

FLIPPING BETWEEN LANGUAGES? AN EXPLORATORY ANALYSIS OF THE USAGE BY SPANISH-SPEAKING ENGLISH LANGUAGE LEARNER TERTIARY STUDENTS OF A BILINGUAL PROBABILITY APPLET

LAWRENCE M. LESSER
The University of Texas at El Paso
Lesser@utep.edu

AMY E. WAGLER
The University of Texas at El Paso
awagler2@utep.edu

BERENICE SALAZAR
Northwest Early College High School
bsalazar@canutillo-isd.org

ABSTRACT

English language learners (ELLs) are a rapidly growing part of the student population in many countries. Studies on resources for language learners—especially Spanish-speaking ELLs—have focused on areas such as reading, writing, and mathematics, but not introductory probability and statistics. Semi-structured qualitative interviews investigated how a purposeful sample of six (Spanish-speaking) ELLs experienced a bilingual coin-flipping simulation applet (NLVM, 2015) and how students might use such resource to confront content misconceptions and language misunderstandings related to probability concepts covered in college introductory statistics courses. We discuss findings, limitations, directions for future research, and implications for teaching, such as handling the phrases “in the long run” and “longest run”.

Keywords: *Statistics education research; Representativeness heuristic; Simulation; Equiprobability bias; Resources*

1. INTRODUCTION

English language learners (ELLs) are a growing K-16 student population in the United States (Cardenas-Hagan, 2010; Lesser, Wagler, Esquinca, & Valenzuela, 2013). By ELL, we mean students who experience “enough limitations that he or she cannot fully participate in mainstream English instruction” (Goldenberg, 2008, p. 10), which includes those beginning to learn English who could benefit from language support and those who are proficient in English but may need additional assistance in social or academic situations (Hoffstetter, 2003). Outside the United States, the academic language may not be English, and correspondingly, we use the term language learner (LL) to indicate a student whose first language is not the dominant academic language in this setting. In the United States, the population of ELLs are linguistically, culturally, and socioeconomically diverse. In Texas, a state with a high population of ELLs, 44% of Hispanic ELLs plan to attend college

(Flores, Batalova, & Fix, 2012) or two-year post-secondary programs with even higher rates of college attendance for other ELL groups with Asian, Black or White ethnicities. Full participation of ELLs into the learning community is essential, not only with regard to issues of equity, but also to recognize the assets ELLs bring to a classroom environment.

This study investigates how a particular bilingual web-based applet may be used by ELLs learning probability concepts from an introductory statistics course. The few studies about language issues in learning probability (e.g., Green, 1984) generally do not involve students learning in a second language. There have been a few studies about second language learners learning probability (Kazima, 2006; Phillips & Wright, 1977), but these did not involve Spanish. Though our focus is on ELLs, we note that the language of probability can also be challenging for many non-ELLs (Lavy & Mashiach-Eizenberg, 2009; Lesser & Winsor, 2009). For example, there is evidence that phrases such as “at least” and “at most” are misunderstood even among English-only students (Nolan, 2002).

2. LITERATURE REVIEW

2.1. ELLS IN THE PROBABILITY CLASSROOM

Research studies have been conducted about resources for ELLs in subjects such as reading, writing, and mathematics, but little research has been conducted on ELLs in probability and statistics (Lesser & Winsor, 2009). ELLs learning probability are faced not only with navigating through new material, but also with learning how to succeed in a course in a different language. Lesser and Winsor illustrated how some ELLs have difficulties understanding the language and concepts utilized in introductory statistics classes and many of these findings are also applicable to learning probability concepts.

Instructional practices targeting ELLs should address students’ linguistic and sociocultural needs without altering the construct of the material being assessed (Willner, Rivera, & Acosta, 2009). When doing assessments, ELLs acquire language processing skills and knowledge of academic English to focus on the assignments and, consequently, some students may take longer to process the language of the assessment and may encounter difficulties due to unfamiliar language, cultural references, or format (Willner et al., 2009). For example, Lesser and Winsor (2009) reported a student’s confusion from a context-rich exercise about correlation because the term “ski resort” was unfamiliar in her high-poverty urban city in a desert region. Similarly, in a probability study, students were provided an example where snowfall could be modelled with a Poisson distribution. However, students unfamiliar with snowfall may assume snow falls “uniformly” and be misled about the true nature of a Poisson distribution and the patterns of snowfall it could model (Borovcnik & Bentz, 1991). This underscores the need to address sociocultural factors in teaching and assessing students in the probability classroom, with a focus on the issues faced by language-minority students.

Appropriate resources improve the performance of ELLs but leave the performance of non-English language learners untouched (Hofstetter, 2003). In fact, experts in bilingual education generally assert that resources and teaching methods directed towards ELLs in academic settings benefit all students who struggle with language use in this domain (Harper & de Jong, 2004). Common supports include working in small groups, giving students more time to think, offering a presentation format which includes translations and linguistic modifications, allowing for broader forms of response, and offering tools such as bilingual word lists, dictionaries or glossaries (e.g., Dragt, 2015), and applets (Abedi & Lord, 2001; Garrison & Mora, 1999; Hofstetter, 2003).

2.2. PROBABILITY MISCONCEPTIONS

Wilensky (1995) laments the practice of “safe probability” in which secondary or tertiary students use rote formulas in formal exercises without attending to the meanings of underlying concepts or questions such as “How can something be both random and structured?” (p. 254). Indeed “random processes lie at the very heart of statistics and probability” and highlight what Jones, Langrall, and Mooney (2007) call probability’s “complementary role in informing statistics and its special place in describing and quantifying the numerous chance phenomena that operate in our world” (p. 910). Attention to the role of language in probability instruction will further deepen the conceptual formation of students as well as clarify the thought processes for ELLs and non-ELLs.

Research suggests that many student difficulties in learning probability are rooted in misconceptions (Cox & Mouw, 1992; Shaughnessy, 1977). Table 1 lists examples of well-known probability misconceptions expressed using the context of coin tossing. We limit ourselves to a single context—and this particular context—because it “allows us to explore student beliefs within a particular context and enables a comparison of student thinking about different probabilistic constructs related to this single context” (Rubel, 2007, p. 536). These misconceptions are also consistent with those in Amir and Williams (1999).

Table 1. Examples of probability misconceptions in a coin tossing context

MISCONCEPTION	EXAMPLE
Equiprobability Bias	A person believes that, say, “exactly three heads” or “exactly one head” are equally likely outcomes for a 3-flip sequence (adapting dice example of Lecoutre, 1992)
Gambler’s Fallacy	A person believes that after nine heads have appeared that the tenth toss is more likely to be tails (Cox & Mouw, 1992; Shaughnessy, 1977)
Law of Small Numbers	A person expects even short runs of coin flips to reflect the fairness of a coin (Kahneman & Tversky, 1972; Shaughnessy, 1977)
Representativeness Heuristic	A person believes that a sequence of coin tossing with a very long streak of heads or with a well-ordered pattern such as THTHTHH is not representative of a random process of coin tossing (Cox & Mouw, 1992; Kahneman & Tversky, 1972)

Having students commit to and then assess their predictions can be used to uncover their prior knowledge, schemas, misconceptions, and intuitions in probability. It provides the opportunity for students to experience cognitive conflict, and thus become aware of their misconceptions and gain the motivation to address them (Lim, Buendia, Kim, Cordero, & Kasmer, 2010; Shaughnessy, 1977). Another method to help students confront misconceptions is using computer-assisted learning environments (Garfield & Ben-Zvi, 2007). Gürbüz and Birgin (2012) conducted quantitative research studying the effect of computer-assisted teaching on remedying probability misconceptions. The results showed that both groups (experimental and control) had gains in student learning, but those using computer assisted teaching had greater gains. The Meletiou-Mavrotheris, Lee, and Fouladi (2007) qualitative research on computer-assisted teaching showed that the technology-based groups enjoyed learning probability but conceptual understanding did not improve.

2.3. THE CONTEXT OF COIN FLIPPING

Coin flipping is common to many cultures for making decisions or resolving disputes. Sporting events such as American professional and college football games begin with the on-field coin flip ceremony to determine who gets the ball first. In addition to such real world connections, a coin flip is a Bernoulli trial and a (fair) coin flip forms the simplest example of an equiprobable sample space. Instructors and textbooks commonly use equiprobable manipulatives for probability explorations: dice, cards, coins, and spinners. Of these, a coin is simplest because a flip has (ignoring the unrealistic possibility the coin lands on its edge) only two possibilities named *heads* and *tails* in the US, but having other names in other countries (Lesser et al., 2013).

While offering a sample space less complex than a deck of playing cards (which Knapp, 1996, used to cover an entire introductory statistics course), coin tossing can describe almost all major probability misconceptions (see Table 1) and be a vehicle to illustrate random walks and many other diverse statistical concepts. Consortium for Mathematics and its Applications (COMAP) (2013) starts its probability chapter with coin tossing illustrating the long-run frequency interpretation of probability, then uses coin tossing to illustrate sample spaces, and ends the chapter with an application of coin tossing to the Central Limit Theorem. Also, see Lesser and Glickman (2009) for discussion of how flipping a coin (whether or not the coin is “rigged”) has been used pedagogically to explore concepts of hypothesis testing. Coin tossing has been used as a conceptual benchmark for randomness, to investigate whether some basketball players are “streak shooters” (Gilovich, Vallone, & Tversky, 1985): “Consider a professional basketball player who makes 50% of his shots....The player’s performance, then, can be compared to a sequence of hits and misses generated by tossing a coin” (p. 296). Another application is simulating whether people correctly identify which of two cups has bottled water (Franklin et al., 2015). Research studies using coin tossing as a context include Falk and Lann (2015), Rubel (2007), Sedlmeier (1998), and Watson and English (2015).

A statistics joke (<https://www.causeweb.org/resources/fun/db.php?id=69>) compares coin flips to guessing on a true-false test. Students are often captivated by the challenge of distinguishing which of two long sequences of coin flips is fabricated and which is obtained from actual flips (e.g., Gnanadesikan, Scheaffer, Watkins, & Witmer, 1997). Students are also interested in when coin tosses are fair but rarely appreciate what that entails (Gelman & Nolan, 2002; Lipkin, 2003). Some activities investigate whether heads and tails are equally likely when a coin is spun (see “Spinning Pennies” in Scheaffer, Gnanadesikan, Watkins, & Witmer, 1996) or even stood on edge on a flat surface that is then hit so the coin falls over (see “Coins on Edge” also in Scheaffer et al., 1996).

It turns out that coin tossing has potential for bias. Klarreich (2004) notes that a coin is more likely to land on its initial face, a tendency that is increased when the flip produces a wobble closer to being a flat toss, a situation hard to distinguish visually. Because someone could manipulate a manual coin flip, there is motivation to use technology to generate virtual flips. Another advantage of virtual flipping is its ease in simulating many trials. Wilensky (1995) notes that everyday experience (e.g., a coin flip) may not help build intuition about the large numbers of instances probability is founded on. Students would not have the patience (of Karl Pearson, John Kerrich, or Count Buffon) to do thousands of tosses manually. The discussion of the fifth recommendation of the college *Guidelines for Assessment and Instruction in Statistics Education* (GAISE College Report ASA Revision Committee, 2016) notes “Technology tools should also be used to help students visualize concepts and develop an understanding of abstract ideas by simulations” (p. 20), and one tool GAISE lists is web applets.

The particular applet in our study is the coin tossing Java-based applet freely available at Utah State University's NSF-supported project National Library of Virtual Manipulatives (NLVM) (2015), in the Data Analysis & Probability category. The applet allows the user to generate and view a coin toss sequence that stops after a given number of flips or after a given "longest run of heads" first occurs. The applet allows the user to change the probability of heads from its default value of $\frac{1}{2}$. As the flipping proceeds, the user sees the proportion, number, and longest run of heads (and these counterparts for tails) update in real time.

2.4. LANGUAGE IN PROBABILITY

Students often find statistics and probability concepts difficult to understand in both formal and everyday contexts, partially due to lexical ambiguity (Garfield & Ben-Zvi, 2007; Kaplan, Fisher, & Rogness, 2009, 2010; Kaplan, Rogness, & Fisher, 2012, 2014; Lavy & Mashlach-Eizenberg, 2009; Lesser & Winsor, 2009; Sharma, 2006). We note that probability learning is greatly affected by matters of language, often in ways that are more difficult for ELLs. For example, many probability statements use words such as "at least", "at most", "fewer than", etc. Students accustomed to a reflexive "key word" association (e.g., treating "at least" as "less than") may make errors interpreting a question such as "What is the probability that Bob will have at least 8 successes in 10 attempts?" Other language conventions that must be learned in probability include reading $P(A|B)$ as "probability of A given B" or building arguments with statements such as "Let X be normally distributed." Additionally, students may encounter probability language using words from an earlier math course in a very different way. It is common that a 50% probability might be described as "the odds are even", a phrase unrelated to even or odd numbers. Furthermore, the word *odds* has been misused in prominent real-world settings (e.g., Lesser, 2013). Decoding probabilities requires more sophisticated language skills. For example, students may not realize how small wording changes produce very different probability events (e.g., $P(\text{male smoker})$ vs. $P(\text{smoker being male})$ vs. $P(\text{male being a smoker})$; $P(\text{all die rolls are not fives})$ vs. $P(\text{not all die rolls are fives})$; $P(\text{someone who tests positive has cancer})$ vs. $P(\text{someone with cancer tests positive})$). Decoding such differences may be more difficult for ELLs. Resources such as <http://www.statlit.org/GC/> let students practice distinguishing, e.g., "10% of men who are black are runners who smoke" from "10% of men who smoke are runners and are black" (Burnham & Schield, 2005).

Kaplan, Fisher, and Rogness (2009, 2010) and Kaplan, Rogness, and Fisher (2012, 2014) considered the role of lexical ambiguities in the statistics classroom. Their findings with regard to student use and understanding of the word "random" are applicable when teaching probability, but their research did not explicitly address ELLs. Another research study, focused on how ELLs encounter language when learning introductory-level probability and statistics, consisted of semi-structured interviews with scenarios and open-ended questions that gave the students the opportunity to describe what statistical terminology meant in their own words (Lesser & Winsor, 2009). This study presented instances where the translation of words would make little difference especially when the meaning of a term was influenced more by the everyday language. Another observation of Lesser and Winsor (2009) was that deficiencies in students' native language could delay students' understanding of probability or statistics concepts and that the role of context is more important for ELLs learning statistics than learning mathematics because statistics inherently involves more context (Lesser & Winsor, 2009). These research findings are likely to be applicable to the teaching of probability as many statistical concepts are rooted in probability concepts. Lesser, Wagler, Esquinca, and Valenzuela (2013) focused on

learning practices of ELLs when learning statistics and discussed many distinctions between ELL and non-ELL learning practices. For example, it was found that ELLs are more likely to feign understanding in the classroom than non-ELLs. These findings collectively inform the research questions for this study, such as the use of the word *appear* in Research Question 2.

3. METHOD

3.1. PURPOSE AND RESEARCH QUESTIONS

The purpose of the study is to investigate how the bilingual coin-flipping simulation applet might help ELLs learn probability concepts in their introductory statistics course.

1. What is the nature of how Spanish-speaking ELLs use a bilingual applet when learning probability?
2. When does it appear that language plays a factor when Spanish-speaking ELLs explore probability with the applet?

3.2. PARTICIPANTS

Participants were recruited through flyers given to five sections of an introductory statistics course at a moderately large doctoral/research university in a large city in the southwestern United States. Roughly three-quarters of the student population is Hispanic. The flyer stated that the research sought Spanish-speaking ELLs for an interview study that could help the researchers better understand how to teach probability and statistics more effectively. Though only females contacted the (female) interviewer to schedule an interview, the interviewees were not unrepresentative of the heavily female enrollment in that course, which was 76.8% female that semester. The six females were at least 18 years old, currently taking an introductory statistics course in that (January–May 2012) semester, and self-identified as ELLs.

3.3. RESEARCH PROTOCOL

The students and interviewer met in an instructor office in the campus mathematics department building. Before the interview, the student read the consent form and had a chance to ask questions before signing. The form disclosed that the interview would be video recorded, take approximately 25-40 minutes, and have no actual names or identifying footage ever viewed or shared beyond the researchers.

Video recording was necessary to record not just what was said but also how the interviewee was interacting with the NLVM applet on the computer screen and to record any non-verbal behavior and gestures. Participants were informed that their name would not be used during the interview, and that a pseudonym would be used in interview transcripts.

Interviews were conducted in March-April 2012 using a protocol (see Appendix 1) with three parts: pre-applet questions (PreQ) to assess student understanding of probability concepts before exposure to the applet, task-based questions (AQ) to explore with the applet, and post-applet questions (PostQ) to assess new understanding. Laurie Rubel (2007) used coin flipping as the main context in her research-based inventory given to 173 grades 5-11 students to assess their probabilistic reasoning abilities. Her work, in turn, informed some of our items: our PreQ5 relates to her “two-coin item”, which invokes the distinction between sample space and sample set (Chernoff & Zazkis, 2011), and our PreQ20 relates

to her “four heads item”. Our survey also had many items without counterpart in Rubel (2007), such as items on numbers and proportions of heads for higher numbers of tosses.

All probability questions used the context of coin tossing and students were reminded (to reinforce the statement on the consent form that the interviewer was bilingual) that they were welcome to ask for any question to be stated in Spanish if it was not clear in English, and that they could give an answer in Spanish if they preferred. The interviewer, who is proficient in both languages, had already translated the questions from English to Spanish using the Spanish version of the applet as reference. After collaboration with other graduate students who are also proficient in both languages, the interviewer finalized the Spanish version of the questions. Just before turning on the video camera to start the interview, the interviewer provided a word list (of English-Spanish translations of probability terms common in an introductory statistics course, not including some specific terms, such as “in the long run” used in the interview— see Appendix 2) and said it could be consulted at any time (no one used it and only one person even looked at it).

During the applet questions, the interviewer showed the students the NLVM coin-tossing simulation applet, including how to toggle between English and Spanish (Figure 1 shows screenshots in those languages). The applet, initially displaying English, could also toggle to French or Chinese versions, but the study focused on the English and the Spanish versions. The students were given the minimum guidance to explore the applet enough to have a basis for answering task-specific questions and deciding if their observations supported or conflicted with their predictions. Finally, during the post-applet questions, students were asked about their experience using the applet.

An overview of the project timeline appears in Table 2. Transcribing the interviews (included in Salazar, 2012) was followed by a modified grounded theory approach (Glaser & Strauss, 1967) in which codes were allowed to emerge inductively, though informed by a few large concepts from the research questions. During this stage of analysis, the first and third authors facilitated a peer debriefing session on October 24, 2012 with a group of 13 mathematics education graduate student researchers focusing on transcript excerpts where interviewees struggled to answer (PreQ9, PreQ10, PreQ17). A second peer debriefing session attended by all three authors was held on April 22, 2015 with six graduate students (a different group from the first debriefing session) who, as an assignment in their course on qualitative methods in mathematics education research, did coding and cross-case analysis on all transcripts. Insights from both sessions informed the Findings and Discussion sections.

Table 2. Research timeline

STEP	TIME
Study Design; IRB process	October 2011-February 2012
Recruitment	February-March 2012
Interviews	March-April 2012
Interview transcription	June-July 2012
Analysis	August-December 2012
Peer debriefing	October 15, 2012; April 22, 2015
Final refinements by authors	April-May 2015

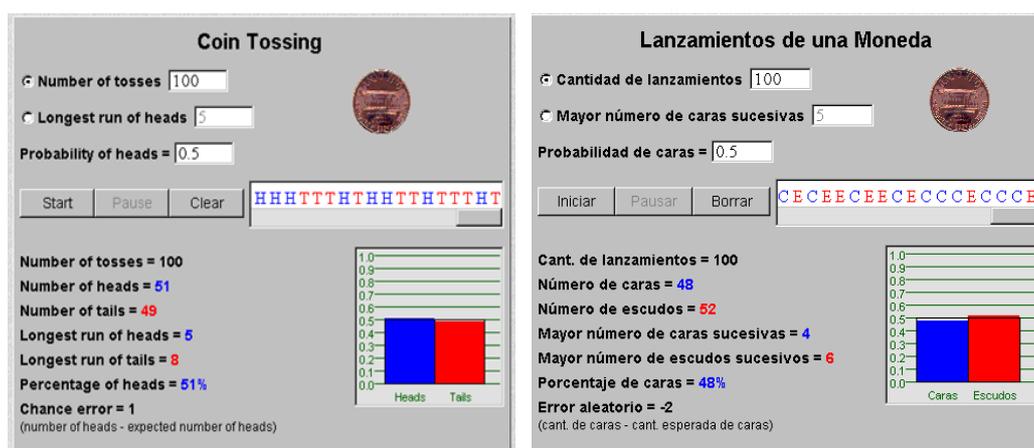


Figure 1. NLVM coin-tossing applet screenshots in English (left) and Spanish (right)

4. FINDINGS

4.1. FINDINGS FOR RESEARCH QUESTION 1: THE NATURE OF HOW ELLS USE A BILINGUAL APPLET

To answer Research Question 1, the third author (denoted B in transcript excerpts) compared the responses the interview participants (P1 through P6) gave before they had any contact with the applet to the responses they gave after being exposed to the applet.

Applet exploration uncovers nature of probability language misconceptions PreQ17 asked interviewees about the meaning of “longest run” (language the applet uses). Five of the six interviewees had difficulty explaining the meaning of the phrase. P1 provides an example of this confusion.

- B: Okay, okay now, in your own words, what does it mean, what does ‘longest run’ mean to you?
 P1: Longest what?
 B: I’m going to repeat it again, in your own words, what does ‘longest run’ mean to you?
 P1: Longest run, hmmm [long pause] longest run, like [short pause] the more the most hmmm the fastest to flip the coin, like many times but so fast [nervous laugh]

The next pre-applet question (PreQ18) asked about “in the long run”.

- B: In your own words, what does ‘in the long run’ mean to you?
 P1: Long run, okay in Spanish [nervous laugh].
 B: En tus propias palabras, ¿Que significa ‘a largo plazo’?
 P1: Oh, tener determinado un plazo pero muy largo [nervous laugh] [To have determined a very long term].

During the applet portion of the interview (AQ5), P1 predicted the longest run would be 100 heads, which would conflict with her applet results.

- B: Okay now I’m going to the next question. Okay. You had predicted in the previous question that there would be a longest run of 100 in a 100-flip sequence, remember?
 P1: Uh huh.

- B: Okay, do you want, do your results seem to confirm or conflict with your prediction? I don't know if you notice in here it actually tells you the longest run [interviewer gets the mouse and points to the longest run]
- P1: Oh and it's four,
- B: [after running applet again] Now we got the longest run is 6, right? Okay so, does this want to make your, you had predicted in a previous question that you would have a longest run of 100 in 100 tosses. Do your results seem to confirm or conflict with your prediction?
- P1: Conflict.

Before exposure to the applet, P3 said that the longest run in a 100-flip sequence would be two. After running the applet (and encountering streaks with greater length than students generally expect from a random sequence) and being reminded that her previous prediction was two, P3 said that the results conflicted with her prediction.

Interviewees were also asked in AQ10a "If we change the probability of heads from 0.5 to 0.7, what change, if any, do you think there would be on the number of heads? What change, if any, do you think there would be on the longest run of heads?" Once the interviewees made their predictions, they were asked to change the probability of heads from 0.5 to 0.7 and to run the applet in the mode where the coin would flip 100 times. All interviewees again experienced confusion with the phrase "longest run", but use of the applet again clarified the meaning of this phrase (in conjunction with a Spanish translation of the phrase for five of six interviewees). Before applet usage, P5 was unable to provide a reasoned response to the question about the longest run of heads due to an apparent confusion of the phrase with longest possible run (where P5 said the longest run would be 100). This exchange occurred during AQ10:

- P5: There would be no change 'cause we are just putting 100 tosses, right?
- ...
- P5: No change either 'cause you don't.
- P5: [after running applet again] Of course there would be more heads because you are changing the probability of having heads to 70% from 50% so it would be more. I think I would change my answer from before.

Similarly, P6 changed her response to AQ10b. This respondent believes the longest run will change, but cannot describe the nature of that change because she mistakenly believes the "longest run" is unaffected by the probability of heads.

- P6: Sí, sí va a cambiar [yes, yes it's going to change].

However, when asked how a change in proportion of heads affects the longest run,

- P6: Ahmmm, no precisamente [not precisely].
- ...
- P6: [after running the applet] Es diferente a mi predicción [it's different from my prediction], sí porque yo dije que no importaría el número de caras y si importa [yes because I said that the number of heads did not matter and it does].

After running the applet, P5 and P6 corrected their predictions. Table 3 summarizes these responses. We note that some interviewees may not have been initially able to interpret the phrase "longest run" correctly. For example, P1 experienced confusion with this phrase at the outset of the interview where she mistook the phrase to be the "fastest flip of the coin" (see first interview excerpt in this section regarding PreQ17). However,

even after obtaining the Spanish translation of this phrase, she did not clearly have the phrase well-defined. This excerpt involves AQ10b.

- B: Do you want me to read it in Spanish? Maybe it makes more sense to you?
 P1: Okay.
 B: It says, en tus propias palabras, ¿Que significa la mayor racha o el mayor número de caras sucesivas?
 P1: Mm-hum, osea, el lo las caras que pasaron más rápido, no que se repitieron [that is, the heads that passed faster, not that repeated]

Table 3. Responses of interviewees for questions about longest run

	AQ10a: Change in number of heads	AQ10b: Change in longest run of heads	Does your result seem to confirm or conflict with your prediction?
P1	We are going to have more faces? [heads]	3	[not asked]
P2	They're probably equal	I guess there would be more heads because there's more probability	Confirm
P3	The heads will be more	7	I think it's very close
P4	There would probably be more heads	I'm not sure, it would show more caras sucesivas [successive heads]	Conflict. I thought it would take less time to get more
P5	There would be no change	No change either 'cause you don't know	Of course there would be more heads because you are changing
P6	Sí, sí va a cambiar [yes, yes, it's going to change]	No precisamente [not precisely]	Es diferente a mi predicción [it is different from my prediction]

She improves her response (“not that repeated”), but this clarification does not lead to a correct prediction. Once she views the applet, she comes around to a clearer idea of the “longest run” both with regard to language as well as concept. Some misconceptions may be attributable to language, rather than probability misconceptions. However, for P1, there appears to be some confusion due to words as well as probability misconceptions. Regardless, the probability applet appears to address these initial misconceptions. Similarly, P2 experienced confusion about the phrase “in the long run” in PreQ18.

- B: Okay, now in your own words, what does ‘in the long run’ mean to you?
 P2: Does it have to, like, relate to math?
 B: Yeah
 P2: In the long run, I don't know.
 B: Or in everyday life.
 P2: In the long run, like, isn't it's [long pause] like, ahmm in the long run, I know what it means but I don't know how to explain it to you. It's kind of self-explanatory, like you know something in the long run, it's like it happen, I don't know how to explain it, but I know what it means.
 B: Okay, now look at sequence A [spontaneously points to example in recently discussed PreQ14 to see if P2's conception becomes more clear when applied to a concrete example].
 P2: Mm-hmm.

- B: And what is the longest run?
 P2: Hmm, the one that has the most T's.
 B: Okay, which is?
 P2: Well there's 2, like this one... 'cause it has the most T's.
 B: The most T's. How many T's in that?
 P2: 4.

ELLs report language support from applet To assess how Spanish-speaking ELLs say that they benefit from using a computer applet simulation in English and Spanish, interviewees were asked whether they thought using applets like the one they experienced would be helpful to them or other students. All interviewees reported that these applets are beneficial, but reasons varied. P5 reports that seeing the equivalent words in the English and Spanish versions of the applet to be helpful.

- P5: Well as far as if you are having problems with the wording of certain terms, it would be easier if you have the ability to change language, 'cause maybe you learned in Spanish like I did and then can you switch it to English, it helps.

ELLs report visualization support from applet Others perceived the usefulness, not as a language resource, but as a visualization tool.

- P2: I think it would be [useful] because to me I'm more of a visual person so that's why I was having a hard time when you were asking me questions, because I was trying to visualize it. So I think this would be better because you can see what's happening.

P3 and P6 agreed that the applet is helpful because it illustrated coin tossing.

- P3: It would be helpful because in one [version of the coin flipping] it is very clear with the toss and two it's showing you a lot of the different types of, ah, data for the coin toss like the actually number that it was tossed [sic] that you get to play with it and figure out whatever kind of information you need and, and it has a lot of information just for that.
 P6: Yes, it would because it explains you everything like the number of heads, hmmm, ósea te dice cuantas veces lo lanzaste o sin la necesidad de que tu lo estés haciendo [it tells you how many times you flipped without having you do it].

4.2. FINDINGS FOR RESEARCH QUESTION 2: WHEN IS LANGUAGE A FACTOR IN APPLET USE?

To assess Research Question 2, the researchers created cross-case tables to look for patterns in language used. The interviewees were shown both English and Spanish versions of the applet and provided the opportunity to choose what version they were more comfortable working with and whether to toggle between English and Spanish. Five out of the six interviewees decided to work with the Spanish version of the applet throughout the interview (P2 chose the English version).

When referring to a coin During the PreQs, the interviewees were asked what they call the sides of a coin in English. P1 and P5 referred to the sides of a coin as faces (English translation of "caras") and the rest of the interviewees said heads and tails. Then they were asked what they called the sides of a coin in Spanish with varying answers: P1 and P6 said "cara o cello", P5 and P6 said "águila o cello", P3 said she didn't know, P2 said "cara o letra", and P4 said "cabeza". Interviewees that correctly named the sides of the coin in Spanish (P1, P5, P6) spoke in Spanish more throughout the interview than P2 and P4 who

provided incorrect Spanish names for the sides of a coin. P3 said she did not know what to call the sides of a coin in Spanish and spoke English throughout the interview.

Flipping between English and Spanish registers There is evidence that three interviewees were thinking in one language while talking in the other. P4 said *cabeza* (which is head in English) when she was asked what she called the sides of a coin in Spanish. Similarly, P1 said “faces” when she was asked what she called the sides of a coin in English. In many Spanish-speaking countries, the sides of a coin are called *cara y cello* or *cara y águila* (face and seal; face and eagle). Also in Spanish, *caras* is used in *las caras de las monedas* which means “sides of the coin” but can translate literally as “faces of the coin”. For example in PreQ15 (Appendix 1), when given two sequences and asked which sequence had higher probability or were they equally likely, P1 responded, “Hmmm, in the B is more equally of the faces than this one [pointing to sequence A] because this one have more Ts.” In the 2012 peer debriefing, the researchers and the graduate students agreed that in this instance P1 was referring to heads. Another instance when P1 said faces was when the interviewer asked “What would you call the sides of the coin in English?” and she responded “Faces?” The interviewer repeated the question and P1 said “Face, hmmm, I don’t know.” In this case, the researchers believe interviewee P1 said faces to refer to the sides of the coin. Similarly, when P5 was asked “What would you call the sides of a coin in English?” she said “The faces, the face of the side, I don’t know,” also referring to the sides of the coin.

There were several occasions when interviewees answered spontaneously in Spanish a question that was asked in English.

- B: Now you had predicted in a previous question that there would be a longest run of 5 in 100 toss sequence, does the result seem to confirm or conflict with your prediction? Look at the longest run. How many does it say?
 P2: El mayor número de caras sucesivas nueve [the number of successive heads nine].

The interviewee may have answered in Spanish because she was viewing the Spanish version of the applet (Figure 1). P4 and P6 gave similar responses that were apparently tied to the language used in the applet.

There were other occasions when the question was given in Spanish because the interviewee requested it or because the interviewer felt the interviewee did not understand the question and asked the interviewees if they wanted to hear the question in Spanish. After hearing the question in Spanish, the interviewees sometimes answered in English. For example, when asked (PreQ9) if you flip 100 times, how different from 50 heads would you have to get to call that result surprising, P6 asked for a translation and then answered seventy. This provides an example of when a question in Spanish was in turn answered in Spanish, but without regard to use of the applet. When asked for a translation of PreQ17 by P5 and P6, the interviewer said:

- B: En tus propias palabras, ¿que significa la mayor racha o el mayor número de caras sucesivas? [providing requested translation of PreQ17 “In your own words, what does the longest run mean to you?”]

P6 responded with,

- P6: Cuando eh obtienes [short pause] ah bastante número de caras continuamente [when you obtain a number of heads continuously].

and P5 responded with,

P5: Las veces que más tuviste cello o eso significa [the times you had most tails or that's what it means].

Interaction between complexity of response and language of communication There were occasions when the interviewees were able to express themselves with more elaboration in Spanish than in English after using the applet. Generally, responses in Spanish were more elaborated and insightful than responses in English. When P5 answered in English, it was not very clear what she was trying to say, but after a long pause she answered in Spanish and her response was clearer and more elaborated for PreQ6.

B: How long do you think the longest run of either heads or tails would be?

P5: I can't just answer here, actually you don't know 'cause it would be a 50-50 chance.

...

P5: [When asked if the applet results confirmed or conflicted with her prediction] It's 'cause you can have more so there would be more flips, so I think it would be could have more.

...

P5: [long pause] Puedes tener más caras sucesivas si tienes más lanzamientos [you can have more successive heads if you have more flips].

In general, the bilingual applet appears to improve conceptual knowledge as well as clarify language use in probability. The extent to which the bilingual aspect of the applet assists with clarifying language is unclear because every student was an ELL. However, this section documents cases where the students relied upon the translation of the applet into Spanish in order to either understand the concept or to construct an answer.

5. DISCUSSION

5.1. SUMMARY

As with other student populations, Spanish-speaking ELLs struggle to learn probability concepts, and using applets such as the one used in this research study may be helpful. Beyond positive self-reported feedback by ELLs on the applet, the interviews showed that the applet may help ELLs by offering a vehicle to identify misconceptions and clarify the meaning of probability phrases. There were instances where the interviewees had made an erroneous interpretation or prediction but were able to improve it after the opportunity to run the applet. For example, three interviewees realized after using the applet that they had been misinterpreting the phrase “in the long run” and then improved their paraphrase. Similarly, students often believed the “longest run” was much shorter than what is typically obtained from a true random sequence. However, upon using the applet, students realized their understanding of this phrase was incorrect and made appropriate revisions. The bilingual applet also supports language “flipping”— e.g., movement between the English and Spanish registers. Support for language transfer, along with visualization tools, are commonly noted resources for teaching statistics to ELLs (Lesser & Winsor, 2009) and this study shows how a bilingual applet can play both roles. Additionally, Falk and Lann (2015) showed participants’ overreliance on the law of large numbers in coin tossing. The interview transcripts concerning the graphs confirm this research result and these results show how an applet can assist in correcting this mistake.

The results support the implicit conjecture in research question 2 that language indeed plays a part when using the applet. There were several occasions when a question asked in

English was answered in Spanish after using the Spanish mode of the applet. We also noted more elaborated and insightful responses when participants used the Spanish mode and were allowed to respond in Spanish, not unlike the findings of Garrison and Mora (1999). We note that Viera (2015) observed ELL hesitation to use the bilingual mode of the web-based educational adaptive environment ALEKS due to a desire to learn the concepts in English to be prepared for assessments written in English and for making connections to future material that would be taught in English. However, those same students did use the bilingual version of ALEKS when taking exams. Our interviews reveal that when the bilingual applet was utilized, it appeared to enhance conceptual knowledge and clarify some language confusion, which is consistent with the findings of Viera (2015). Finally, we noted during a 2015 peer debriefing session that our findings are consistent with those of Amir and Williams (1999) who studied how language informed “informal knowledge” of probability and identified many of the same misconceptions as outlined in Table 1.

Though the research questions reference how ELLs approach using bilingual applets and how language is a factor when using these applets, we are limited in what we can say about how students perceive the usefulness of the bilingual aspect of the applet. Recall in Section 4.2, that P6 appears to have experienced conceptual and language confusion with regard to the phrase “the longest run”. In the post applet questions, P6 was asked about the use of the bilingual applet (in questions PostQ1-PostQ3) in the following excerpt:

- B: How did you decide when to use the English version and when to use the Spanish version of the applet?
- P6: Because there are some words I do not understand in English.
- B: What difference between the two versions did you notice or experience?
- P6: Well in [my] personal opinion, the Spanish is easier because my, my first language is Spanish and I, I'm not sure in English so.
- B: Ok, do you think that using applets like this would or would not be helpful to you or other students?
- P6: Yes it would.
- B: Why?
- P6: Because it explains you everything like the number of heads, hmm, osea te dice cuantas veces lo lanzastes o sin la necesidad de que tu lo estes haciendo [it tells you how many times you flipped without having you do it].

This excerpt suggests that the bilingual applet clarifies both language and concepts. Used in conjunction with word lists (which 5 out of 6 participants agreed are helpful to bilingual students), bilingual applets appear to clarify the use of language and provide concrete representations of these probability terms in a meaningful and impactful manner.

5.2. LIMITATIONS

One limitation of this study is that after transcribing the videos, we were unable to clarify some things that were later deemed unclear. Though we had permission to contact the interviewees via email to ask for clarification, the interviewer was unable to get a response once the need was identified. Member checking would have been very useful, especially when analyzing the transcripts from the first interview. P1 used the word *face(s)* seven times throughout the interview, and clarification may have determined whether she was referring to *cara o cello* [heads or tails] or *las caras de la moneda* [the sides of the coin]. Another place where clarification was needed concerned PreQ16, for which P1 responded that sequence A was more real “because is more normal for the probability”. There were four other occasions where clarifications from P1 were needed and one occasion each for P2 and P4. Review of the transcripts during a 2015 peer debriefing

session revealed that the bilingual word list should include more task-specific words such as “absolute difference”, “made up”, “longest run”, “in the long run”, “streak”, “chance error”, and “prediction”. Finally, we note that interviewees gave minimal justification for their answers, so a pre-interview questionnaire (such as Rubel, 2007) could better assess prior understanding.

5.3. DIRECTIONS FOR FUTURE RESEARCH

This study reveals how a bilingual resource such as this applet may benefit ELLs in an introductory statistics class. Future studies could build upon these results to investigate which characteristics of bilingual resources are most beneficial and systematically study how students use these resources to help build their English language statistics register. Some unexpected results of this study include confusion with frequently-used phrases involving the word “run” and the potential of students to use their knowledge of Spanish to help resolve this confusion.

In particular, five of six interviewees had difficulty interpreting “longest run” in English and either needed to run the applet or ask for a Spanish translation. In fact, the Spanish version for the “longest run” (*el mayor número de caras sucesivas*) can be translated into English as “what is the largest number of successive [or consecutive] heads,” as noted during a 2012 peer debriefing session. Such paraphrases of “longest run” should be included in future studies that assess student understanding of alternative phrases so as to inform future instruction and curricula for *all* students, not just ELLs.

A similar research strategy could be employed to assess the effectiveness of using “in the long term” (the English translation of “en el largo plazo”) instead of the more idiomatic “in the long run.” Other stock phrases in statistics are likely to also prove difficult to students, and particularly ELLs, and use of bilingual resources can clarify language and reveal misconceptions rooted in language. During the 2015 peer debriefing, several research directions were identified. First, the use of virtual manipulatives to help students make the transition from concrete objects to pictorial representations should be investigated in probability learning with a focus on ELLs (Moyer, Niezgodá, & Stanley, 2005). Also, some interviewees admitted that they used the applet solely in Spanish, but didn’t ask for Spanish translations. Reasons for using the applet as a language support could be better understood in a study with more extensive and sustained interviewing and observation.

5.4. IMPLICATIONS FOR TEACHING

Language use Specific findings from the present study suggest that instructors should take particular care with the way several probability concepts are worded (whether in the context of instruction, assessment, or technology exploration), including “runs,” “in the long run,” and “longest run.”

These results show how applets may be effective in addressing word confusion for phrases such as “longest run.” For instruction, phrases that are difficult, ambiguous, or idiomatic should be explicitly discussed during instruction or augmented with a more descriptive term (e.g., “largest number of successive heads” or “largest number of consecutive heads”). The Spanish translation of “in the long run,” “en el largo plazo” [in the long term] is more descriptive than the standard English phrase and this is an excellent example where the Spanish language is a resource for students’ conceptual understanding of probability. We recommend that instructors use “long term” in addition to or instead of “long run”, and Sullivan (2010, p. 259) is an example that refers to the “long-term proportion” rather than probability in the long run. This phrase is important to the very

definition of probability, beyond any specific context. When having the class use software packages (e.g., Probability Explorer, Stohl 2002) that have a stick figure running as the icon to *run* the simulation, we recommend that instructors proactively and explicitly bring this to students' attention to help prevent confusion regarding these phrases involving the word *run*.

Beyond the more descriptive paraphrases of phrases involving "run" (run, long run, longest run) in our study, instructors can also offer a table (as does Sullivan, 2010, in a section on the binomial probability distribution) of common phrases (more than, no more than, at least, at most, etc.) in probability questions and their corresponding mathematical inequality symbol.

Applet use Instructors teaching in a computer-equipped classroom could demonstrate operations with the applet during class time and assign students applet-based tasks to complete for homework outside of class time (as one way to "flip" the classroom), or simply make students aware of the applet as one of many (bilingual) content-specific resources that can support learning in the course.

Instructors should consider allowing small group discussion of applet results in students' dominant languages. This can help students clarify the meaning of probability phrases such as *longest run* so that concepts and questions can be discussed in the students' native language before navigating the English statistical register. As suggested by the interviews, encouraging students to paraphrase in another language clarifies concepts so that they may then transfer this knowledge into English and more accurately communicate what they understand.

The findings of this study also support using this applet in conjunction with pre-applet predictions. This appears effective in creating cognitive dissonance whereby students must recognize conflict between their thinking and applet results, not unlike Shaughnessy (1977). Finally, these interviews make clear that ELLs "flip" between languages frequently when thinking about probability concepts. This might manifest as thinking in a language other than English and transferring these ideas into English or needing to see translations of terminology or questions in another language. Flipping between languages was not always tied to use of the bilingual applet, but this study revealed how the applet supports ELLs and can effectively lead to correct understanding of probability terminology in English.

ACKNOWLEDGEMENTS

This study's data are from the thesis of the third author (Salazar, 2012), and the first two authors served as thesis committee chair and committee member, respectively. This *SERJ* paper is also informed by helpful coding and/or peer debriefing from these graduate students as part of a mathematics education research course taught by the first author: Jorge Anaya, Erica Bernal, Fanny Bernal, James Bilbe, Israel Chavez, Sandra Cuevas, Jennifer Esparza, Susana Gonzales, Adam Gutierrez, Heidi Heinrichs, Carlos Hernandez, Suzana Jackson, Cole Joslyn, Perla Lampinstein, Norma Martinez, Frank Prieto, Blanca Rubio, Jeremy Sneed, and Saul Soto. The authors also express great appreciation to the anonymous referees whose feedback strengthened this paper. This work was presented in part at the 2016 conference of TODOS: Mathematics for ALL.

REFERENCES

- Abedi, J., & Lord, C. (2001). The language factor in mathematics test. *Applied Measurement in Education*, 14(3), 219–234.
- Amir, G. S., & Williams, J. S. (1999). Cultural influences on children's probabilistic thinking. *Journal of Mathematical Behavior*, 18(1), 85–107.
- Borovcnik, M., & Bentz, H.-J. (1991). Empirical research in understanding probability. In R. Kapadia & M. Borovcnik (Eds.), *Chance encounters: Probability in education* (pp. 73–105). Dordrecht, The Netherlands: Kluwer.
- Burnham T., & Schield, M. (2005). An online grammar-parsing program that decodes ordinary English descriptions and comparisons of percentages and rates. *Proceedings of the Joint Statistical Meetings, Section on Statistical Education* (pp. 2202–2210). Alexandria, VA: American Statistical Association.
- Cardenas-Hagan, E. (2010). Response to intervention: Implications for Spanish-speaking ELLs. *Perspective on Language and Literacy*, 36(2), 1–13.
- Chernoff, E. J., & Zazkis, R. (2011). From personal to conventional probabilities: From sample set to sample space. *Educational Studies in Mathematics*, 77(1), 15–33.
- Consortium for Mathematics and its Applications (COMAP) (2013). *For all practical purposes: Mathematical literacy in today's world* (9th ed.). New York: W. H. Freeman.
- Cox, C., & Mouw, J. T. (1992). Distribution of the representativeness heuristic: Can we be perturbed into using correct probabilistic reasoning? *Educational Studies in Mathematics*, 23(2), 163–178.
- Dragt, J. (Ed.) (2015). *ISI Multilingual Glossary of Statistical Terms*. International Statistical Institute. [Online: <http://isi.cbs.nl/glossary/>]
- Falk, R., & Lann, A. L. (2015). Numbers defy the law of large numbers. *Teaching Statistics*, 37(2), 54–60.
- Flores, S. M., Batalova, J., & Fix, M. (2012). *The educational trajectories of English language learners in Texas*. Washington, DC: Migration Policy Institute.
- Franklin, C. A., Bargagliotti, A. E., Case, C. A., Kader, G. D., Scheaffer, R. L., & Spangler, D. A. (2015). *The statistical education of teachers*. Alexandria, VA: American Statistical Association.
[Online: <http://www.amstat.org/education/SET/SET.pdf>]
- GAISE College Report ASA Revision Committee (2016). *Guidelines for assessment and instruction in statistics education: GAISE college report*. Alexandria, VA: American Statistical Association.
[Online: <http://www.amstat.org/education/gaise/GAISECollege.pdf>]
- Garfield, J., & Ben-Zvi, D. (2007). How students learn statistics revisited: A current review of research on teaching and learning statistics. *International Statistical Review*, 75(3), 372–396.
- Garrison, L., & Mora, J. K. (1999). Adapting mathematics instruction for ELLs: The language-concept connection. In W. G. Secada, L. Ortiz-Franco, N. G. Hernandez, & Y. De La Cruz (Eds.), *Changing the faces of mathematics: Perspectives on Latinos* (pp. 35–48). Reston, VA: National Council of Teachers of Mathematics.
- Gelman, A., & Nolan, D. (2002). You can load a die, but you can't bias a coin. *The American Statistician*, 56(4), 308–311.
- Gilovich, T., Vallone, R., & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology*, 17(3), 295–314.
- Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory*. Chicago: Aldine.

- Gnanadesikan, M., Scheaffer, R. L., Watkins, A. E., & Witmer, J. A. (1997). An activity-based statistics course. *Journal of Statistics Education*, 5(2).
[Online: <https://ww2.amstat.org/publications/jse/v5n2/gnanadesikan.html>]
- Goldenberg, C. (2008). Teaching ELLs: What the research does— and does not— say. *American Educator*, 33(2), 8–19, 22–23, 42–44.
- Green, D. R. (1984). Talking of probability... *Bulletin of the Institute of Mathematics and its Applications*, 20(9/10), 145–149.
- Gürbüz, R., & Birgin, O. (2012). The effect of computer-assisted teaching on remedying misconceptions: The case of the subject “probability”. *Computers & Education*, 58(3), 931–941.
- Harper, C., & de Jong, E. (2004). Misconceptions about teaching English-language learners. *Journal of Adolescent & Adult Literacy*, 48(2), 152–162.
- Hofstetter, C. H. (2003). Contextual and mathematics accommodation test effects for ELLs. *Applied Measurement in Education*, 16(2), 159–188.
- Jones, G. A., Langrall, C. W., & Mooney, E. S. (2007). Research in probability: Responding to classroom realities. In F. K. Lester, Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 909–956). Charlotte, NC: Information Age Publishing.
- Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, 3(3), 430–454.
- Kaplan, J. J., Fisher, D. G., & Rogness, N. T. (2009). Lexical ambiguity in statistics: what do students know about the words association, average, confidence, random and spread? *Journal of Statistics Education*, 17(3), 1–19.
[Online: <https://ww2.amstat.org/publications/jse/v17n3/kaplan.pdf>]
- Kaplan, J. J., Fisher, D. G., & Rogness, N. T. (2010). Lexical ambiguity in statistics: How students use and define the words: Association, average, confidence, random and spread. *Journal of Statistics Education*, 18(2), 1–22.
[Online: <https://ww2.amstat.org/publications/jse/v18n2/kaplan.pdf>]
- Kaplan, J. J., Rogness, N. T., & Fisher, D. G. (2012). Lexical ambiguity: Making a case against *spread*. *Teaching Statistics*, 34(2), 56–60.
- Kaplan, J. J., Rogness, N. T., & Fisher, D. G. (2014). Exploiting lexical ambiguity to help students understand the meaning of *random*. *Statistics Education Research Journal*, 13(1), 9–24.
[Online: [http://iase-web.org/documents/SERJ/SERJ13\(1\)_Kaplan.pdf](http://iase-web.org/documents/SERJ/SERJ13(1)_Kaplan.pdf)]
- Kazima, M. (2006). Malawian students’ meanings for probability vocabulary. *Educational Studies in Mathematics*, 64(2), 169–189.
- Klarreich E. (2004). *Science News Online*, 165(9), 131.
[Online: <http://www.sciencenews.org/articles/20040228/fob2.asp>]
- Knapp, T. R. (1996). *Learning statistics through playing cards*. Thousand Oaks, CA: Sage.
- Lavy, I., & Mashiach-Eizenberg, M. (2009). The interplay between spoken language and informal definitions of statistical concepts. *Journal of Statistical Education*, 17(1), 1–9. [Online: <https://ww2.amstat.org/publications/jse/v17n1/lavy.pdf>]
- Lecoutre, M. (1992). Cognitive models and problem space in “purely random” situations. *Educational Studies in Mathematics*, 23(6), 557–568.
- Lesser, L. M. (2013). Letter to the editor: The odds of academic usage of statistics terms in everyday contexts such as lotteries. *Journal of Statistics Education*, 21(1), 1–5.
[Online: https://ww2.amstat.org/publications/jse/v21n1/Lesser_letter.pdf]
- Lesser, L. M., & Glickman, M. E. (2009). Using magic in the teaching of probability and statistics. *Model Assisted Statistics and Applications*, 4(4), 265–274.

- Lesser, L., Wagler, A., Esquinca, A., & Valenzuela, M. G. (2013). Survey of native English speakers and Spanish-speaking ELLs in tertiary introductory statistics. *Statistics Education Research Journal*, 12(2), 6–31.
[Online: [http://iase-web.org/documents/SERJ/SERJ12\(2\)_Lesser.pdf](http://iase-web.org/documents/SERJ/SERJ12(2)_Lesser.pdf)]
- Lesser, L., & Winsor, M. (2009). ELLs in introductory statistics: Lessons learned from an exploratory case study of two pre-service teachers. *Statistics Education Research Journal*, 8(2), 5–28.
[Online: [http://iase-web.org/documents/SERJ/SERJ8\(2\)_Lesser_Winsor.pdf](http://iase-web.org/documents/SERJ/SERJ8(2)_Lesser_Winsor.pdf)]
- Lim, K. H., Buendia, G., Kim, O. K., Cordero, F., & Kasmer, L. (2010). The role of prediction in the teaching and learning of mathematics. *International Journal of Mathematics Education in Science and Technology*, 41(5), 595–608.
- Lipkin, L. (2003). Tossing a fair coin. *College Mathematics Journal*, 34(2), 128–133.
- Meletiyou-Mavrotheris, M., Lee, C., & Fouladi, R. T. (2007). Introductory statistics, college student attitudes and knowledge: A qualitative analysis of the impact of technology-based instruction. *International Journal of Mathematical Education in Science and Technology*, 38(1), 65–83.
- Moyer, P. S., Niezgoda, D., & Stanley, J. (2005). *Technology-supported mathematics learning environments*. Reston, VA: National Council of Teachers of Mathematics.
- National Library of Virtual Manipulatives (NLVM) (2015). Coin tossing [online applet]. Utah State University. [Online: http://nlvm.usu.edu/en/nav/frames_asid_305_g_4_t_5.html?from=search.html]
- Nolan, V. (2002). Influence of attitude towards statistics, English language ability and mathematical ability in the subject quantitative techniques at the Vaal Triangle Technikon, South Africa. In B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching Statistics: Developing a statistically literate society*, Cape Town, South Africa. [CD-ROM]. Voorburg, The Netherlands: International Statistical Institute.
[Online: http://iase-web.org/documents/papers/icots6/8a5_nola.pdf]
- Phillips, L. D., & Wright, G. N. (1977). Cultural differences in viewing uncertainty and assessing probabilities. In H. Jungermann & G. De Zeeuw (Eds.), *Decision making and change in human affairs* (pp. 507–519). Dordrecht, The Netherlands: Reidel.
- Rubel, L. H. (2007). Middle school and high school students' probabilistic reasoning on coin tasks. *Journal for Research in Mathematics Education*, 38(5), 531–556.
- Salazar, B. (2012). *Qualitative analysis of the usage by Spanish-speaking English language learners of resources for learning probability* (Unpublished master's thesis). The University of Texas at El Paso.
- Scheaffer, R. L., Gnanadesikan, M., Watkins, A., & Witmer, J. (1996). *Activity-based statistics*. New York: Springer.
- Sedlmeier, P. (1998). The distribution matters: Two types of sample-size tasks. *Journal of Behavioral Decision Making*, 11(4), 281–301.
- Sharma, S. (2006). Personal experiences and beliefs in early probabilistic reasoning: Implications for research. *International Electronic Journal of Mathematics Education*, 1(1), 33–54.
- Shaughnessy, J. M. (1977). Misconceptions of probability: An experiment with a small group, activity-based, model building approach to introductory probability at the college level. *Educational Studies in Mathematics*, 8(3), 295–316.
- Stohl, H. (2002). *Probability Explorer*. Software package (v.2.0).
- Sullivan, M. (2010). *Statistics: Informed decisions using data* (3rd ed.). Boston: Pearson.

- Viera, J. (2015, March). Creating self-directed English language learners in an online precalculus course. Paper presented at the 12th International Sun Conference on Teaching and Learning, El Paso, TX.
- Watson, J., & English, L. (2015). Confronting expectation in Grade 4: Tossing two coins. *International Journal for Studies in Mathematics Education*, 8(3), 1–30.
- Wilensky, U. (1995). Paradox, programming and learning probability: A case study in a Connected Mathematics framework. *Journal of Mathematical Behavior*, 14(2), 253–280.
- Willner, L. S., Rivera, C., & Acosta, B. D. (2009). Ensuring accommodations used in content assessments are responsive to ELLs. *The Reading Teacher*, 62(8), 696–698.

LAWRENCE M. LESSER
Department of Mathematical Sciences
The University of Texas at El Paso
500 W. University Avenue
El Paso, TX 79968

APPENDIX 1: INTERVIEW QUESTIONS

Pre-Applet Questions (“PreQ”)

- PreQ1 What do you already know about probability?
¿Qué es lo que ya sabes acerca de probabilidad?
- PreQ2 What probability or stats courses have you taken, are you taking, and plan to take?
¿Qué cursos de probabilidad o estadística has tomado, estas tomando, y planeas tomar?
[Give a penny to the student]
- PreQ3 What would you call the sides of a coin in English?
¿Cómo le llamarías a los lados de una moneda en inglés?
- PreQ4 What would you call the sides of a coin in Spanish?
¿Cómo le llamarías a los lados de una moneda en español?
Remember that you are always welcome to ask for a question to be given to you in Spanish if it is not clear in English. The English version will refer to coin sides as heads(H) and tails(T), and the Spanish version with cara(C) and escudo(E).
Recuerda que siempre está a tu disposición pedir que se te haga una pregunta en español si no entiendes en inglés. La versión en inglés se referirá a los lados de la moneda como heads (H) y tails (T), en la versión en español se les referirá como cara (C) y escudo (E).
- PreQ5 If you flip the coin twice, what do you think is the probability you get exactly one head? Why?
Si lanzas la moneda dos veces, ¿cual crees que sea la probabilidad de obtener una cara exactamente? ¿Por qué?
- PreQ6 If you flip the coin four times, what do you think is the probability you get exactly 2 heads? Why?
¿Si lanzas la moneda cuatro veces, cual crees que sea la probabilidad de obtener dos caras exactamente? ¿Por qué?
- PreQ7 If you flip the coin 100 times, how many times do you expect to get heads?
¿Si lanzas la moneda 100 veces, cuantas veces anticipas que caiga en cara?
- PreQ8 What do you think is the probability of getting exactly that many heads in 100 flips?
¿Cual crees que sea la probabilidad de obtener exactamente esa cantidad de caras en 100 lanzamientos?
- PreQ9 If you flip 100 times, how different from 50 heads would you have to get to call that result surprising?
¿Si lanzas la moneda 100 veces, que tan diferente de 50 caras tendrías que obtener para que el resultado sea sorprendente?
- PreQ10 If you flip 10 times, what do you think is the probability if you flip 10 times of getting between 40%-60% heads— in other words, exactly 4, 5, or 6 heads?
Si lanzas la moneda 10 veces ¿Cual crees que sea la probabilidad de obtener entre 40%-60% caras- en otras palabras, exactamente 4, 5, o 6 caras?
- PreQ11 If you flip 100 times, what do you think is the probability of getting between 40%-60% heads— in other words, exactly 40, 41, 42..., 58, 59, or 60 heads?
Si lanzas la moneda 100 veces, ¿Cual crees que sea la probabilidad de obtener entre 40%-60% caras- en otras palabras, exactamente 40, 41, 42,... , 58, 59, 60 caras?
- PreQ12 Why do you think your answers to the last 2 questions were [based on their answers to #10 and #11, choose the appropriate ending from: the same; different]?
¿Por qué piensas que tus respuestas a las últimas 2 preguntas son [basado en las respuestas de las preguntas 10 y 11, elije el final apropiado entre: igual, diferente]?
- PreQ13 Suppose there was a graph of the proportion of heads so far after the 100th flip, 200th flip, 300th flip, etc. Make a sketch of what you imagine it might look like on this graph paper. [Provide interviewee a pen and sheet of graph paper with vertical axis going

from 0 to 1 in jumps of 0.1 and the horizontal axis going from 0 to 1000 in jumps of 100.]

Supón que hay una grafica de la proporción de caras que han caído hasta el momento después de 100 lanzamientos, 200 lanzamientos, 300 lanzamientos, etc. Has un dibujo de lo que imaginas que podría ser la grafica en este papel.

[If the interviewee has “no idea”, show him/her five pre-made graphs and ask which seems more likely and why]

- PreQ14 Suppose there was a graph of the absolute difference between the number of heads so far and number of tails so far after the 100th flip, 200th flip, 300th flip, etc. Make a sketch of what you imagine it might look like on this graph paper. [Provide interviewee a pen and sheet of graph paper with vertical axis going from 0 to 50 in jumps of 5 and the horizontal axis going from 0 to 1000 in jumps of 100.]

Supón que hay una grafica de la diferencia absoluta entre el número de caras y escudos hasta el momento después de 100 lanzamientos, 200 lanzamientos, 300 lanzamientos, etc. Has un dibujo de lo que imaginas que podría ser la grafica en este papel para graficar.

[If the interviewee has “no idea”, show him/her five pre-made graphs and ask which seems more likely and why]

For the next couple of questions, consider these two sequences of coin flips:

Para las próximas preguntas, ten en cuenta estas dos secuencias de lanzamientos de la moneda:

(A) H T T T H T T H T H T T T H T T T T H H T T T H T T H T T T H T T T T H T T T H T

(B) H T H T T H H T H T H H T T H T T H H T T H T H H T T H T H T H T H T H T H T H T

- PreQ15 What sequence has a higher probability of happening or are they equally likely?
¿Qué secuencia tiene una mayor probabilidad de ocurrir o son igualmente probables?
- PreQ16 If one of the sequences were made up and the other actually obtained from real coin flipping, which do you think was which and why?
Si una de las secuencias fue inventada y la otra realmente obtenida lanzando una moneda, ¿Cuál crees que fue cual y por qué?
- PreQ17 In your own words, what does “longest run” mean to you?
En tus propias palabras, ¿Que significa la “mayor racha” o el mayor número de caras sucesivas?
- PreQ18 In your own words, what does “in the long run” mean to you?
En tus propias palabras, ¿Que significa “a largo plazo”?
- PreQ19 In sequence A, what is the longest run? In sequence B, what is the longest run? For a 100-flip sequence, how long do you think the longest run (of either all H’s or all T’s) would be?
En la secuencia A, ¿Cuál es la mayor racha o el mayor número de caras sucesivas? En la secuencia B, ¿Cuál es la mayor racha o el mayor número de caras sucesivas? Para una secuencia de 100-lanzamientos, ¿Qué tan larga crees que sea la racha más larga de número de caras o escudos?
- PreQ20 If you saw a streak of 4 heads in a row for a fair coin, what do you think is the probability of heads for the next flip? [If they need choices, give them these 3 choices: less than 0.5, equal to 0.5, greater than 0.5] Why?
Si vez una racha de cuatro caras de una moneda justa, ¿Cuál crees que es la probabilidad de que el próximo lanzamiento sea cara? [Si necesitan opciones: menor que 0.5, igual a 0.5, mayor que 0.5] ¿Por qué?
[On the computer screen get to the following address:
http://nlvm.usu.edu/en/nav/frames_asid_305_g_3_t_5.html?from=topic_t_5.html
(easy way: Google the words: NLVM coin tossing) and show the interviewee how he/she can toggle with the pull down menu to the Spanish version— briefly show it, and then toggle right back to the English version]

Applet Questions (“AQ”)

- AQ1 Run the applet (to do a 100-flip sequence) about 5 times. How does this change or confirm what you predicted about how far off (i.e., the “chance error” number) from 50 heads you would reasonably expect to get in a 100-flip sequence?
Ejecuta el programa (para realizar una secuencia de 100 lanzamientos) alrededor de 5 veces. Como cambia esto o confirma lo que predijiste sobre que tan lejos (i.e., el número de “error casual”) de 50 caras se puede considerar razonable en una secuencia de 100-lanzamientos.
- AQ2 Do you think a 1000-flip sequence would have a larger, smaller, or equal “chance error” compared to a 100-flip sequence? Why?
¿Crees que una secuencia de 1000-lanzamientos tendría un mayor, menor o igual “error casual” en comparación con una secuencia de 100-lanzamientos? ¿Por qué?
- AQ3 Does exploring with the applet make you feel that you want to keep your answer or change your answer to the pre-applet question #11? Explain.
Explorar el programa ¿Te hace que quieras conservar o cambiar tu respuesta a la pregunta #11 antes de usar el programa? Explica.
- AQ4 Does exploring with the applet make you feel that you want to keep your answer or change your answer to the pre-applet question #12? Explain.
Explorar el programa ¿Te hace que quieras conservar o cambiar tu respuesta a la pregunta #12 antes de usar el programa? Explica.
- AQ5 You had predicted in a previous question [pre-applet question #19] that there would be a longest run of ____ in a 100-toss sequence. Do your results seem to confirm or conflict with your prediction?
Habías predicho en la pregunta anterior que la mayor racha o el mayor número de caras sucesivas es de ____ en una secuencia de 100-lanzamientos. ¿Tus resultados parecen confirmar o entran en conflicto con tu predicción?
[Change the mode of the applet to “longest run of heads” and insert 4]
- AQ6 How many tosses do you think it will take before we see a streak of 4 heads in a row?
¿Cuántos lanzamientos crees que tomara para que veas una racha de cuatro caras consecutivas?
[After prediction is made, have the interviewee do the simulation 6 times]
- AQ7 Do your results seem to confirm or conflict with your prediction?
Tus resultados ¿Parecen confirmar o entran en conflicto con tu predicción?
- AQ8 If we change the probability of heads from .5 to .7, what change, if any, do you think there would be a difference on how many tosses it takes to get a streak of 4 heads in a row?
Si cambiamos la probabilidad de caras de .5 a .7, ¿Que fue lo que cambió, si es que hubo algún cambio?, ¿Crees que habrá una diferencia en cuantos lanzamientos se necesitan para obtener una racha de cuatro caras consecutivas?
[Change the probability from .5 to .7 and have the interviewee simulate this 6 times]
- AQ9 Do your results seem to confirm or conflict with your prediction?
Tus resultados ¿Parecen confirmar o entran en conflicto con tu predicción?
[Change the mode back to number of tosses, keep the probability at .7, and have interviewee run the applet 6 times]
- AQ10 If we change the probability of heads from .5 to .7 ...
(a) What change, if any, do you think there would be on the number of heads?
(b) What change, if any, do you think there would be on the longest run of heads?
Si cambiamos la probabilidad de caras de .5 a .7 ...
(a) ¿Que cambios, si los hay, crees que habrá en el número de caras?
(b) ¿Que cambios, si los hay, crees que habrá en la mayor racha de caras?

Post-Applet Questions (“PostQ”)

- PostQ1 How did you decide when to use the English version and when to use the Spanish version of the applet?
¿Cómo decidiste cuando usar la versión en inglés y cuándo usar la versión en español del programa?
- PostQ2 What difference between the two versions did you notice or experience?
En tu experiencia, ¿Que diferencias entre las dos versiones notaste o experimentaste?
- PostQ3 Do you think that using applets like this would or would not be helpful to you or to other students? Why?
¿Crees que el uso de programás como este serían o no serían útiles para ti u otros estudiantes? ¿Por qué?
- PostQ4 Do you think it would or would not be helpful to have a word list such as this one [*show list from EXCEL file*] that shows English and Spanish equivalents for probability terms? Why?
¿Crees que sería o no útil tener una lista de palabras como esta [*mostrar lista de archivos EXCEL*] que muestra términos equivalentes de probabilidad en inglés y español? ¿Por qué?

APPENDIX 2. ENGLISH-SPANISH PROBABILITY WORD LIST

English Word	Spanish Translation
Average	media aritmética ; promedio aritmético ; promedio ; media
Bar Graphs	gráfico de barras ; diagrama de barras
Bell Shaped Curve	curva campanular ; curva acampanada ; curva campaniforme
Categorical Variable	variable categorico
Correlation	correlación
Data	datos
Discrete Variable	variable discreta
Independent	independencia
Mean	media aritmética ; promedio aritmético ; promedio ; media
Median	mediana
Mode	moda ; valor modal
Normal Distribution	distribución normal ; distribución de Gauss
Personal Probability	probabilidad subjetiva
Probability	probabilidad
Proportion	proporcion
Random Numbers	números del muestreo al azar ; números aleatorios para muestreo
Relative Frequency	frecuencia relativa
Variability	mutabilidad ; variabilidad