THE POWER OF PERCENT

Jane Watson
University of Tasmania
<jane.watson@utas.edu.au>
Lyn English
Queensland University of Technology
<l.english@qut.edu.au>

Jane Watson and Lyn English use a chance activity exploring expectation and variation with coin tossing to highlight the importance of understanding the part–whole relationship embodied in percentage and its power to measure and compare for different wholes, in this case different sample sizes.

Introduction

Which statistic would you use if you were writing the newspaper headline for the following media release: "Tassie's death rate of deaths arising from transport-related injuries was 13 per 100,000 people, or 50% higher than the national average"? (Martain, 2007). The rate "13 per 100,000" sounds very small whereas "50% higher" sounds quite large. Most people are aware of the tendency to choose between reporting data as actual numbers or using percentages in order to gain attention. Looking at examples like this one can help students develop a critical quantitative literacy viewpoint when dealing with "authentic contexts" (Australian Curriculum, Assessment and Reporting Authority [ACARA], 2013a, pp. 37, 67).

The importance of the distinction between reporting information in raw numbers or percentages is not explicitly mentioned in the *Australian Curriculum: Mathematics* (ACARA, 2013b, p. 42).[1] Although the document specifically mentions making "connections between equivalent fractions, decimals and percentages" [ACMNA131] in Year 6, there is no mention of the fundamental relationship between percentage and the raw numbers

1 The activities in this study reinforced the Chance descriptors for Year 4 (ACARA, 2013b, p. 33) including ordering chances of events, and identifying events that cannot happen at the same time and events that do not affect each other. They also required the recognition of variation in results as suggested in the General Capability of Numeracy Learning (ACARA, 2013a, p. 46).

represented in a part–whole fashion. Such understanding, however, is fundamental to the problem-solving that is the focus of the curriculum in Years 6 to 9. The purpose of this article is to raise awareness of the opportunities to distinguish between the use of raw numbers and percentages when comparisons are being made in contexts other than the media. It begins with the authors' experiences in the classroom, which motivated a search in the literature, followed by a suggestion for a follow-up activity.

Context: Exploring probability

As part of a research project on beginning inference with Year 4 students, students undertook an exploration of the chances of getting a head when a normal coin is tossed, the expectation of the number of heads in 10 tosses, and the percentage of heads in larger and larger numbers of tosses. The purpose of the investigation was to experience *variation* and *expectation* (Watson, 2005) when they arose in the chance context, as a foundation for later work drawing informal inferences. Students had an *expectation* of obtaining half heads in a number of trials ('theoretical' chance of $\frac{1}{2}$) but experienced much *variation* from this *expectation* as they carried out small numbers of tosses. After conducting some trials 'by hand', the software *TinkerPlots* (Konold & Miller, 2011) was used to simulate larger and larger numbers of tosses. Students recorded their outcomes in tables with the number of heads and the percentage of heads for each number of tosses in side-by-side columns. They were then asked to calculate the range of percentage outcomes for simulations of size 10, 100 and 1000. Except for one pair of students, all groups found a decreasing range as the number of trials increased. Students were able to write summaries of this observation in their workbooks.

A few students, who finished early, were asked to plot their results on number lines to demonstrate the reduction in range. This request resulted in some surprises for the authors, such as the student who drew number lines for the actual numbers of heads each time. Although realising that the scales for the lines would be different for 100 and 1000 trials to fit on the paper, the result made the outcomes look similar, giving an inappropriate impression of the reduction in variation (similar to Figure 1). The first step for the student was confusion and then the realisation that numbers reporting frequencies do not tell the story: 517 – 463 = 54 is much bigger than 57 – 43 = 14 — but the variation is supposed to be smaller! It is the part of the whole that is important, not the actual numbers. The student's second attempt (see Figure 2) showed the appropriate percentages but the scales were different. When asked why the plot did not agree with the reduction in the range, the student had an 'aha' moment about the scale on the plots and proudly produced the equivalent of Figure 3.

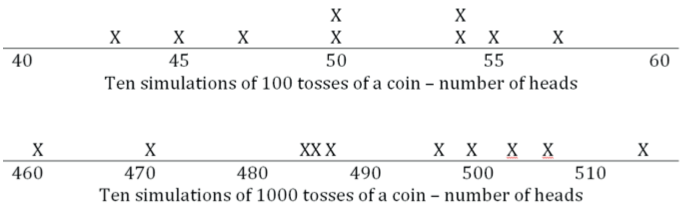

Figure 1. Number of heads in 100 tosses and 1000 tosses.

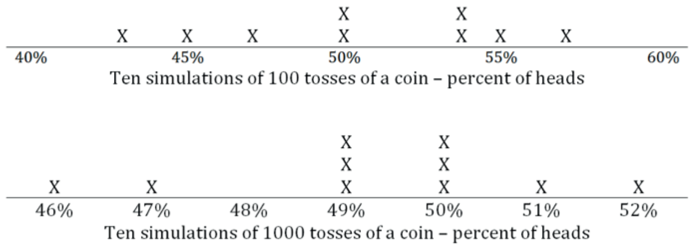

Figure 2. Percentage of heads in 100 tosses and 1000 tosses.

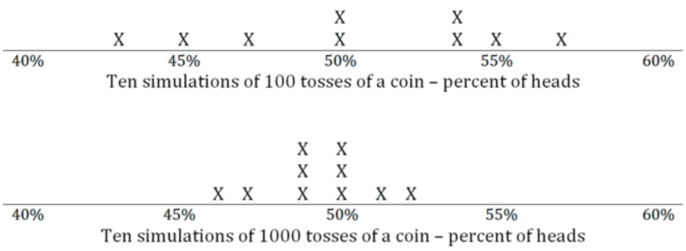

Figure 3. Percentage of heads in 100 tosses and 1000 tosses on the same scale.

Existing work on percentages

The experience with this student and several others led the authors to seek evidence in the research literature on the understanding of the basic relationship of frequency and percentage, and the power of percentages to provide a linear order. Of particular interest to us is not the reporting of students' attempts to complete computation tasks related to percentage, but the consideration of what percentage actually is as a representation when encountered in a descriptive context.

There has been much written for teachers on proportional reasoning, but much less specifically on percentage. The focus of writing in the area of percentage, however, appears to be on solving problems rather than appreciating the basic power of percentage to tell part–whole stories. Hawera and Taylor (2011), for example, consider working out the original mass of a product if the new version of the product is 20% larger, using a strategy similar to that suggested by Dole (2004). Watson and Beswick (2009) look in detail at the calculation of percentages but also use the percentages to order data on salt content of foods on a linear scale, with a similar purpose to the use shown in Figure 3 displaying variation.

A very useful source for background on the power of percentages to convey part–whole situations is the extensive literature review of Parker and Leinhardt (1995). In considering students' difficulties in learning about percentage, they explored a number of ways in which the concept can be applied. Of relevance here is Parker and Leinhardt's consideration of percentage as a number, commencing with the number-like extensive aspect of quantity, perhaps as a certain quantity out of 100. In this sense, percentages can be ordered linearly for ease of comparison. Percentages can also be added if the context is right and if they represent portions of the same whole (e.g., some probability tasks). Ordering is an important feature in the example we have offered due to the need to display the range of variation.

Parker and Leinhardt further examined percentage as an intensive quantity, one showing a relationship; of importance here, it is a part–whole relationship, perhaps embodied as a fraction or ratio. Lastly, they described percentage as a statistic, with the purpose of either reporting a relationship between known pieces of data or computing a functional expression such as taxes or discounts. It is the number-like qualities of ordering and showing a part–whole relationship of percentage as a statistic that are again a feature in the present example. In the remainder of this article, we describe another probability activity where this notion of percentage plays a key role.

Further exploration of percentage in a probability context

The issue related to percentage that initiated this exploration was based in a statistical investigation of creating simulations to confirm (or otherwise) a theoretical probability model. As one increases the number of simulations, one expects the outcomes to approach more closely the expected probability. Expressed another way: as the sample size increases, one expects the variation to decrease between the relative frequency of the outcomes and the theoretical value.

As an example, Figure 4 shows three simulations for tossing a regular six-sided die, where one would expect equal numbers for each of the six outcomes (a uniform distribution).

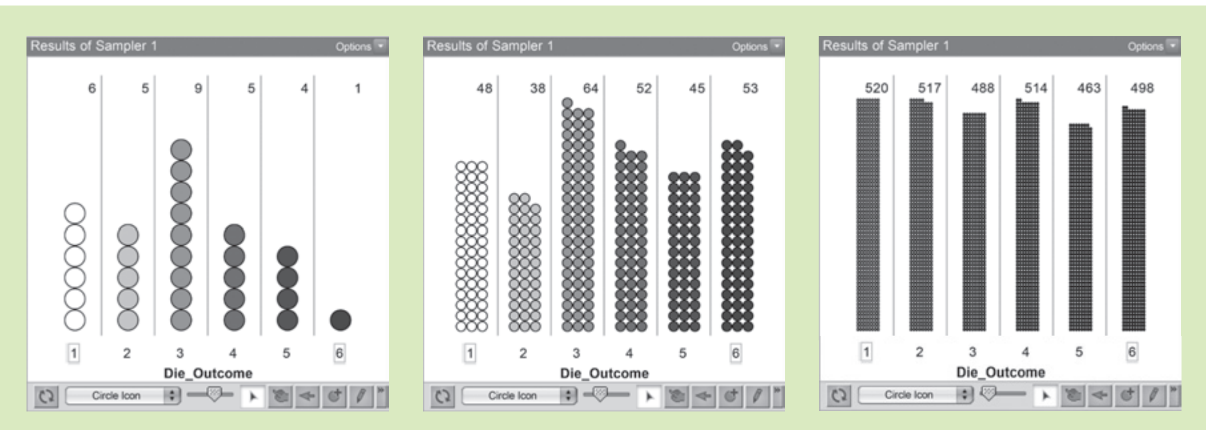

Figure 4. Numbers of outcomes for 30, 300 and 3000 simulations of 6-sided die tosses.

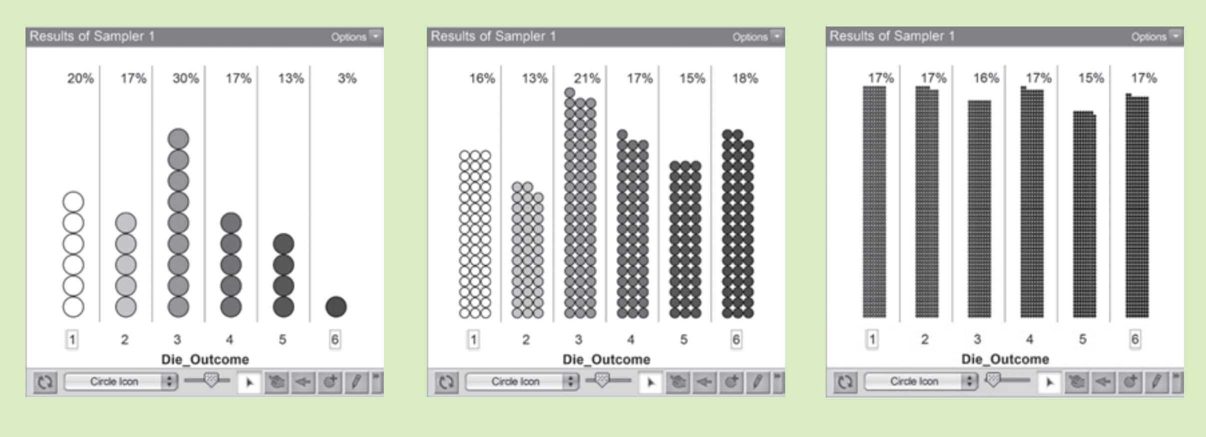

Figure 5. Percentages of outcomes for 30, 300 and 3000 simulations of six-sided die tosses.

On the left are outcomes for 30 tosses of the die; in the centre are outcomes for 300 tosses; and on the right, outcomes for 3000 tosses. Each of these outcomes, 1 to 6, is labelled with the number of times it occurs. Variation can hence be seen between the numbers of outcomes for 3 and 4, over the three simulations: $|n(3) - n(4)| = 4$ for 30 trials; $|n(3) - n(4)| = 12$ for 300 trials; and $|n(3) - n(4)| = 26$ for 3000 trials. Looking at the stacked dots, it does appear that they are approaching each other in height, but the numerical differences are getting further apart.

This can be very confusing for students who are likely to claim by looking at the numbers that the variation is increasing not decreasing. What is, of course, missing from the analysis is the transition from a frequency-count way of comparing the outcomes to a relative frequency that acknowledges the part–whole relationship between the individual outcomes and the total number of trials. This relationship is encased in a percentage, as shown for the same three data sets in Figure 5. The power of percentage in this context is its ability to provide a relative measure that can be compared with others, in this context to show results of the simulations approaching a theoretical value, in this case a difference of zero. For example for 30 trials, $|\%(3) - \%(4)| = 13\%$; for 300 trials, $|\%(3) - \%(4)| = 4\%$; and for 3000 trials, $|\%(3) - \%(4)| = 1\%$.

The conceptual dilemma in this situation seems to be the transition to seeing the whole in a part-whole relationship as being as important as the part. When presented with the results as in Figure 5 with the percentage for each outcome, the hope is that students will see the value of the part–whole representation and what it means for the purpose of the investigation. Although playing a proportional role in representing the relative frequencies, the percentages also play an additive role in being able to rank the differences linearly to observe them approaching the theoretical value of 0 based on the probability model. In this case, many plots are possible showing the decreasing variation as the number of tosses increases. Similar to what occurred in Figure 3, the differences of the percentages of threes and fours, for repeated sample sizes of 30, 300, and 3000 are shown in Figure 6. Alternatively, the results for any one of the six outcomes can be followed to approach 17% (approximating $\frac{1}{6}$). Different groups in a class could be given one of the six numbers to trace and plot. The class could then compare the results.

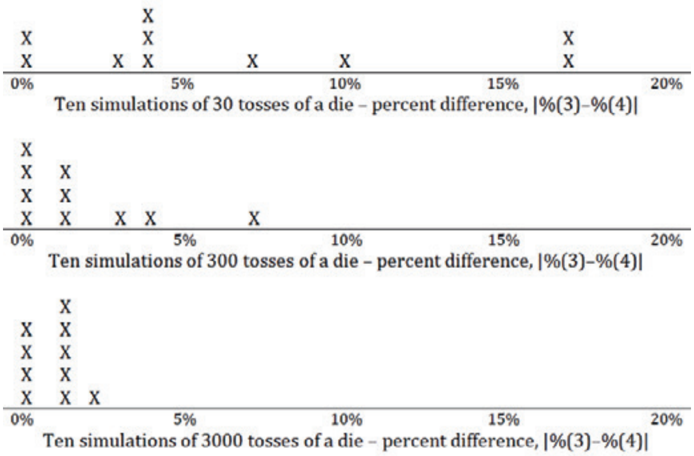

Figure 6. Difference in the percentages of threes and fours in 10 simulations of 30, 300 and 3000 tosses of a die.

Conclusion

The problem that gave rise to this discussion was the display of decreasing variation with increasing sample size in a statistical investigation. Using percentage meaningfully was an essential ingredient of success. This led to thinking of ways to use probability investigations to enhance both the part–whole and additive features of percentage. The question then arises as to whether other mathematics educators have explored this transition and the usefulness of percentage itself as comparable measure. More examples would be useful for teachers.

Returning to the question at the beginning of this paper, being a critical quantitative literacy thinker (ACARA, 2013a, p. 67) requires questioning of each presentation of either raw numbers or percentages. It is likely that one representation does not tell the entire story without information on the population total from which a frequency or a percentage is reported. Kluger (2006), in writing about people's understanding and assessment of risk, claimed two of the issues were (1) difficulties in people's intuitions in interpreting percentages and (2) deliberate stating of numerical values rather than percentages by those who want to increase perception of hazard (p. 45). Watson (2007) explored further the issue of reporting frequencies and rates with examples related to deaths of elephants and to fatal shark attacks. Developing the understanding explored in this article may assist students in asking critical questions of reports in many contexts.

Acknowledgement

This report arose from a research study supported by a Discovery Grant (DP120100158) from the Australian Research Council (ARC). Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the ARC. We wish to acknowledge the enthusiastic participation of the classroom teachers and their students, as well as the excellent support provided by our senior research assistant, Jo Macri.

References

Australian Curriculum, Assessment and Reporting Authority (ACARA). (2013a). *General capabilities in the Australian Curriculum, January, 2013*. Sydney, NSW: ACARA.

Australian Curriculum, Assessment and Reporting Authority (ACARA). (2013b). *The Australian curriculum: Mathematics, Version 4.1, 1 February 2013*. Sydney, NSW: ACARA.

Dole, S. (2004). *Mental computation: A strategies approach. Module 6: Ratio and percent*. Hobart: Department of Education Tasmania and the University of Tasmania.

Hawera, N. & Taylor, M. (2011). "Twenty percent free!" So how much does the original bar weigh? *Australian Primary Mathematics Classroom, 16*(4), 3–7.

Kluger, J. (2006, December 4). Why we worry about the things we shouldn't … and ignore the things we should. *Time, No. 48*, 40–45.

Konold, C. & Miller, C.D. (2011). *TinkerPlots: Dynamic data exploration* [computer software, Version 2.0]. Emeryville, CA: Key Curriculum Press.

Martain, T. (2007, January, 13). The unsafe state… *The Mercury* (Hobart, Tasmania), p. 11.

Parker, M. & Leinhardt, G. (1995). Percent: A privileged proportion. *Review of Educational Research, 65*, 421–481.

Watson, J. M. (2005). Variation and expectation as foundations for the chance and data curriculum. In P. 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, Melbourne, pp. 35–42). Sydney: MERGA.

Watson, J. M. (2007). Linking the big ideas in middle school maths. In K. Milton, H. Reeves & T. Spencer (Eds), *Mathematics: Essential for learning, essential for life* (Proceedings of the 21st biennial conference of the Australian Association of Mathematics Teachers, Hobart, pp. 82–94). Adelaide: AAMT Inc.

Watson, J. & Beswick, K. (2009). Which is bigger: 250 tonnes or 17%? A tale of salt. In C. Hurst, M. Kemp, B. Kissane, L. Sparrow & T. Spencer (Eds), *Mathematics: It's mine* (Proceedings of the 22nd biennial conference of the Australian Association of Mathematics Teachers Inc., Fremantle, pp. 175–184). Adelaide: AAMT Inc.