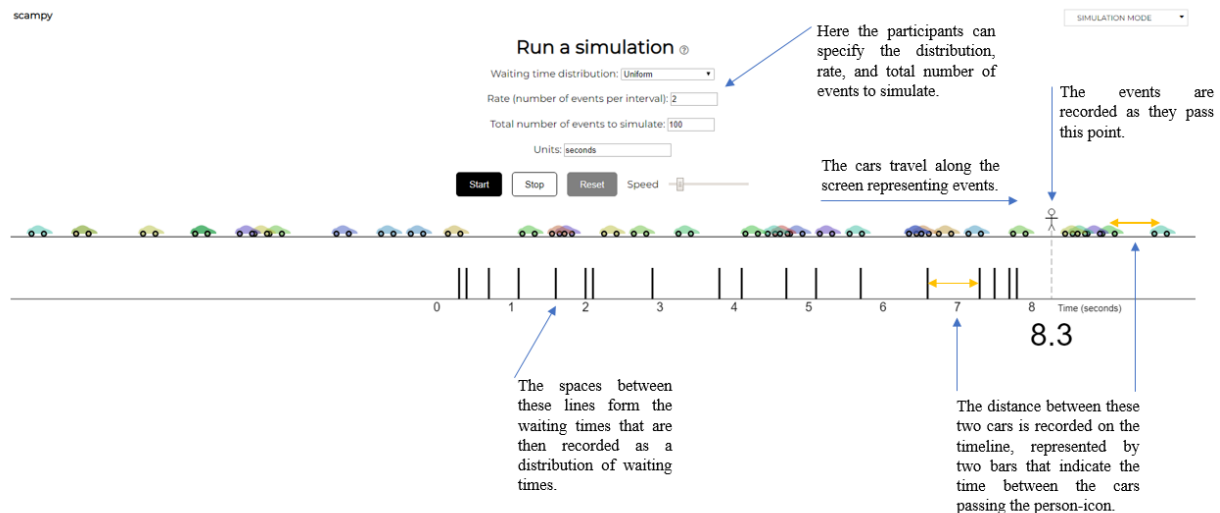
Figure 1. One annotated frame from running a simulation of waiting times generated by the uniform distribution using the *Scampy Tool*.

Auditory Cues and the Task

The use of Simulation Mode (*Figure 1*) allows the participants to select a distribution and specify its parameters before watching a dynamic visualisation of the waiting time data being generated, accompanied by sounds and colour-coordinated connections. This approach was designed to provide a platform for investigating participant perspectives of randomness through auditory cues. Together with the visual cues indicated in *Figure 1*, such as the spacing between the cars and the distance between data points on the timeline, an additional possible avenue for participants to gauge a sense of the data in a more intuitive way was provided through sound effects that can be heard during the generation of the waiting time data. As each simulation unfolded, conversations between the researchers and participants were recorded and later analysed, with the aim of delineating possible participant perspectives of the tool and of the randomness experienced during the simulation. The participants were asked by the researchers to think about the uniform distribution as a means of generating the waiting time between cars passing the person-icon. During the simulation, the *Scampy Tool* produced auditory cues in accordance with the generation of these waiting times. The participants were then asked to consider the pattern of sound produced by the simulated sequence and comment on how this compared to their expectations of a simulation of the uniform distribution. Following this, they were asked to repeat the task, but instead generating constant waiting times. The two simulated sound patterns were then compared. Please see the following link for a recording of the tool being used, and see Appendix A for details of the step-by-step process the participants undertook: https://tinyurl.com/y2zdupzv.

DISCUSSION

Throughout the session, the participants exhibited common randomness misconceptions. When asked to define and explain randomness, the participants suggested that it was unpredictable and related to a change in probability between what was anticipated and what actually occurred. It was unclear where the notion of changing probabilities originated. As a commonly accepted example of randomness, a fair coin will randomly produce an outcome of heads or tails without the underlying probability of the event of heads or tails changing. This perception of changing probabilities, along with the perception of a lack of randomness in distributions that arose later in the session, indicated the presence of randomness misconceptions. The misconception regarding a lack of randomness in distributions is the focus of this paper, and while other misconceptions were evidenced, these will not be discussed here.

Simulation Mode provided an opportunity for some of these misconceptions to be explored. During the session, participants were asked to conjecture what the distribution of waiting times may look like if they were generated by the uniform distribution. The following is an excerpt from conversations with the participants (P1 and P2) as they articulated their expectations of waiting times generated by the uniform distribution before running the simulation.

> *P2: Uniform is… It's like the same number of events per interval every time, I think… Not random.*
> *P1: The waiting time is the same.*

The suggestion of a lack of randomness present in the uniform distribution is indicative of randomness misconceptions. It appears that the participants have taken what they expected the uniform distribution to look like and determined that any random observation, in this case referring to the cars as events (rather than the waiting times observed), generated from a uniform distribution is equally likely meaning that randomness could not be present. This connected with the participants' definition of randomness relating to changing probabilities. Following the above discussion regarding their perception of the uniform distribution, the participants were invited to observe and listen to the simulation using the *Scampy Tool*. The pattern of sound produced by the simulated sequence was heard during this time and this evoked a conversation regarding the uniform distribution. Enquiring about the sequence produced from the data being recorded, the researcher (Res) prompted the participants to consider the intuitiveness of these sounds relative to the underlying distribution giving rise to the data. This question was unscripted and the response was unexpected:

> *Res: Just wondering about the sounds, with the beeps, did it sound uniform?*
> *P1: Nope.*
> *Res: Does that help with your intuition as to why that would be uniform?*
> *P1: Yeah, it helped with intuition, it didn't sound uniform.*
> *Res: It didn't sound uniform?*
> *P1: It was obvious that it wasn't.*
> *Res: Why, why are you saying that it didn't sound uniform?*
> *P1: Oh because if it's uniform it's kind of obvious, it's like a beat of a metronome, 1, 2, 3, 4.*

The participants had previously expressed the expectation that the uniform distribution would not exhibit randomness, and hence anticipated a constant "metronome" sequence of sound. We conjecture that the participants determined that variation in the sequence of sound would correspond to the presence of randomness. Since they suggested that waiting time observations generated from the uniform distribution would not exhibit randomness, the resulting sequence of sound produced from the simulation must have invariable timing. The participants' initial expectations were challenged because the simulation produced a non-constant pattern of sound. To demonstrate with a different example, the participants repeated the process generating constant waiting times. In doing so, sounds with a regular beat were heard and the participants were then able to see that their expectations of the uniform distribution were misplaced.

To take a sequence of observations displayed as numbers and deliberate what this may sound like over a timeline would be an unfamiliar task to most students. A search through the literature failed to uncover articles considering the implications of this, particularly within a statistical setting. It would appear that attending to the sounds produced by a sequence of events corresponding to observations generated from a specific distribution is not common classroom practice, and many students may not have had the opportunity to notice patterns in this way before. Whether it is reasonable to expect the participants to correctly articulate the sound sequence of a uniform distribution when this kind of reflection is generally unfamiliar to students is unclear. As a new avenue to explore, we are not yet sure what response from the students would have been deemed appropriate. Whilst randomness forms part of a student's understanding of simulating from distributions, it may be that the response was due to an incomplete conception of probability distributions. However, this small exploratory study found that auditory cues may be beneficial in challenging students' intuitions of randomness. The participants were unlikely to have thought about the simulated sequence in this way without being prompted. Further research is therefore needed to establish whether, despite being a typically non-considered aspect, the use of sounds assists in learning or whether it causes an excess of information. Further research may then consider whether the use of sounds provides an intuitive way of noticing randomness and whether this would be beneficial in a classroom setting.

Literature on the use of multiple senses frequently deliberates the implications of Cognitive Load Theory (Chance & Rossman, 2006). Cognitive Load Theory relates to situations where excess information overwhelms students, negatively impacting on their learning ability (Paas et al., 2003). Although research conducted by Shams and Seitz (2008) found that cognitive load can be reduced by congruent multisensory inputs, this depends on the multisensory elements contributing positively to a student's experience and requires that these additional sources of information assist in providing clarity for students. There was no evidence of cognitive overload being an issue in this study. The participants did not appear to exhibit signs of confusion or over-stimulation during the task. It seems that the amalgamation of dynamic representations, participant interaction, colours, and sound effects created a coherent picture of the situation concerned, with this congruency perhaps reducing the prevalence of cognitive overload. While findings from this very small study suggest that cognitive overload may not be an issue, further research is needed to confirm this with a wider variety of participants.

CONCLUSION

The information collected in this study is not conclusive, nor are the tentative results representative of all potential users of the tool. However, the perceptions of these two participants may be used to inform further research. Whether auditory cues can be used to challenge student misconceptions of randomness and if classroom implementation is both possible and advantageous is yet to be established. Consideration of how such an approach may work in a classroom setting may lead to the development of a task so as to facilitate a larger number of students and still promote the asking of questions around how randomness may be observed. Continuation of research aims to improve the external validity through increasing the number of participants and introducing variation in the educational background of participants. The two participants in this study appeared to exhibit randomness misconceptions, some of which were successfully challenged through the auditory cues produced in Simulation Mode of the *Scampy Tool*. Further research into using this tool to help students make connections between different representations of distributions is to commence this year.

REFERENCES

Batanero, C., Arteaga, P., Serrano, L., & Ruiz, B. (2014). Prospective primary school teachers' perception of randomness. In E. Chernoff & B. Sriraman (Eds.), *Probabilistic thinking: Presenting plural perspectives* (pp. 345–366). Dordrecht, The Netherlands: Springer.

Batanero, C., Chernoff, E. J., Engel, J., Lee, H. S., & Sánchez, E. (2016). Research on teaching and learning probability. In *Research on teaching and learning probability* (pp. 1–33). Cham, Switzerland: Springer.

Batanero, C., & Serrano, L. (1999). The meaning of randomness for secondary school students. *Journal for Research in Mathematics Education, 30*, 558–567.

Budgett, S., & Pfannkuch, M. (2018). Modelling and linking the Poisson and exponential distributions. *ZDM Mathematics Education*, *50*(7), 1281–1294.

Chance, B. L., & Rossman, A. J. (2006). *Investigating statistical concepts, applications, and methods*. Belmont, CA: Cengage Learning.

Engel, J., & Sedlmeier, P. (2005). On middle-school students' comprehension of randomness and chance variability in data. *ZDM Mathematics Education*, *37*(3), 168–177.

Gougis, R. D., Stomberg, J. F., O'Hare, A. T., O'Reilly, C. M., Bader, N. E., Meixner, T., & Carey, C. C. (2017). Post-secondary science students' explanations of randomness and variation and implications for science learning. *International Journal of Science and Mathematics Education*, *15*(6), 1039–1056.

Holmes, K. (2009). Planning to teach with digital tools: Introducing the interactive whiteboard to pre-service secondary mathematics teachers. *Australasian Journal of Educational Technology*, *25*(3). doi:10.14742/ajet.1139

Mitchel, A. D., & Weiss, D. J. (2011). Learning across senses: Cross-modal effects in multisensory statistical learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(5), 1081.

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist*, *38*(1), 1–4.

Reimers, S., Donkin, C., & Le Pelley, M. E. (2018). Perceptions of randomness in binary sequences: Normative, heuristic, or both? *Cognition*, *172*, 11–25.

Shams, L., & Seitz, A. R. (2008). Benefits of multisensory learning. *Trends in Cognitive Sciences*, *12*(11), 411–417.

APPENDIX A

The following is an outline of the process the participants undertook in completing the simulation task, refer to Figure 1 or the link below for a visual reference:

1. The participants were asked to articulate their expectations of what the waiting times generated by the uniform distribution might look like before running the simulation. This conjecture regarding their expectation of the waiting times sparked the later auditory-based questions.

2. Participants selected the uniform distribution from the drop-down menu. Other distributions are available but for this part of the task the participants were requested to select the uniform distribution. This is the underlying distribution from which the tool will simulate *waiting times*. The participants then entered the number of objects per interval (rate of cars) and the total number of observations to simulate (n waiting times) (see Figure 1 or the link below).

3. The participants then click "start" and the simulation begins. The simulation generates a sequence of cars travelling along a timeline with the waiting time, or time between consecutive cars, determined by the specified distribution (see Figure 1 or the link below). That is, the simulation mimics traffic passing a point with a specified underlying waiting time distribution, in this case the uniform distribution. As the cars pass the person-icon, a sound is heard.

4. Following the simulation, the participants were asked to reflect on the pattern of sounds generated. The participants were asked if the sounds they heard were uniformly distributed. The participants were expecting equally spaced sounds and could hear that this was not the case.

5. Once the simulation has finished, the participants move on to look at the distribution of waiting times. Here, the waiting times generated by the simulation are recorded on a graph below the timeline (as seen in the link below), building up a dot plot of waiting times. For the uniform distribution simulation, we expect the shape of the dot plot to be approximately uniform when a large number of events (waiting times) have been simulated. Disparities from this indicate the presence of randomness.

6. The participants then move on to look at the distribution of counts, that is, the number of objects (cars) observed within a time interval. A time interval is specified, with the timeline then split into segments of this specified length. Counts are created by counting how many cars occurred during each time interval. The number of cars in each segment is then recorded on a bar graph (as seen in the link below).

7. Lastly, the summary page presents the distribution of waiting times and the distribution of counts next to each other and numerical values can be viewed. Participants can click to see the mean waiting time as well as the theoretical information next to the simulated information (such as means and counts, etc.).

See this in action at: https://tinyurl.com/y2zdupzv