Probabilistic Thinking: An initial look at students' meanings for probability

Neil J. Hatfield
Arizona State University

Probability is the central component that allows Statistics to provide a useful tool for many fields. Thus, the meanings that students develop for probability have the potential for lasting impacts. Using Thompson's (20015) theory of meanings, this report shares the results of examining 114 undergraduate students' conveyed meanings for probability after they received instruction.

Key words: probability, statistics, meanings, introductory statistics course

Probability is the engine that makes inferential Statistics run. This first statement is one that hardly any practitioner of Statistics will disagree with. Statisticians and Statistics Educators freely acknowledge that the central ideas of probability allow us to move beyond merely describing a data set to using the data set as evidence for supporting/refuting claims. This is even one area in which Frequentists, Bayesians, and Subjectivists all agree. However, how practitioners think is often vastly different from how students think before, during, and after instruction. In a recent discussion with a university instructor about introductory statistics courses, I was surprised to hear this individual say "I skip by probability because my students don't really need it and we need the time to talk about doing hypothesis tests." This statement caught me off-guard for two reasons: 1) this instructor had a Ph. D. in Statistics, and 2) the instructor continued to talk about how she wanted her students to develop "rich and productive meanings for hypothesis tests and *p*-values". While I believe that students *can* and *will* develop meanings for hypothesis tests in the absence of a way of thinking about probability, I challenge the claim that students can develop "rich and productive" meanings.

Cobb and Moore (1997) took the position that "first courses in statistics should contain essentially no formal probability theory" (p. 820). I agree with the spirit of their position. While this may seem like I am in the camp of the aforementioned instructor, there is a critical distinction. Cobb and Moore's position is *not* that first courses should avoid discussing probability, but rather emphasis on formal rules, such as the rules for $P\left(A \cup B\right)$, are of little consequence in these courses. Rather, they suggest that "informal probability" is sufficient, especially if the course focuses on the idea of sampling distributions. Thus, rather than skipping over probability entirely, an introductory course should skip over calculational rules of probability and focus on helping students construct *ways of thinking* about probability. I agree with Liu and Thompson (2002) that trying to debate the question of "What *is* probability?" is a fruitless endeavor in a first course. Rather, in a first course on statistics and probability, our focus should be on what we (us and our students) *mean* by the term "probability".

Introductory statistics texts that cover probability focus almost entirely on *how to calculate* rather than *how to think about*. The introductory text *Statistics for the Life Sciences,* 4th edition (Samuels, 2012) devotes 15 pages to probability. However, there are only two sentences related to how to think about probability. Out of the 29 exercises provided for the students to use for homework, 26 ask for students to calculate a value of the probability of some event, 3 ask students to make a claim about whether or not two events are independent, and 0 questions ask students to interpret/make use of a way of thinking about probability. Likewise, *Introduction to the Practice of Statistics*, 7th edition (Moore, McCabe, & Craig, 2012) devotes 18 pages to probability and randomness. Of these pages, only 3 sentences (all variations of each other) focus on how to think about probability. There are only two questions of the 45 that focus on something other than a calculation of probabilities or judgment of independence; one asks whether or not a probability value is applicable to a larger set of colleges, and the other asks students to explain what a probability value means. In both of these cases, students' major takeaway is that probability is a calculation.

The above issues created a backdrop for an informal, observational study that aimed to serve as a first step in looking at how undergraduate students think about probability after enrollment in an introductory course on Statistics. In the spring of 2014, students in the four sections of an introductory statistics course designed for life science majors at a large, public, Southwestern university responded to two questions related to their thinking about probability. A senior lecturer taught two sections of the course, and two graduate students (one Ph. D. Statistics and one Ph. D. Mathematics/Statistics Education) each taught one section. While the aforementioned *Statistics for the Life Sciences* served as the official text for the course, only three sections followed this text. The fourth section (taught by the Ph. D. Math/Stat Ed. graduate student) moved away from the text. This section followed an experimental design curriculum intended to support students developing ways of thinking about statistics beyond procedures and placed a heavy emphasis on meanings. In addition to the students answering the questions, the three instructors also answered the questions.

Methodology and Theoretical Background

I conducted this observational study at a large, public university located in the Southwestern region of the United States in the spring of 2014. Given that this study only serves as a first step to a more formal study, the selection of course and school was convenient to the researcher. The instructors of the selected introductory statistics course for life science majors asked their enrolled students to respond to three questions after the students had already received instruction on probability. Two questions related directly to the general purpose of this study: how the students think about probability.

Question 1: How do you think about probability? That is, how would you explain probability to another person?

Question 2: Consider the following statement:

The probability of observing a value of 4 when looking at the product of two dice is 3/36.

How should someone think about (interpret) 3/36 given the above statement?

Given the qualitative nature of the written responses to these questions, I made use of a coding methodology consistent with that of Strauss and Corbin (1990). I initially used open coding for the responses and then I made use of an axial coding system. I used the axial codes in my analysis.

In coding, I focused on the meaning conveyed by the students' responses. *Meaning* refers to the space of implications (including actions, images, and other meanings) that results from an individual assimilating some experience and thereby forming some understanding of that experience (Thompson, 2015). Using a student's responses, we may postulate the meaning that he/she has for probability. Just as responses may be viewed as more/less productive, so can meanings. I view *productive meanings* as those meanings that provide coherence to ideas that students have and those meanings which afford students a frame that supports the students in future learning (Thompson, 2015). *Productive meanings* are clear, widely applicable (within reason), and rely on explicated assumptions. To help demonstrate this, let us look at an example using Question 2. Suppose that we have two students, George and Sally. George's meaning for probability deals with notions of long-run relative frequency while Sally's meaning is a blend of circular and Classical (see Results for details on these meanings). George's meaning orients him to view 3/36 as a measure of the relative frequency of seeing a value of 4 when carrying out the dice experiment an indefinite, large number of times. George's thinking supports him in making statements such as "That 3/36th of the time, we will observe a product of four" or perhaps "About 8% of the time, we'll observe a product of four". George's meaning is flexible enough so that if his teacher adds information that the two die have an unequal number of sides, or the dice are not fair, he won't feel a need to alter his initial response. For George, his initial interpretation works in light of this new information.

On the other hand, Sally's meaning supports her in thinking about 3/36 as a statement that there are 3 ways to get a product of 4 when rolling two dice out of 36 total ways of getting products. Implicit to Sally's thinking is that the two dice are standard, six-sided dice. This fact enables Sally to make sense of the 36. If the teacher were to reveal that one die was four-side while the other was a twenty-side die, Sally would struggle to make sense of the 36. Likewise, should the teacher state the two dice are unfair, but not state how they are unfair; Sally's meaning for probability does not necessarily enable her to give an interpretation in light of unfair dice. Additionally, the use of equivalent fractions could create issues for Sally's way of thinking. We know that 3/36 reduces to 1/12. However, the interpretation could change significantly for Sally; "there is only 1 way to get a 4 and 12 possible outcomes". The underlying process is no longer the same; a Classical meaning appears to allow individuals to ignore/forget the process all together. Additionally, 3/36 could be re-written as 4/48, 7/84, or even 30/360. The same issue with reducing still applies. George's meaning, particularly if he moves from the fraction to a percentage, will have no issue with using an equivalent fraction.

The productivity of George's and Sally's meanings for probability has broader implications. Consider the statement "The probability of selecting a random US man, 20+ years old, who is under six foot tall is 2/3." in place of the dice statement. George's meaning still supports him reasoning about the relative frequency of observing US men in the age group whose height is under 6ft. However, Sally either has to reason that there are 2 heights under six foot out of 3 total possible heights that US men can be or she needs to have a completely separate meaning for probability in continuous contexts. Having separate meanings for probability in different contexts does not lend itself to the student building a coherent way of thinking about probability.

Results

I characterized students' responses to Question 1 in five broad categories. The first category of responses deals with thinking about probability as being about the long-run relative frequency of some event (L.R.R.F.). For these students, they seem to think about probability as something that emerges after imagining carrying out some process a large number of times. While these students may speak about the probability of some event, from discussions they do not appear to think that the event is the next outcome of the process. Rather, they always reference needing to imagine the process carried out many, many times.

The second category is "Frequency" and contains all of the cases where the students appeared to focus on the frequency (or relative frequency) of some event occurring, but their responses do not clearly indicate that the student imagines the frequency stemming from repeating a process an indefinite number of times.

The third category covers those students' responses that dealt with prediction. The responses that fall into this category are reminiscent of the outcome-approach of probabilistic thinking (Konold, 1989). Often these students only spoke about the very next time you carry out some process.

The fourth category of responses I called "Circular". Typical responses that fall into this category are "Probability is the chance that something happens", or "the likelihood of some event". The descriptor of "circular" is highly indicative of how these students seem to think. During discussions, students who spoke of probability as being "chance" or the "likelihood" of some event, would often answer the follow up question of "What is chance/likelihood?" with the statements along the lines of "well, chance is, umm, just probability." The way students thought about probability appeared to be a near unending cycle of labels with little meaning behind those labels. The seemingly only way these student broke out of this cycle was when they had to deal with a concrete situation, a specific value for them to speak about, and, occasionally, restrictions on what words they could use (i.e. not use the words "chance", "likelihood", "probability"). The fifth category, "Other", covers those responses not captured by the other categories.

The following bar chart (Figure 1) shows the frequency of responses that fall into these categories. Overwhelmingly, 89 students (78.1%) gave a response that seems indicative of circular thinking. Nineteen students (16%) appear to think about probability in terms of frequency/relative frequency. Of these students, 15 think about probability as the long-run relative frequency of some process.

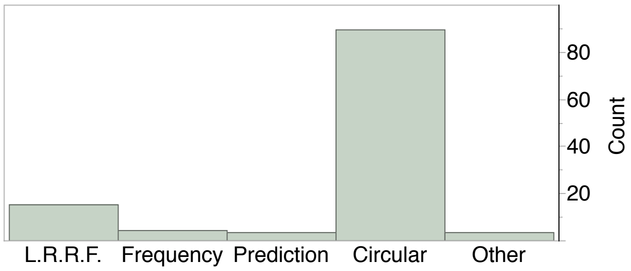

Figure 1. Students' responses for Question 1.

How do students appear to interpret a specific value of probability?

I used five codes to characterize students' responses for interpreting the probabilistic value 3/36. The first category consists of those students who seemed to think of 3/36 as one number rather than two numbers separated by a bar. These students spoke about 3/36 as representing the percent of the time you would see a product of 4 if you carried out the process of rolling two dice an indefinite number of times ("many, many times").

The second category, "Classical", covers those responses where students appeared to view the fraction 3/36 as two numbers. The upper number represented the number of ways to get the outcome of interest while the second number represented the total number of different outcomes. This way of thinking is exactly like that used in "classical" probability. In this school of thought, the sample space consists of a finite number of unique outcomes, which we assume as having the exact same probability of happening. In addition to the equi-probability assumption, the students also must make the assumption about details of the stochastic process. Namely, that there are 36 distinct outcomes.

Similar to the "Classical" category, another group of responses reflected thinking about the probability value 3/36 as telling us that either we already had observed 36 rolls of two dice and saw exactly 3 products of 4 or if we were to roll the dice 36 times, we would then see exactly 3 products of 4 ("fixed number of rolls"). These students also appear to view the fraction as two numbers. In both cases, students appear to think that the probability value tells us exactly how many outcomes of interest we saw for a set number of trials.

The forth category, "Chance", are those students whose response to the question was to essentially say that 3/36 was the chance getting a product of 4. The fifth category, serves as the catch-all for responses that did not fall into any of the other categories. This includes students who repeated the given statement (4 students), either simplified or wanted simplification of 3/36 (2 students), or expressed the need for dice (2 students) among the responses. I did not include the three students did not respond to this question in the final count.

As shown in the following bar chart (Figure 2), a majority of students interpreted the probability value as being about a fixed number of rolls of the dice and a fixed number of 4's (40.5%). Only 17.1% (19) of the students thought about 3/36 as representing the percent of the time we would see a product of 4. Fifteen students appeared to use a "classical" way of thinking, while 19 just substituted "chance" for "probability".

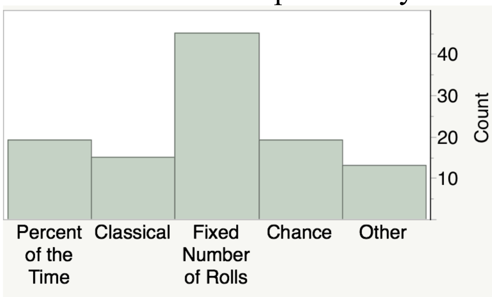

Figure 2. Students' responses for Question 2.

How does apparent student thinking about probability in general relate to how they interpreted a specific value of probability?

A natural question that follows from the previous two questions, is how do the students' responses to each question relate to one another? Table 1 shows the two-way contingency table for students' responses to both questions. The vast majority of individuals who appeared to think about probability as the long-run relative frequency of some event interpreted the given probability value as the percent of the time we would see some event happen in the long run. The majority of students who interpreted 3/36 as being two number separated by a bar (either Classical or Fixed Number) or as a "measure of chance", gave a circular meaning for probability. The wide range of interpretations given by students with a circular meaning is not surprising. Given that the students' meaning for probability appears related to a word-exchange, the students would need to draw upon some other meanings to help make sense of the value 3/36. All but one student who explained 3/36 as the "chance" of getting a product of 4, gave responses that indicated a circular meaning to Question 1.

Table 1. Students' responses to Question 1 by their responses to Question 2.

	Percent of the Time	**Classical**	**Fixed Number of Rolls**	**Chance**	**Other**	total
L.R.R.F.	12	1	0	1	0	14
Frequency	1	0	1	0	2	4
Prediction	2	1	0	0	0	3
Circular	4	12	42	18	10	87
Other	0	0	1	0	2	3
total	19	14	44	19	18	111

Is there a difference between how students appear to think about probability in general when accounting for the instructor?

To explore this question, Table 2 provides a good visualization of how students appeared to have thought about probability in regards to Question 1. A striking aspect to notice is that all of the students who appeared to think about probability as a long-run relative frequency all have Instructor A. Additionally, the vast majority of students for both Instructor B and Instructor C gave responses that appear indicative of a circular meaning for probability in general.

Table 2. Students' Responses to Question 1 by Students' Instructor

	L.R.R.F.	Frequency	Prediction	Circular	Other	total
Instructor A	15	2	2	8	0	27
Instructor B	0	2	1	58	1	62
Instructor C	0	0	0	23	2	25
total	15	4	3	84	8	114

To further explore this difference, I conducted a Kruskal-Wallis test with $\alpha = 0.05$. The test statistic has a value of 58.0382. Thus under a χ_2^2 distribution, the approximate probability of observing the differences we did or ones more extreme is $p \leq 0.0001$. A post-hoc analysis using the Steel-Dwass method shows that Instructor A's students' responses are significantly different from the responses of Instructor B's students ($p \leq 0.0001$) and significantly different from Instructor C's students ($p \leq 0.0001$). However, the responses from Instructor B's and Instructor C's are not significantly different from each other ($p \approx 0.1752$).

Is there a difference between how students interpret a specific value of probability when accounting for the instructor?

Much like the prior question, a two-way contingency table provides insight into answering the question about the difference in how students interpret a given probability value in relation to the students' instructor. Notice in Table 3 that the vast majority of students who interpreted 3/36 as a percent of time have Instructor A and two-thirds of Instructor A's students gave this type of interpretation. None of Instructor C's students and only 1 of Instructor B's students gave a response that fell into this category. Given that the majority of Instructor B's and Instructor C's students appeared to have a meaning for probability that was circular (see Table 2), the spread of their students' interpretations is not surprising.

Table 3. Students' Responses to Question 2 by Students' Instructor

	Percent of the Time	Classical	Fixed Number of Rolls	Chance	Other	total
Instructor A	18	3	3	1	1	26
Instructor B	1	10	27	12	11	61
Instructor C	0	2	15	6	1	24
total	19	15	45	19	13	111

I conducted a second Kruskal-Wallis test (with $\alpha = 0.05$) to test the difference between the students' responses in relation to instructor. The test statistic has a value of 32.2145. Under a χ_2^2 distribution, the approximate probability that we observe the differences we did or one greater is $p \leq 0.0001$. Post-hoc analysis using the Steel-Dwass method indicates that Instructor A's students' responses are significantly different from those of Instructor B's students ($p \leq 0.0001$) and Instructor C's students ($p \leq 0.0001$). The responses of Instructor B's students are not significantly different from Instructor C's students ($p \approx 0.9801$).

Discussion

The vast majority (78.1%) of students describe thinking about probability in circular ways, with roughly 13% (of 114) describing probability as being about the long-run relative frequency of some event (given a stochastic process). Similarly, a majority of students

(40.5%) interpreted the given probability value, 3/36, as being about a fixed number of rolls (of dice) and observing exactly 3 outcomes that were the event of interest. Even after students had received instruction on probability, there was still variation in the responses that students gave. This being said, there appeared to be clear distinctions between the majority response for each instructor's students. In the case of Instructor A, the majority response for probability (both in general and for interpreting) students gave is consistent with thinking about probability as the long-run relative frequency of an outcome of some repeatable process. For Instructor B's and Instructor C's students, the dominant responses for probability seemed to focus on a circular word-exchange and a fixed number of trials. The coded responses for the three instructors appear in Table 4. Like with the students, the teachers' responses offer insight into the meaning for probability that each teacher has. While further investigation into each teacher's actual meanings for probability is necessary, their responses appear to match up with the prevailing responses of their students. This seems logical, given that a teacher's mathematical meanings serve as one of the key components of how that teacher teaches (Thompson, 2013).

Table 4. Instructor's Responses to Question 1 and Question 2.

	Response to Question 1 (probability in general)	Response to Question 2 (interpret 3/36)
Instructor A	L.R.R.F.	Percent of the Time
Instructor B	Circular	Classical
Instructor C	Circular	Classical

A limitation to this study is that responses to two questions do not necessarily provide enough information to confidently describe an individual's meanings for a mathematical topic. While some informal discussions with students have taken place, interviews with more students will help to support the claims about the possible meanings students might operate with when they give particular responses. Additionally, given that this was an observational study, we cannot definitively say that Instructor A is the cause for stark differences between the three sets of students' responses. However, given that Instructor A made the decision to follow a curriculum centered on assisting students in developing productive ways of thinking, there is evidence of a strong causal relationship. Further research could substantiate this claim.

This study serves as but a first step in examining how undergraduate students think about probability after receiving instruction. While only drawing upon data from two questions, the inclusion of similar questions can help to refine items that serve as a means to measure a progress variable for probability. Progress variables represent "(a) the developmental structures underlying a metric for measuring student achievement and growth, (b) a criterion-reference context for diagnosing student needs, and (c) a common basis for interpretation of student responses to assessment tasks" (Kennedy & Wilson, 2007, pp. 3–4). Establishing a progress variable for probability along with items that measure such a variable has the potential to change how we teach probability at all levels. Additionally, a progress variable for probability is of use for other areas of statistics education research including students' notions of *p*-values, hypothesis testing, and distributions of random variables.

The present study into how a set of undergraduates thought about probability has shown that there are some stark differences between different sections of the same course. Sadly, the dominant meanings that these students appear to use for probability are circular and calculationally oriented. One section of the course, which used a "reformed" curriculum, does have a number of students who appear to have a highly productive meaning for probability. Further work needs to be done in order to help more students develop a rich and deep meaning for probability that is coherent and does work for the students in statistics.

References

Cobb, G. W., & Moore, D. S. (1997). Mathematics, Statistics, and Teaching. *The American Mathematical Monthly*, *104*(9), 801–823.

Kennedy, C. A., & Wilson, M. (2007). Using Progress Variables to Interpret Student Achievement and Progress (BEAR Technical Report No. 2006-12-01). *Retrieved from the Berkeley Evaluation & Assessment Research Center Website: Http://bearcenter. Berkeley. edu/publications/UsingProgVars_Kennedy_Wilson. Pdf*. Retrieved from http://drupal.bear.berkeley.edu/sites/default/files/Kennedy_Wilson2007.pdf

Konold, C. (1989). Informal Conceptions of Probability. *Cognition and Instruction*, *6*, 59–98. http://doi.org/10.1207/s1532690xci0601_3

Liu, Y., & Thompson, P. W. (2002). Randomness: Rethinking the foundation of probability. In *Proceedings of the Twenty Fourth Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education, Athens, GA* (Vol. 3, pp. 1331–1334). Retrieved from http://www.patthompson.net/PDFversions/2002Randomness.pdf

Moore, D. S., McCabe, G. P., & Craig, B. A. (2012). *Introduction to the Practice of Statistics* (7th ed.). New York, NY: W. H. Freeman.

Samuels, M. L. (2012). *Statistics for the life sciences* (4th ed). Boston: Pearson Education.

Strauss, A. L., & Corbin, J. M. (1990). *Basics of qualitative research: grounded theory procedures and techniques*. Newbury Park, Calif.: Sage Publications.

Thompson, P. W. (2013). In the Absence of Meaning... In K. Leatham, *Vital Directions for Research in Mathematics Education* (pp. 57–93). New York, NY: Springer. Retrieved from http://bit.ly/Ztg3Hm

Thompson, P. W. (2015). Researching mathematical meanings for teaching. In L. D. English & D. Kirshner (Eds.), *Third Handbook of International Research in Mathematics Education* (pp. 435–461). New York: Taylor and Francis.