The Classicist and the Frequentist Approach to Probability within a *TinkerPlots*2 Combinatorial Problem

Theodosia Prodromou
The University of New England
<tprodrom@une.edu.au>

This article seeks to address a pedagogical theory of introducing the classicist and the frequentist approach to probability, by investigating important elements in 9th grade students' learning process while working with a *TinkerPlots2* combinatorial problem. Results from this research study indicate that, after the students had seen the systematic construction of the event space via combinatorial analysis, they viewed the sample space as an essential property that regulated the results of the distribution of each sum's theoretical frequency.

Introduction

The biggest leaps forward in the next several decades- in business, and society at large- will come from insights gained through understanding data. It is of paramount importance to acquire the mathematical skills and understandings required to enable citizens to become informed about our world and to understand the data that politicians, advertisers and other advocated are using to promote particular causes that will impact on the future of our planet. This embraces the capacity of not only understanding the underlying messages that the data attempt to reveal but also critically examining the probabilistic statements presented in news media in terms of data and data representations such as charts, tables, and graphs.

These skills and everyday understanding of mathematics might be best described as quantitative literacy (Steen, 1997). Quantitative literacy was defined as "an aggregate of skills, knowledge, beliefs, habits of mind, communication capabilities, and problem-solving skills that people need in order to engage effectively in quantitative situations arising in life and work" (Steen, 2001, p. 7). Statistical literacy is an important part of quantitative literacy skills. As Wallman (1993) suggested in her Presidential Address to the American Statistical association, there is a need to enhance citizens' "ability to understand and critically evaluate statistical results that permeate daily life and to acknowledge the contributions that statistical thinking can make in public and private, professional and personal decisions." (p. 1).

Australian Curriculum and Reporting Authority K-10 (ACARA, 2010) recognizes that the twenty-first century world is information driven, and through Statistics and Probability students can make informed judgments about events involving chance. The role of probability as a central component for statistical investigations has engendered the need for "probability literacy" to deal with a variety of real-world situations that encompass interpretation or generation of probabilistic messages as well as decision-making. Elements of probability-related knowledge and some dispositions that may be the building blocks comprise probability literacy were proposed by (Gal, 2005), but they have received little explicit attention in discussions of people's numeracy and statistical literacy. Gal (2005) listed these elements as follows: (1) several foundational big ideas such as randomness, independence, and variation; (2) Figuring probabilities of events in order to estimate the probabilities of events, (3) Language of chance, that is the terms and phrases related to chance and the various ways to represent and communicate about actual events; (4) Understanding the role and the implications of probabilistic issues and messages in different

In J. Dindyal, L. P. Cheng & S. F. Ng (Eds.), Mathematics education: Expanding horizons (Proceedings of the 35th annual conference of the Mathematics Education Research Group of Australasia). Singapore: MERGA.

contexts and in personal and public discourse; and (5) Critical questions to reflect upon when dealing with probabilities.

The proposed view of probability literacy illustrates a "scope of probability at the school level to reflect the study of random events, the development of appropriate probabilistic intuitions, a basic understanding of language and simple events, an appreciation of distribution and the addressing of misconceptions" (Watson, 2006, p. 127-128). This scope can be seen as consistent with the tendency to place less emphasis on the knowledge pertaining to theoretical probability when addressing important issues relevant to teaching data-based statistics (Moore, 1997). Watson (2006) claims that the aim in statistics teaching is to enhance data handling through an empirical frequency-based approach to probability that is an essential foundation for later work in the study of theoretical probability based on sample spaces. Continuous classroom-based investigations based on frequencies may reinforce the building of an appreciation of a frequency approach to probability when performing trials and comparing favourable outcomes to total outcomes. Watson (2006) also claims that it is not always appropriate to introduce an experiment to calculate relative frequencies before suggesting a theoretical model based on the possible outcomes of a sample space. Ultimately, introducing the environment of theoretical probability where data will be collected should not be left behind or dismissed. Hands-on simulations and simulation software provide students with opportunities to explore the nature of the sample space and probability distributions, particularly in the light of refining understanding of variation.

In order to better understand pedagogical theory, we have developed an experiment that uses Watson's theory of pedagogy, and will analyse the students' behaviour using Radford's (2009) theory of cognition to guide our attention to important elements in the students' learning processes. Hence, this article examines Watson's (2006) claim, by having students explore the generation of a sample space within *Tinkerplots2*.

Radford's (2009) theory of cognition is the knowledge objectification, a theoretical perspective on teaching and learning in which learning is taking place as a social process through which students become progressively conversant with cultural forms of reflection (Radford, 2009). Within the theory of objectification, the distinctive sensuous and artefact-mediated nature of thinking emphasizes the semiotic means of objectification through which knowledge is objectified. The semiotic means of objectification include kinaesthetic actions, gestures, signs (e.g. mathematical symbols, graphs, inscriptions, written and spoken language), and artefacts (e.g. rulers, calculators). Thus, in what follows, in the practical investigation of 9th grade students' probabilistic thinking, attention will be paid to important elements in the students' learning processes when students explore the generation of a sample space within *Tinkerplots2*.

Data Collection

The data come from an ongoing research study conducted in a rural secondary school. The data have been collected during regular mathematics lessons. In these lessons the students spent extensive time working in pairs. The researcher interacted continuously with the pairs of students during the pair work phase in order to probe the reasons that might explain their thinking. The data collected included audio recordings of each pair's voices and video recordings of the screen output on the computer activity using Camtasia software. When students' body language or facial expression appeared to be indicative of their conceptual evolution, notes were kept. The researcher (Re) prompted students to use the mouse systematically to point to objects on the screen when they reasoned about computer-based phenomena in their attempt to explain their thinking. Plain paper was also available for students' use in case students needed to explain their thinking in a written form. This article focuses on one pair of students, Rafael (Ra) and George (Ge).

Data Analysis

In the first lesson, students watched an instructional movie that shows how to use *TinkerPlots2* features to build a simulation of rolling two dice. They then built a simulation of rolling four dice where the dice were presented by spinners (top right in Figure 1). They ran the simulation many times and graphed the sum of four dice (bottom in Figure 1).

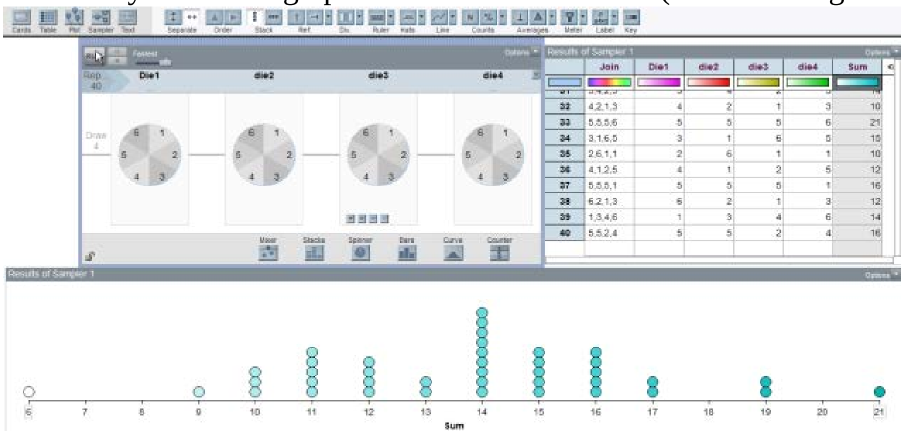

Figure 1. Simulation of rolling four dice in *TinkerPlots2* and the graph the sum of four dice.

In the second lesson, students watched a *TinkerPlots2* movie. *TinkerPlots2* provides a sampler that is essentially a non-conventional form of probability distribution. The *TinkerPlots2* movie showed how to use two counters to generate a sample space of rolling two dice (Figure 2). Each counter had numbers one through six. As the simulation ran, the

Figure 2. Two counters generate a sample space of rolling two dice

right counter circled though the numbers one through six, but the left counter stayed on one. Then the left counter advanced to two and the right counter circled through one to six again until all the possible outcomes were produced. The students observed the systematic listing of all possible outcomes of rolling two dice and the creation of a graph that shows the sum of two dice and the use of the sample space to calculate theoretical probabilities. The students were asked to use counters to build a sample space of rolling four dice (Figure 3) and then, after answering some questions, to come up with general rules applied.

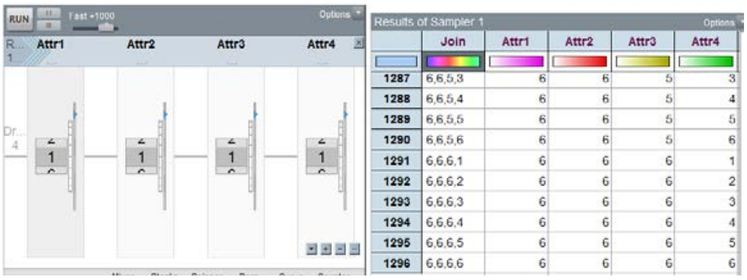

Figure 3. A sample space of rolling four dice

In tune with our theoretical framework, a multi−semiotic data analysis was conducted to investigate students' probabilistic thinking. At the first stage, the audio recordings were fully transcribed and screenshots were incorporated as necessary to make sense of the transcription. The most salient episodes of the activities were selected. Focusing on the selected episodes with the support of the transcript, along with a detailed account of significant actions, gestures, and artefacts used; I study the role of spoken language with the gestures including students' actions when pointing to objects on the computer screen and artefacts.

Results

The students built a simulation of rolling four dice and made a graph showing the sum of four dice (left, in Figure 4). They ran the simulation and observed the possible outcomes of rolling four dice. The researcher spent the rest of the first day discussing with the students the graph that showed the sum of four dice. They run the simulation 80 times.

1. Ra: We see 14 is the most likely sum. (Gesture−uses the mouse to point to the 14th column of the graph on the screen)
2. Ge: Run the simulation again.
3. Ra: This time it's 15. (G− uses the mouse to point to the 15th column of the graph on the screen)
4. Ge: Run it again, 80 times.
5. Ra: Okay, it's16 now. (G− uses the mouse to point to the 16th column of the graph on the screen)
6. Ge: It's a bit inaccurate.
7. Ra: It's in that general area. 13, 14, 15, 16 but they're not all high values. Like see, 14, 15 taken a drop here (right, Figures 4). (G-uses the thumb and the index finger of his right hand to point to the interval from the 13th to 16th columns and then uses the mouse to point to the 14th and 15th columns of the graph individually).

Rafael changed from using the mouse to point to individual columns of the graph, to using two fingers (the thumb and the index) to indicate and emphasize an interval on the graph and then again he used the mouse to point to individual columns of the graph. Gestures dominated Rafael's respond to the researcher's question about the most likely 4-dice sum or a range of 4-dice sums. The coordinated use of spoken language and the use of gestures served as semiotic means of objectification for Rafael. The boys continued their discussion about the most likely 4-dice sum.

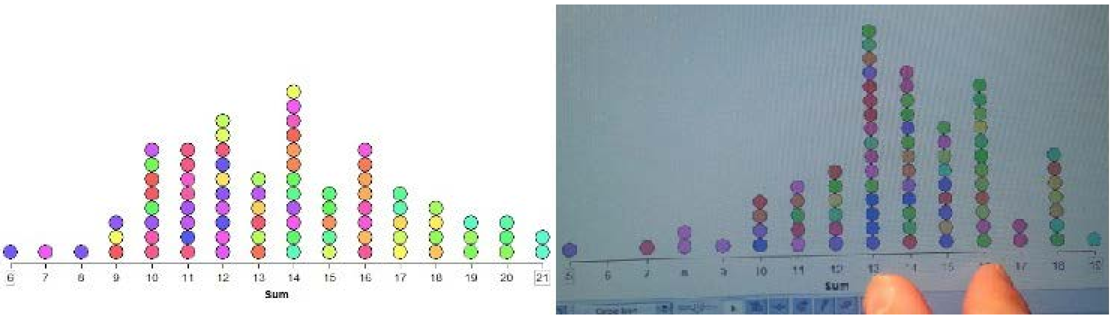

Figure 4. Graphs of the sum of four dice

8. G: Um, probably 14 because the first time it was the highest. (G− uses the mouse to point to the 14th column of the graph on the screen)
9. Ra: It's remained constant.
10. G: Second time wasn't yeah. It's remained constantly fairly high.
11. Ra: It's, um, generally it's usually the same area most likely so that would seem the most likely one that would come up. 13, 14, 15, 16 yeah, but particularly 14's always stayed constant. Like even though it's not the highest here, it's always stayed in the range of the top highest.
12. G: Yeah, it's in the top 3.

The boys reflected on previously observed graphs. Rafael became consciously aware of a region in which the most likely outcomes were included. This region was placed around the 14th column of the graph. Although the 14th column was not always the highest, it was, according to the boys, constantly included in that region. The boys continued to observe features of the graph while running the simulation several times (100 rolls each time).

13. Ra: Yeah, but 13, 14, 15, 16 they always get high. (G-uses the thumb and the index finger of his right hand to point to the interval from 13th to 16th column).
14. G: When 15 was our highest, 15 was the most likely outcome by overall.
15. Ra: If you're viewing in consistency standards, it would be 14. It seems to be the most consistently high, or the most consistently high or most consistent and reliably one to stay around that level or that range.
16. Re: Are you talking about the distribution of the outcomes.
17. Ra: Well it's generally seems to stay between um… 12, 13-ish to about 17. (G-uses the thumb and the index finger of his right hand to point to the interval from 12th to 17th column).

The boys ran the simulation several times increasing the number of rolls from 200 to 250. They both concluded that 14, 15, 16 have the highest consistency. George added:

18. G: That seems to be what's happened. That they are the most likely you're going to get.

The next day (2rd lesson) the researcher drew students' attention to permutations without using any explicit mathematical terminology.

19. R: Which 4-dice sum can be made up by adding different ordered numbers?
20. Ra: It should be … There is only 4 chances out of the possible outcomes. Like if I was to get 24 I would have to get 4 six's
21. G: That's very unlikely. We didn't get it there (referring to the graph they observed during the previous lesson).
22. Ra: The reason, the range of those numbers (14 to 15 range)…. were so high because there's so many dice. You don't need to get much of each dice. Those numbers there allow you to have a wider variety of combinations like um if I… 13, 14, 15 being the highest ones.

The boys were able to appreciate the connection between the outcomes from the previously observed graphs, with the different faces of 4 dices that summed up to those outcomes. While they attended to the shape of the graph, the permutations were not perceptually distinguishable as actual experimental outcomes. When the researcher asked students to list (on paper) those permutations that added up to 15, the boys began to randomly list the possible outcomes for rolling four dice.

23. Ra: If I write it down once, do you want me to write it down every other way I could write it down? I mean if I have, I have 6 – 4 – 3 – 2 do you want me to also write 3 – 4 – 6 – 2? Is it important?
24. Re: Yes. The order of the inscriptions is important. Why do you think that the order is important?
25. Ra: I'm thinking, I guess, I'm trying to like answer but I don't know if it's right but it's like it's, they vary…
26. Ra: … it's just umm, like the fact that it's variabled. Like it's varied umm, I think is important because like there is a better chance of getting that because the amount if variables that you get with the equation that equals up to 15….If that sort of makes sense, I don't know, I'm just thinking.

As the previous dialogue shows, Rafael seemed able to understand the uniqueness of permutations (i.e., that [3,5,6,1] and [5,3,1,6] were distinct) but he was unable to articulate why permutations are distinct. When the researcher mentioned that the order of inscriptions is important, Rafael seemed aware of the significance, but was unable to explain it. The researcher instead of impressing upon the students the importance of the ordered outcomes in a combinatorial experiment, asked students to watch a *TinkerPlots* movie which showed how to use two counters to generate a sample space of rolling two dice. The students observed the systematic listing of all possible outcomes of rolling four dice and made the graph that showed the sum of four dice (Figure 5).

27. Ra: That's how I would have worked out the four but I would have never ... it would take someone too long to do by hand anyway. And I suppose they also do this with locks and stuff. Yeah, when you get your bike lock. The 4 combinations 1 to 6. It says how many combinations. I honestly never thought there would be that many. Pretty, amazing!

When Rafael was previously asked to list all the possible outcomes of the sample space

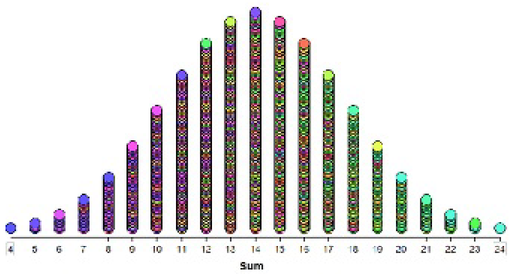

Figure 5. The that showed the theoretical outcomes of the sum of four dice

he had difficulties with constructing combinatorial type outcomes, because he either did not exhaust the sample space or duplicated possibilities. When Rafael watched how *TinkerPlots* features systematically listed all the possible outcomes of rolling four dice using the odometer strategy, he was surprised by the number of possible outcomes. The animated generation of a sample space made it possible for the students to see the use of various representations in solving the combinatorial problem situation including the use of animations, systematic listings, and a table holding the all possible outcomes of rolling 4-

dice; these are semiotic means of objectification. When the students attempted to talk about the graph that showed the theoretical outcomes (as opposed to the histograms generated by running a sample, like they were looking at earlier in the paper):

28. Ra: Look at this peak, being 13, 14, 15 high (G-uses the thumb and the index finger of his right hand to point to the interval from 13th to 16th column). We noticed that before (refers to the graph they observed during the first lesson).
29. G: The difference between these two is that one (the graph created using all the outcomes of the sample space) is a lot more accurate as it's going through every single stage compared to the other one.
30. Ra: This is a lot more thorough, as it goes through every single possibility you could have … It would probably be really helpful if you knew this, if you were gambling or placing bets… It shows you the possible outcomes if you're playing a game of chance. If you want to know the possible outcomes of the ace coming up. Like you want to know if there is more chance than it won't come up, than it will not. You could do that also, at horses.

The gesture (uses the thumb and the index finger of his right hand to point to the interval from 13th to16th column) reminded boys of the previously observed graph, so they began comparing the two graphs before the researcher would ask them (the researcher intended to ask such a question). After this interplay, we observe that although R and G had never been taught at school about the classicist and frequentist approaches to probability, the means of objectification helped them to develop understandings about the two approaches to probability. The students expressed a preference for using the classicist approach to conceptualise the probabilistic experiment. They viewed the classicist approach as a "more accurate" process, because it is based on all the possible outcomes of the sample space. The sample space appeared to the students as an essential property that can regulate the results of the graph, which shows the theoretical frequencies of the possible sums of 4-dice. After the students had seen the systematic construction of the event space via combinatorial analysis performed by features of *TinkerPlots2*, both Rafael and George made sense of the role of the event space in relation to the experiment it is said to model. They saw connections between the classicist approach and games of chance and suggested how to apply such a probabilistic approach to the solution of new problems.

Discussion

Our data offer an initial glimpse of the development of probabilistic thinking about combinatorial problems. It shows how the *TinkerPlots2* combinatorial problem implemented the pedagogical objective of enabling Grade 9 students with minimal probabilistic knowledge to understand the generation of the event space of rolling four dice. Students' "spontaneous" perception was successfully transformed through the interaction of students with the *TinkerPlots*2 combinatorial problem and the systematic listing of all possible outcomes of rolling four dice. This interaction might be conceptualized as occurring in the zone of proximal development out of which the students constructed new understandings about the relevance of order (permutations) in the construction of the event space. The researcher questioned students as they explored the combinatorial problem, to promote students' combinatorial understanding. The interaction of spoken language with gestures (using fingers to point to the graph, using the mouse to point to the graph on the screen), signs (graphs, written inscriptions, tables) and artefacts (computer animation) led students to modify, or refine some of their original structuring relationships between the semiotic means of objectification, or to consolidate new structuring relationships and understandings of probabilistic concepts related to both the classicist and frequentist approaches.

In this paper, there is evidence that the students expressed a preference for using the classicist approach to conceptualise the probabilistic experiment. The results should not be interpreted to indicate that activities that enable students to encounter classicist approaches to key principles of combinatorial concepts are sufficient to achieve desired pedagogical objectives.

References

Australian Curriculum, Assessment and Reporting Authority. (2010). *Australian Curriculum: Mathematics*. Version 1.1. Retrieved March 15, 2011, from http://www.acara.edu.au

Gal, I. (2005). Towards "probability literacy" for all citizens: building blocks and instructional dilemmas. In G. A. Jones (Ed.), *Exploring Probability in School: Challenges for Teaching and Learning* (p. 39-63). New York: Springer.

Moore, D. S. (1997). New pedagogy and new content: The case of statistics. *International Statistical Review*, *65*(2), 123-165.

Radford, L. (2009). Why do gestures matter? Sensuous cognition and the palpability of mathematical meanings. *Educational Studies in Mathematics*, *70*, 111-126.

Steen, L. A. (Ed.). (1997). *Why numbers count: Quantitative literacy for tomorrow's America*. New York: College Entrance Examination Board.

Steen, L. A. (Ed.). (2001). *Mathematics and democracy: The case for quantitative literacy*. Washington DC: Woodrow Wilson National Fellowship Foundation.

Wallman, K. K. (1993). Enhancing statistics literacy: Enriching our society. *Journal of the American Statistical Association*, *88*, No. 421, 1-8.

TinkerPlots: *Dynamic data exploration* (Version 2.0) [Computer software]. Emeryville: CA: Key Curriculum Press.

Watson, J. M. (2006). *Statistical literacy at school: Growth and goals*. NJ: Lawrence Erlbaum Associates.