RESEARCH PAPER

Exploring the Potential for Student Development of the Big Ideas of Statistics with Random Trials: The Case of the Mystery Spinner

Jane Watson
University of Tasmania
jane.watson@utas.edu.au

Noleine Fitzallen
University of Tasmania
noleine.fitzallen@utas.edu.au

This paper reports on the potential for engaging students in an activity that considers the interconnectedness of the five Big Ideas of Statistics in the context of conducting random trials. In the study, two classes of Year 6 students (aged 11-12 years) used *TinkerPlots* to determine the sample space of a "Mystery Spinner." Analysed for this paper were data collected from entries made in completed workbooks while engaged in the learning activity and responses to relevant questions in an end-of-year questionnaire from 27 students. The results indicated using an activity that included a probability model contributed to students working mathematically with percentages and frequencies and supported the development of intuitions about randomness and informal inference. This was promoted by students analysing the variation in the distribution of data and describing their expectations about an unknown sample space.

It has been common for primary students to use hands-on spinners when investigating probability in the primary years of schooling (e.g., Chick & Baker, 2005; Torok, 2000). In most situations the frequency and relative frequency of events occurring, the confirmation of the outcomes according to the known sample space (actually seen on the spinner), and fairness/randomness of the devices are the focal points of such investigations. These investigations tend to emphasise that many trials will ensure the experimental results reflect the known sample space (theoretical expectations). In this case, variation and expectation are fundamental concepts associated with randomness. Developed from early experiences of probability that usually involve tossing dice or coins, the students appreciate that the experimental results can *vary* from the theoretical *expectations* when only a few trials have been undertaken but tend to be confident that conducting more trials will ensure "success" in getting the outcomes expected, with little acknowledgement of the variation that can still occur. There are benefits of using hands-on materials to develop and extend mathematical knowledge and problem-solving strategies (Chick, 2018) but the scope for artefacts such as dice, coins and spinners is limited. Traditionally, success in these situations is measured by getting what would be considered "the correct answer." Taking the traditional approach in the classroom constrains students' opportunities to engage in the productive struggle required to develop higher order thinking, such as critical thinking and reasoning skills.

Probability and statistics are often taught as two discrete subjects within the mathematics curriculum. This does not assist in connecting the areas of probability and statistics, which is required when questioning data and making inferences from statistical information (Batanero et al., 2016; Watson et al., 2018). With the goal of probability activities to compare experimental results to theoretical outcomes determined by the known sample space (Lee et al., 2010), the opportunity of developing understanding of probability concepts encountered in statistics later in students' studies and future working situations (e.g., *p*-values) is scant.

Probability activities in the early years of schooling are essentially devoid of the exploratory nature of statistical enquiries that meet goals of student development of higher-order thinking. There is, however, scope to deepen students' learning about probability concepts by introducing elements of cognitive conflict into learning opportunities. Cognitive conflict is the disagreement between cognitive structures, such as knowledge and mental representations, and experience (Waxer & Morton, 2012). The challenge is to develop activities that build on established knowledge and take advantage of what students would consider familiar in order to promote engagement in deeper and more complex problem solving and

2022. N. Fitzallen, C. Murphy, V. Hatisaru, & N. Maher (Eds.), *Mathematical confluences and journeys* (Proceedings of the 44th Annual Conference of the Mathematics Education Research Group of Australasia, July 3–7), pp. 530–537. Launceston: MERGA.

reasoning. Learning activities need to be not only cognitively accessible for students but also cognitively challenging to ensure opportunities for growth in learning are maximised. Therefore, it is important to develop learning sequences that establish prior knowledge, consolidate learning, and extend that initial learning in anticipation of targeting long-term goals of critical thinking. Important research problems in this regard are:

> (a) clarifying the way in which probabilistic thinking could contribute to improving mathematical competencies of students, (b) analysing how different probability models and their applications can be presented to the students, (c) finding ways in which it is possible to engage students in questions related to how to obtain knowledge from data and why a probability model is suitable, and (d) how to help students develop valid intuitions in this field. (Batanero et al., 2016, p. 25)

Hence, the research question for the study reported in this paper is:

In terms of the Big Ideas of Statistics, what is the learning potential of an activity that requires students to use randomly generated data to make conjectures (hypotheses) about the sample space of a hidden spinner?

Related Literature

Five Big Ideas underpin statistics education at school and beyond (Watson et al., 2018). In relation to probability, carrying out *Random* trials, students experience *Variation*, which is viewed through a summary *Distribution*, resulting in an *Expectation* about the underlying sample space. As trials increase, more *Variation* occurs and *Expectation* about the ultimate outcome is likely to change. The goal is to make an *Informal Inference* about the underlying probability of the elements of the sample space. The interconnectedness of the five underpinning concepts is shown in Figure 1.

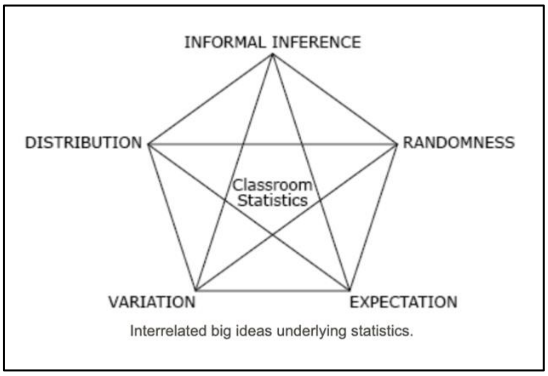

Figure 1. The Big Ideas of Statistics (Watson et al., 2018, p. 126).

Randomness and informal inference are the more complex of the Big Ideas, as each requires two components for understanding. For randomness, there is more than the individual outcome of being uncertain; there also needs to be a progression towards a fixed outcome with many trials (Lee et al., 2010). For informal inference, there needs to be both the expression of an expectation and a level of confidence expressed in the expectation because the result is not known for certain (Watson et al., 2018). Classroom experiences demonstrate that “random” is not an easy principle for students to understand fully (Chick, 2018; Lee et al., 2010; Watson & Fitzallen, 2019). Part of this may be associated with the activities employed. Chick discussed the pros and cons of many of the main activities that were used in linking probability and randomness, including one based on two spinners, and pointed out the importance of recognising the learning opportunities afforded by activities used to foster intuitions about randomness. This imperative concurred with suggestions made by Batanero et al. (2016) and Lee et al. (2010).

Probability, as taught in school, is basically the endpoint of random phenomena, based on the "sample spaces" that are observable on the sides of a die or coin that is tossed. In this case, activities do not produce challenges for two of the big ideas: the expectation is *seen* on the device and there is no informal inference to be made because the *answer* is known at the beginning. Specific approaches to the learning and teaching of probability are needed to enhance the learning outcomes of traditional probability activities (Batanero et al., 2016; Chick, 2018; Chick & Baker, 2005). As suggested in the *Australian: Mathematics Curriculum* (Australian Curriculum, Assessment and Reporting Authority [ACARA], 2021), frequency and sample space tend to focus on the Proficiency Strands of understanding and fluency. It is interesting that although 'random number' and 'random sample' were defined in the Australian Curriculum Glossary for Mathematics (ACARA, 2021), the term 'random' itself was not defined. Random is associated directly with the relationship between expectation and variation.

Research Approach

The methodology employed was exploratory, which warranted using qualitative research and data analysis strategies to look for possibilities and opportunities in a pragmatic fashion (Mackenzie & Knipe, 2006). The learning sequence in the classroom included establishing prior knowledge, consolidating foundational learning, recognising issues arising, and carrying out the main investigation of random trials with a mystery spinner to extend learning. Data were collected from student workbooks and, as part of measuring adoption of the intended outcomes, responses to questions in the end-of-year questionnaire.

Context of the Study

As part of a classroom intervention with Year 6 students (aged 11-12 years) near the end of a 4-year project (Fitzallen & Watson, 2020), 56 students in two classes took part in two extended lessons, taught by the classroom teacher, addressing the Chance curriculum for Year 6 (ACARA, 2021):

> Conduct chance experiments with both small and large numbers of trials using appropriate digital technologies (ACMSP145)
>
> Compare observed frequencies across experiments with expected frequencies (ACMSP146)

The students had been involved in data handling activities to answer statistical questions, but the activity reported in this paper was their first encounter as part of the project with the Chance part of the curriculum. The activity was hence planned to establish students' previous knowledge and explore an extension encompassing an unknown sample space, allowing for an informal inference to be made (Batanero et al., 2016). The activity reported in this paper was designed to allow for the focus to be on the random nature of trials, the variation encountered, the expectation created, the distribution of results to assist with the expectation, and the inference made about the underlying sample space being examined with spinners (cf. Figure 1). Although advances in technology have made it possible for students to complete many trials of a spinner (e.g., Chick, 2018; Lee et al., 2010), often these activities still display the final goal of the trials as trials are taking place. The *TinkerPlots* software (Konold & Miller, 2015), however, offered the option of a Mystery Spinner, where the proportions of the spinner were not visible as trials were performed. The purpose of the blank spinner was to introduce cognitive conflict into the learning experience to provide a situation where students were required to determine the sample space from the data generated. As Batanero et al. (2016) suggested,

> These "black-box" types of simulations may assist students in thinking about probability from a subjective or frequentist perspective where they can only use data generated from a simulation to make estimates of probabilities that they can use in inference or decision-making situations. (p. 19)

Learning Sequence

Establishing prior knowledge in terms of Expectation. The goal was to establish that students could relate to the term expectation and appreciate what it meant in relation to their learning and real-life contexts. Could they make reasonable predictions and be creative and imaginative in describing what would be considered realistic expectations? The lessons began with a review of likelihood in terms of expectation, with students in groups of four, writing down on an A3 sheet of paper divided into quarters, and discussing:

1. Something I expect will **definitely** happen this weekend …
2. Something I expect **might** happen this weekend …
3. Something I expect **may** happen this weekend but I do not think it is very likely …
4. Something I definitely **do not** expect to happen this weekend …

Figure 2 illustrates the typical responses of students in these categories. The responses established that the students could relate to the term expectation and could appreciate what it meant in relation to their school and real-life contexts. The students made reasonable predictions and were creative and imaginative in describing what would be considered impossible expectations. A few of the apparently less appropriate responses were discussed within the groups and with the class. This was then combined with a review of the general language of likelihood with a clothesline labelled from 0 to 1. Students drew words/phrases out of a bag and attached them to the clothesline with discussion about their placement in terms of the probability represented.

1. The sun will come up … Won't go to school … Humans will exist … Baby will be born (every 19 secs) … Eat food … Speak to my family.	2. Someone might watch a Marvel movie … People will be on their phones … Going to church … Homework … Go rock climbing.
3. I will clean my room … WW3 … MacDonald's will shut down … Donald Trump gets killed … Go to the park … Might have a 4-day weekend.	4. The sun to fall out of the sky… The moon will rain water … Go to Mars … Meet an alien … To play tag around the park … Go to school.

Figure 2. Examples of students' expressions of expectation.

Consolidating foundational learning. The initial data collection activity included using conventional spinners. The aim was to establish mathematical expectations in terms of the outcomes of spinners according to the sample space with the benefit of many, many trials. After the teacher reviewed the expression of probabilities as fractions and did trials of a 50-50 black and white spinner, students conducted trials with hands-on spinners as seen in Figure 3, recognising $Pr(\text{red}) = Pr(\text{yellow}) = Pr(\text{green}) = Pr(\text{blue}) = ¼$. A four-colour spinner was chosen as a hands-on beginning of the data collection because of the possible extensions using the technology, *TinkerPlots*, to extend the number of trails as implied in the curriculum statement.

The students worked in groups of four, individually conducting 20 spins of the spinner, recording the results in their workbooks, and commenting on how close they were to each other and to the probabilities built into the spinner. They then combined their results in pairs, reporting whether their results were closer to the probabilities than before or further away. Finally, they combined results for the four students in the group, again commenting on their results and the expected probability for 80 spins. Figure 3 shows the outcomes for the four-part spinner from two students in different groups. Obvious from this, and many cases, the results of the approximations for some groups got worse over the trials but most got closer. For some students it took until the data from the whole class were combined before they were convinced that the desired percentages for the probabilities were approached.

Number of trials	Red		Yellow		Green		Blue		Comment
	n	%	n	%	n	%	n	%	
20	6	30%	4	20%	5	25%	5	25%	Very close
40	16	40%	8	20%	8	20%	8	20%	No (not closer). Much variation
80	31	38.75%	20	25%	16	20%	13	16.25%	More variation

Number of trials	Red		Yellow		Green		Blue		Comment
	n	%	n	%	n	%	n	%	
20	9	45%	4	20%	1	5%	6	30%	Spinner doesn't like me.
40	14	35%	11	27.5%	5	12.5%	10	25%	Closer
80	23	28.75%	19	23.75%	15	18.75%	23	28.75%	More closer

Figure 3. Results from four-part spinner trials.

Recognising issues arising. There are two issues, however, that arise here: after an introduction to these devices, assuming they are fair, (1) students *know* the answer they are trying to approach: it is just a matter of how long it takes to get there! Also, (2) it can get tedious doing many trials by hand and mistakes can creep in using a calculator to make calculations.

For Issue (2), the technology employed was *TinkerPlots* (Konold & Miller, 2015) to save the tedium of performing hundreds of trials. Students had used *TinkerPlots* in other activities as part of the project (e.g., Watson et al., 2022) and were keen to use the program for performing the spins of the spinner. *TinkerPlots* had a "Sampler", which could repeat the trials (pseudo-randomly) for a model of the spinner, producing increasingly large numbers of trials quickly and displaying the results in a table or plot. Once using the technology, it was not necessary for the spinner to be separated into four equal parts. The classroom teacher then did a demonstration to show the possibility of using other divisions of the spinner, as seen in Figure 4. As well *TinkerPlots* displayed the distribution of the results in a format different from the tables that the students had used in their workbooks with the hands-on spinners (see Figure 3). The teacher then increased the number of trials, eventually reaching several hundred, with discussion of how close the percentages in the Results of Sampler 1 distribution were getting to the values that could be seen on the spinner.

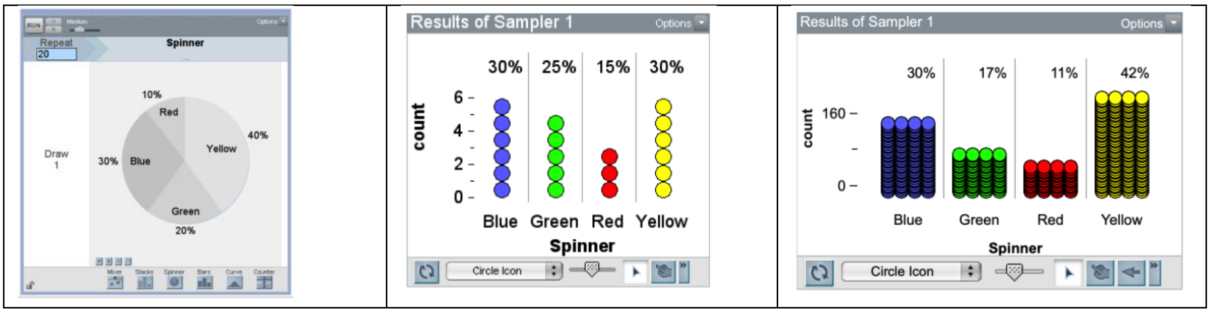

Figure 4. Teacher demonstration with 20 spins and then with 500 spins.

Extending learning. In relation to Issue (1) of knowing the makeup (sample space) of the spinner at the beginning, a blank spinner was introduced. *TinkerPlots* had the option of being able to hide the proportions of each colour. The aim was to introduce cognitive conflict as variation and expectation were juxtaposed in inferring the construction of the spinner. Students were given a "Mystery Spinner" with unequal sections, with the goal to work out the proportion of each colour. Six different Mystery Spinners (appearing in *TinkerPlots* as Figure 5, left) were prepared for the students to explore individually. The spinners were divided into three segments, labelled Apple, Banana, and Cherry and the Sampler was labelled "Fruit_Spinner".

In their workbooks, students started by recording the outcomes for 20 spins, then drawing a spinner with their estimates of the percentage of each fruit as their expectations of its construction (Figure 5, centre). They were then asked to mark their level of confidence on the arrow (Figure 5, right) and explain their reasoning. The number of trials was then increased to 50, 100, and 500, with similar data recorded and diagrams drawn. Student also answered the question, “Has your prediction or level of confidence changed? Why/why not?” This activity exhibited the five Big Ideas of Statistics (Figure 1). It was a (pseudo)*random* process taking place with the technology; there was *variation* as the trials progressed; after each trial students expressed their *expectation*s by sketching in the divisions of the spinner in Figure 5; this was based on the *distribution* of *TinkerPlots* results similar to that seen in Figure 4; and each time an *informal inference* (prediction) was made, the students marked their level of confidence on the accompanying arrow image. The level of confidence was an indicator of the degree the students expected the results at each set of spins reflected what they expected to be the makeup of the actual spinner. The students’ confidence for one set of spins was influenced by how much the result varied from the previous set.

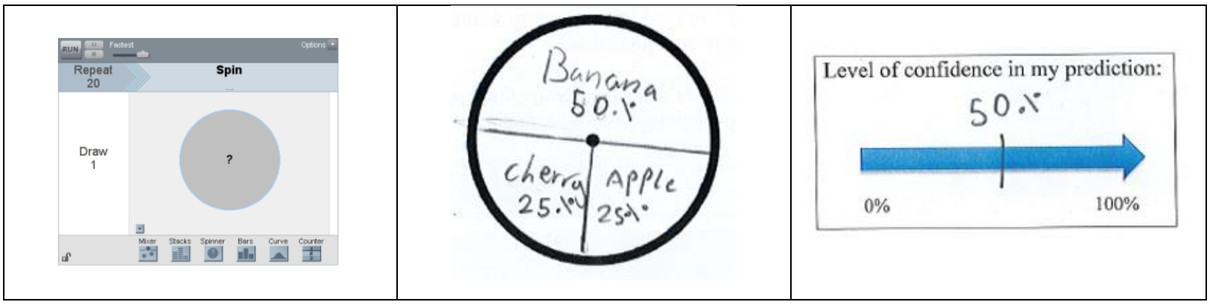

Figure 5. Trials with an unknown spinner.

Table 1.
Student confidence levels and comments

ID	Confidence: 50 spins	Confidence: 100 spins	Confidence: 500 spins
ID143	MC: There’s more spins.	LC: Because they are changing a lot. So I’m not as confident.	MC: I’m confident because the spinners in the last 2 examples have been very close.
ID129	LC: Yes and no, I’m not very sure.	MC: Yes, because I have really thought about them and I’m a bit more confident.	LC: Cherry has done better. We are confident that Cherry is really low but we are not confident what % to give it!
ID122	NC: It’s gone down a bit but my results were quite similar still.	NC: It hasn’t changed because I think that my prediction is correct or very close because the results are all similar.	NC: It still hasn’t changed because I think I’m close.
ID175	NC: Because I think there is more fruit [only two different fruits appeared on the first 20 spins]	MC: Yes, because I think there might only be 3 fruits.	MC: Yes, I think there is definitely only 3 fruits.

The results reported here are from a subset of the workbook data. The data were included if students had completed and reported all spins and written responses (n = 27). Based on their marked confidence arrows and/or their comments, after 50 spins eight students had no change

(NC) in confidence in their expected result from the first 20 spins, seven had less confidence (LC), and 12 had more confidence (MC). After 100 spins there were four with no change in confidence, 18 with more confidence, and five with less confidence than before. After 500 spins results were similar with three students indicating no change in confidence, 19 more confident, and five less confident. The reasons for these judgments varied across the trials. Typical responses are presented in Table 1.

Students could then choose as many trials as they liked and record their results making a final prediction of the percentage of each fruit, answering the question, “What makes you confident that your final prediction is correct? Not many students took up this option but all who did recognised that the often-huge number of trials produced the required percentages.

Measuring adoption of intended outcomes. On the end-of-year questionnaire, approximately eight weeks later, students were presented with the distribution of outcomes for two Mystery Spinners of the type they had investigated in the classroom, one from 30 spins and the other from 600 spins (Figure 6), asked to make a conjecture—hypothesis about the theoretical probability—about the sizes of the three parts of the spinner each time, and to explain why they made the decision each time.

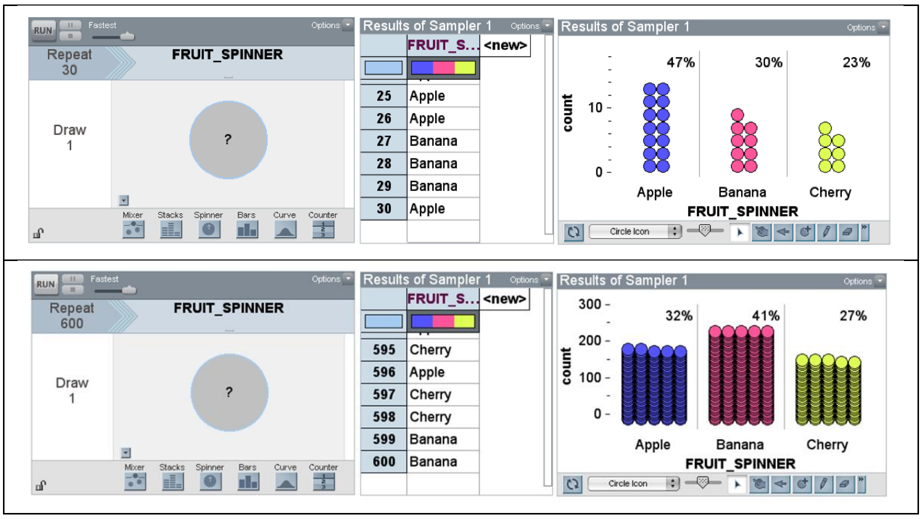

Figure 6. Distribution of data for a mystery spinner used in the end-of-year questionnaire.

For the 30 spins trial, 30% of responses reflected the type of rounding to multiples of 5% or 10%, with 20% of reasons being less precise rounding or “adding to 100%”, and 52% saying “that’s what the plot tells me.” Similar results occurred for 600 spins, with 20% of responses reflecting appropriate approximations and 46% remaining exact values from the figure or suggesting more extreme values. When then asked in which results they had more confidence, 61% said 600 spins and 39% said 30 spins. For the reasons for the choice of the number of spins, 43% specifically noted “more data”, “more spins” or “more chance” for choosing the 600 spins. The other 57% either chose 30, with reasons like, “cause it’s quicker” or “because it’s hardly any spins”, or chose 600, with nebulous reasons like, “I just feel more confident”. Hence, although 61% of students chose the appropriate number of spins only 70% of them also justified the choice based on the sample size.

Discussion and Conclusion

The learning potential embedded within the Mystery Spinner activity was associated with all five Big Ideas of Statistics (Figure 1). This allowed for students to express expectations during trialling and to revise them as variation was seen in successive random trials, which increased in size progressively throughout the activity. An informal inference was then made, accompanied by a degree of confidence. In this case the confidence level expressed was an indicator of student expectation. The opportunity to shift learning to a technological environment demonstrated the potential of using technology to expose students to extended learning opportunities not possible within the constraints of using physical manipulates (Batanero et al., 2016; Lee et al., 2010). The Mystery Spinner activity provided a rich learning opportunity that included outcomes in terms of frequency and percentage through attempting to balance variation and expectation to make judgments about the sample space of the mystery spinner. From a research perspective, this study illustrated that the students were able to make connections between the relative frequency of data and the theoretical probability, but some students only reiterated the raw data—expressed as counts or percentages—when making conjectures about the makeup of the spinner. This suggested that the learning potential of the activity was not totally realised for these students, who, in many cases, had unstable ideas about making inferences from data and did not yet appreciate that the theoretical model embedded in their mystery spinner could potentially vary from the results of the trials. More research is needed that focuses on student understanding of the relationships among the Big Ideas of Statistics to inform effective learning and teaching of statistical concepts, thereby addressing outcomes for the Proficiency Strands: problem solving and reasoning (ACARA, 2021).

Acknowledgements. This study was funded as part of an Australian Research Council (ARC) Discovery Grant (DP150100120). Jane and Noleine acknowledge the assistance of Helen Chick, Ben Kelly, and Suzie Wright in the implementation of the activity.

References

Australian Curriculum, Assessment and Reporting Authority [ACARA]. (2021). *Australian Curriculum: Mathematics (Version 8.4).* Sydney. https://www.australiancurriculum.edu.au/f-10-curriculum/mathematics/

Batanero, C., Chernoff, E. J., Engel, J., Lee, H. S., & Sánchez, E. (2016). *Research on teaching and learning probability.* Springer. https://doi.org/10.1007/978-3-319-31625-3_1

Chick, H. (2018). Understanding randomness: From gimmicks to effective learning. In M. A. Sorto & A. White, & L. Guyot (Eds.), *Looking back, looking forward* (Proceedings of the 10th International Conference on the Teaching of Statistics, Kyoto, Japan, July 8–14, 2018). International Statistical Institute.

Chick, H. L., & Baker, M. (2005). Teaching elementary probability: Not leaving it to chance. In P. C. Clarkson, A. Downton, D. Gronn, M. Horne, A. McDonough, R. Pierce, & A. Roche (Eds.), *Building connections: Theory, research and practice.* (Proceedings of the 28th annual conference of the Mathematics Education Research Group of Australasia, pp. 233–240). MERGA.

Fitzallen, N., & Watson, J. (2020). Using the practice of statistics to design students' experiences in STEM. In B. Shelley, K. te Riele, N. Brown, & T. Crellin (Eds.), *Harnessing the transformative power of education* (pp. 74–99). Koninklijke Brill NV. https://doi.org/10.1163/9789004417311_007

Konold, C., & Miller, C. D. (2015). *TinkerPlots: Dynamic data exploration* [Version 2.3.2]. Learn Troop.

Lee, H. S., Angotti, R. L., & Tarr, J. E. (2010). Making comparisons between observed data and expected outcomes: Students' informal hypothesis testing with probability simulation tools. *Statistics Education Research Journal, 9*(1), 68–96. https://doi.org/10.52041/serj.v9i1.388

Mackenzie, N., & Knipe, S. (2006). Research dilemmas: Paradigms, methods and methodology. *Issues in Educational Research, 16*(2), 193–205.

Torok, R. (2000). Putting the variation into chance and data. *Australian Mathematics Teacher, 56*(2), 25–31.

Watson, J., & Fitzallen, N. (2019). Building understanding of randomness from ideas about variation and expectation. *Statistics Teacher*. https://www.statisticsteacher.org/2019/09/19/building-understanding-of-randomness-from-ideas-about-variation-and-expectation/

Watson, J., Fitzallen, N., Fielding-Wells, J., & Madden, S. (2018). The practice of statistics. In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International handbook of research in statistics education* (pp. 105–137). Springer.

Watson, J., Fitzallen, N., & Wright, S., & Kelly, B. (2022). Characterizing student experience of variation within a STEM context: Improving catapults. *Statistics Education Research Journal, 21*(1), Article 9. https://doi.org/10.52041/serj.v21i1.7

Waxer, M., & Morton, J. B. (2012). Cognitive conflict and learning. In N. M. Seel (Ed.), *Encyclopedia of the sciences of learning.* Springer. https://doi.org/10.1007/978-1-4419-1428-6_280