The Mathematics Enthusiast

Volume 16
Number 1 *Numbers 1, 2, & 3*

Article 11

2-2019

The Expected Value of a Random Variable: Semiotic and Lexical Ambiguities

Hyung Won Kim

Tim Fukawa-Connelly

The Expected Value of a Random Variable: Semiotic and Lexical Ambiguities

Hyung Won Kim[1]
University of Texas Rio Grande Valley

Tim Fukawa-Connelly
Temple University

Abstract: In calculus-based statistics courses, the expected value of a random variable (EVORV) is discussed in relation to underlying mathematical notions. This study examines students' understanding of the mathematical notions of EVORV in connection with its semiotic and lexical representations. It also assesses students' computational competency revolving around EVORV. We collected qualitative data via surveys and interviews from eight students enrolled in a calculus-based university statistics course. The results suggest that while the students in general had the computational accuracy to correctly calculate EVORV, they struggled to understand the notion, and in particular to make sense of the term "random" in "random variable" and the symbol $\mathrm{E}(X)$ in the mathematical context. The study provides a basis for understanding potential challenges to students' learning of EVORV and other related statistics topics and how such challenges may emerge from the semiotic and lexical ambiguities inherent in terms and symbols used in statistics.

Keywords: Expected value of a random variable; Calculus-based statistics; Mathematical understanding; Lexical ambiguity; Semiotic ambiguity

Introduction

Many documents setting goals for the teaching and learning of statistics emphasize statistical literacy and reasoning (e.g., Aliaga et al.'s 2005 *Guidelines for Assessment and Instruction in Statistics Education* [GAISE] *College Report*; Carver et al.'s 2016 GAISE *College Report*; Ben-Zvi & Garfield, 2004; Rumsey, 2002). Developing statistical literacy refers to "acquiring basic and important skills in understanding statistical information or research results" (Ben-Zvi & Garfield, 2004, p. 7), and it requires students to understand statistics terms and symbols (Garfield, 1999; Rumsey, 2002). Recent research, however, suggests that students face challenges in doing both of these in college level elementary statistics. One such challenge relates to the lexical ambiguities inherent in statistical terms (Kaplan, Fisher, & Rogness, 2010). For example, reliance on the "everyday" definition of terms such as "association," "average," "confidence," "random" and "spread" can lead to student misconceptions. Furthermore, students often have trouble making sense of symbolic expressions composed of symbols that encompass opposing notions (Kim, Fukawa-Connelly & Cook, 2016). For example, students tend to mix up symbols for statistics with symbols for parameters, and to associate the symbol for the population standard deviation of a sampling distribution with the sample standard deviation, due to the conflicting meanings of the symbolic expressions' components.

[1] hyung.kim@utrgv.edu

Developing statistical reasoning refers to "understanding and being able to explain statistical processes and interpret statistical results" (Ben-Zvi & Garfield, 2004, p. 7). According to the most recent *GAISE College Report* (Carver et al., 2016), having a good understanding of statistics concepts will make it easier for students to use necessary tools and procedures to answer particular questions about a dataset. That is, students need conceptual understanding of statistics topics in order to understand and be able to explain statistical processes and interpret statistical results. As a way to enhance the understanding of statistics concepts, Aliaga et al. (2005, p. 18) and Carver et al. (2016, p. 16) recommend instructors "use formulas that enhance the understanding of concepts" rather than those that inhibit understanding. For example, to enhance students' understanding of the role of standard deviation as a measure of spread, the report suggested the use of the expression $s = \sqrt{\sum(y-\bar{y})^2/(n-1)}$ rather than the expression $s = \sqrt{(\sum y^2 - \frac{1}{n}(\sum y)^2)/(n-1)}$ because the former "helps students understand the role of standard deviation as a measure of spread and see the impact of individual y values on s" whereas the latter "has no redeeming pedagogical value" (p. 18). That is, with the former, one can establish understanding of how the procedural sequence of subtracting $\bar{y}$ from y, summing up the squares of $(y-\bar{y})^2$, dividing the sum of the squares, $\sum(y-\bar{y})^2$, by $n-1$, and so on contributes to yielding a numerical value that measures the spread of a data distribution. Being able to connect the meaning of a statistical expression with the notions that underlie the operations comprising the procedure in the expression is particularly important in calculus-based statistics courses. In an overview, "On a Calculus-based Statistics Course for Life Science Students," Watkins (2010) noted that pedagogy should take advantage of the quantitative capabilities of the students. We can reasonably assume that students in calculus-based statistics courses (*calculus statistics courses* hereafter) are expected to have a stronger mathematical background than students in non-calculus-based statistics courses. The implication is: in order to promote students' statistical reasoning in calculus statistics courses, instructors must introduce statistics concepts by emphasizing the underlying mathematical notions and must rely on students' mathematical background.

In short, the importance of making sense of statistical terms and symbols to develop statistical literacy, and the importance of learning the mathematical notions of statistical expressions to develop statistical reasoning are well established. Nevertheless, the authors' experiences of teaching statistics suggest that calculus statistics students struggle to connect statistics expressions with relevant mathematical notions and to make appropriate sense of statistics terms and symbols. Hence, this study explores (1) how calculus statistics students make sense of statistics terms and symbols and (2) how they make sense of the mathematical notions of statistical expressions in their learning of the concept of *expected value of a random variable* (EVORV).

Major textbooks for calculus statistics courses, particularly for science and engineering (e.g., Montgomery, Runger, & Hubele, 2009; Walpole, Myers, Myers, & Ye, 2014), emphasize two aspects of EVORV: the mathematical basis of the definition of EVORV and its computational role in the learning of *bias*. The definition draws on the ideas of *weighted average* and *probability function* either by means of summation (when the random variable is discrete) or by means of definite integration (when the random variable is continuous) (Montgomery, Runger, & Hubele, 2009). At the calculus statistics level, students should be able to connect the notion of EVORV to these underlying mathematics ideas: weighted average, summation and definite integration. Regarding EVORV's computational role in the learning of *bias*, algebraic computations using EVORV are frequently dealt with in calculus statistics textbooks used for engineering and science majors (e.g., Montgomery, Runger, & Hubele, 2009; Walpole, Myers, Myers, & Ye, 2014). While routine computations are discouraged at the college level of statistics learning (Aliaga et al., 2005; Carver et al., 2016), the authors consider computations that use EVORV (as a linear operator) to be a critical component of such statistics courses due to EVORV's role as a tool in learning the notions of both *estimator* and *bias*, as in "unbiased estimator." For this reason, this study is also concerned with students' understanding of topics such as *bias*, *variance* and *random variable*. By focusing on students' understanding of EVORV and these related topics, this study addresses the following research questions in the context of calculus statistics courses:

- How do students make sense of the phrase "expected value of a random variable" in conjunction with its lexical features, and how do they make sense of the symbol $\mathrm{E}(X)$ in conjunction with its semiotic features?
- How well do students connect the concept of EVORV to the underlying mathematical notions? And how does their ability to do so compare to their computational proficiency with EVORV and related topics (e.g., *bias*, *variance* and *random variable*)?

Research Background

A line of research has investigated statistics students' understanding of means at different levels. Mokros and Russell (1995) found that 4^{th}–8^{th} graders had difficulty understanding an average as representative of data, and that those who depended on algorithms in calculating averages were likely to have difficulty in construction problems. In particular, according to Mokros and Russell, the students commonly mastered the necessary algorithm for calculating averages without developing a sense of representativeness of data or balance, and without developing a mathematical perspective. Mayén, Diaz and Batanero (2009), in their study of how secondary students make sense of means and medians, found that students often mixed up the symbols for statistics and parameters and misinterpreted means and medians. Kim, Fukawa-Connelly and Cook (2016) found that students often have trouble connecting the concepts and the symbols associated with means. The current study continues this line of research by investigating students' understanding of *expected value* or $\mathrm{E}(X)$, which is an advanced concept of average in that it is defined as a long-term average in the context of repeated experiments. The statistics education literature shows very little attention to this topic. This study approaches the topic by considering the lexical and semiotic features of expected value's representations, and by drawing upon Godino and Batanero's (2003) construct of *semiotic function*.

Lexical and Semiotic Ambiguity

Lexical Ambiguity. Lexical ambiguity arises where more than one meaning can be attributed to a particular word or phrase (Simpson, 1981). Barwell (2005) suggested the use of everyday language in the math classroom creates ambiguities when mathematical language and everyday language overlap. In the same way, ambiguity can arise in statistics classrooms if statistical language and everyday language overlap. In addition, because the discipline of statistics differs from the discipline of mathematics in various ways (Kim & Fukawa-Connelly, 2015; Wild & Pfannkuch, 1999), lexical ambiguity can arise when the two disciplines have different definitions and uses for the same term.

A few studies have considered the issue of lexical ambiguity in statistics in various languages: English (Kaplan, Fisher, & Rogness, 2010), Hebrew (Lavy & Mashiach-Eizenberg, 2009) and Korean (Jung & Hwang, 2016). These studies focused on the five English words “association,” “average,” “confidence,” “random” and “spread,” or their Hebrew or Korean equivalents, which have different meanings in everyday language and statistics language. They investigated how these lexical ambiguities create challenges for elementary statistics students’ understanding of the terms’ definitions in statistics. For example, when the word “random” is used in the phrase “random sampling,” it means each individual in the sample was chosen entirely by chance so that all individuals have the same probability of being chosen in the sampling process (Yates, Moore, & Starnes, 2008). The studies’ findings indicate that students often misuse the term by invoking its everyday meaning of uncertainty. Furthermore, there is evidence that ambiguity can arise between two disciplines when their terminologies overlap. Kim, Fukawa-Connelly and Cook (2016) found that mathematically strong students with an appropriate understanding of the definition of a variable in mathematics could have trouble understanding how the phrase “random variable” is used in statistics. In the current study, “lexical ambiguity” refers to ambiguities that arise due to the overlapping meaning between statistics language and everyday language or between statistics language and language in other fields.

Semiotic Ambiguity. A different kind of challenge addressed in the literature on collegiate elementary statistics is students’ misconceptions of symbols. For example, elementary statistics students tend to mix up symbols used for statistics and for parameters (Mayén, Diaz, & Batanero, 2009). In particular, students misuse the symbols that refer to *the standard deviation of the sampling distribution of a sample mean* due to the opposing notions between of the synthesizing symbols σ (population standard deviation) and $\bar{x}$ (sample mean) in the semiotic function of $\sigma_{\bar{x}}$ and between the symbol σ (population standard deviation) and the procedure of dividing σ by $\sqrt{n}$ (the common practice used to find the standard error of the sample mean) in the formula $\sigma/\sqrt{n}$ (Kim, Fukawa-Connelly, & Cook, 2016). In general, students’ ability to correctly interpret words and symbols is important because it impacts students’ ability to correctly interpret texts. In this paper, for consistency and convenience, we refer to the pedagogical conflicts caused by notational features of a symbolic expression as *semiotic ambiguity*.

Ambiguities in EVORV and $E(X)$. The potential for lexical ambiguity in EVORV lies in the phrase “random variable.” According to Kaplan, Fisher and Rogness (2010), students often associate the everyday meaning of “random” (i.e., without order or reason) with their understanding of “random sample.” Analogously, the multiple meanings of “random” may give rise to lexical ambiguity in students’ learning of the phrase “random variable” which is defined as “a numerical variable whose measure value can change from one replicate of the experiment to another” (Montgomery, Runger, & Hubele, 2009, p. 54). The constituents of the symbol $\mathrm{E}(X)$ also carry the potential for semiotic ambiguity. The symbol represents the expected value of a random variable X, and therefore is a fixed value. However, due to the common use of the symbol $f(x)$, which typically represents a function of a variable x, and the similarity in appearance between $\mathrm{E}(X)$ and $f(x)$, students could potentially misunderstand $\mathrm{E}(X)$ as a function. The potential for such ambiguity has been indicated in an empirical study (Kim, Fukawa-Connelly, & Cook, 2016), where students were found to develop misconceptions when the constituting symbols of a concept suggest contradictory meanings.

Understanding EVORV or $E(X)$ as Semiotic Functions

Godino and Batanero (2003) viewed a symbol and a term (or phrase) used in the mathematics context as a semiotic function, conceiving of it as a correspondence among three sets: a set of signs (or "an expression plane"), a set of contents (or "a content plane") and a set of interpretive codes (or "a correspondence plane"). They held that, depending on the nature of the "content," a semiotic function generates four types of meaning: ostensive, intensive, actuative and extensive (Table 1). The description column in Table 1 shows the content represented by a semiotic function, and the examples are modified versions of Godino and Batanero's original examples.

Table 1. *Types of Meaning of Semiotic Functions (Godino & Batanero, 2003)*

Type	Description	Examples
Ostensive	Names or indications of objects	• 5! denotes $5 \cdot 4 \cdot 3 \cdot 2 \cdot 1$.
Intensive	Notions of mathematical properties or interpretation of objects	• The word "angle" denotes a shape formed by two rays diverging from a common point. • $f(x)$ denotes a continuous function.
Actuative	Subject's action	• $(2/3)/12$ implies the calculation of dividing 2 by 3 and dividing its results by 12.
Extensive	Situation-problem	• An urn model is used to represent a particular probabilistic problem.

This study considers the ostensive, intensive and actuative meanings of EVORV and the related topics of random variable and bias. If one can define a concept given either the term (or phrase) or the symbol that denotes the concept, one has the ostensive meaning. If one can use the critical features inherent in the definition in an applied context, one has the actuative meaning. And if one understands the definition in connection with either the underlying mathematical notion (if applicable) or the reasoning structure, one has the intensive meaning. The extensive meaning, which could be considered as a particular problem-solving situation, is beyond the scope of this study. In the paragraphs that follow, we discuss the three types of meaning for EVORV based on this formal definition of EVORV (Montgomery, Runger, & Hubele, 2009):

> (Continuous case) Suppose X is a continuous random variable with pdf $f(x)$. The mean or expected value of X, denoted as μ or $\mathrm{E}(X)$, is $\mu = \mathrm{E}(X) = \int_{-\infty}^{\infty} x f(x)\, dx$. (p. 64)
>
> (Discrete case) Let the possible values of the random variable X be denoted as $x_1, x_2, \dots, x_n$. The pmf of X is $f(x)$, so $f(x_i) = \mathrm{P}(X = x_i)$. The mean or expected value of the discrete random variable X, denoted as μ or $\mathrm{E}(X)$, is $\mu = \mathrm{E}(X) = \sum_{i=1}^{n} x_i\, f(x_i)$ (p. 90)

Ostensive Meaning. The ostensive meaning of EVORV involves mathematical expressions, as shown in this definition. To evaluate the students' knowledge of the ostensive meaning of EVORV, this study assessed whether students could write these mathematical expressions both in computational questions in problem-solving settings and in questions asking them to explicitly state the definition of the phrase/symbol. In this way, the study is able to explore student misconceptions in their sense-making of the phrase/symbol at the ostensive level.

Actuative Meaning. The final object of EVORV's actuative meaning is the act of doing the computations to find the expected value. The process of computing the EVORV may embody the ostensive meaning. For example, for a discrete random variable X, step (1) of the sequence of equations in Table 2 indicates the ostensive meaning and steps (2) through (4) indicate the actuative meaning.

Table 2. *Example: Sequence of Equations to Solve for E*(X)

Equations	$\mathrm{E}(X) = \sum_{i=1}^{n} x_i\, f(x_i) = 1 \cdot f(0) + 2 \cdot f(1) = 1 \cdot (1/4) + 2 \cdot (3/4) = 7/4$
Steps	(1) (2) (3) (4)

Making sense of the actuative meaning of the expected value of a random variable plays an essential role in the study of calculus statistics in that certain statistics notions rely on the computations in steps (2) through (4) in this equation sequence. For example, the process of performing the computations in these steps is an essential component of developing understanding of the notion of *bias*. Montgomery, Runger and Hubele's (2009) definition of *bias* is: the point estimator $\hat{\Theta}$ is an unbiased estimator for the parameter θ if $\mathrm{E}(\hat{\Theta}) = \theta$, and if the estimator is not unbiased, then the difference $\mathrm{E}(\hat{\Theta}) - \theta$ is called the bias of the estimator $\hat{\Theta}$. To determine whether an estimator is biased, one needs to consider this difference and should be able to do the computations shown in Table 2 (or its equivalent in the continuous case) to find the expected value. For this reason, this study examines students' fluency in computing $\mathrm{E}(X)$.

Computational fluency is considered in terms of three factors: *efficiency*, *accuracy* and *flexibility* (Montiel, 2005; National Council of Teachers of Mathematics [NCTM], 2000). According to Montiel, *efficiency* refers to the development of strategies that take less time and effort; *accuracy* to arriving at the correct answer through the correct usage, manipulation and interpretation of mathematical symbols; and *flexibility* to the ability to recognize when a strategy is not applicable and then change the strategy. In this study, we limit our assessment of computational fluency to the aspect of *accuracy* because it relates directly to students' interpretation of mathematical symbols and their immediate responses to their interpretations. Developing accuracy in computations pertaining to the expected value of a random variable may provide the basis for developing the statistics notion of bias. Therefore, we measure students' acquisition of actuative meaning by the accuracy of their computation.

Intensive Meaning. The intensive meaning of an object, in the learning of mathematics, connotes the mathematical ideas and abstractions of the object including definitions, concepts, propositions, procedures and theories (Godino & Batanero, 2003). In the teaching of statistics, we need to help students understand the intensive meaning in conjunction with the goals of developing statistical literacy and reasoning. Whereas statistical literacy merely refers to understanding the basic language of statistics and acquiring basic knowledge needed for understanding statistical information or research results (Aliaga et al., 2005; Ben-Zvi & Garfield, 2004), statistical reasoning involves a higher level of understanding of the discipline: it refers to the way people reason with statistical ideas and make sense of statistical information (Garfield & Chance, 2000) as well as understanding statistical processes and being able to explain and interpret statistical results (Ben-Zvi & Garfield, 2004). The authors of this paper hold that statistics reasoning relates to understanding mathematical ideas that underlie statistics concepts and making connections between statistics concepts and the underlying mathematical ideas. While developing such mathematical ideas is critical in the study of statistics, students' analysis of statistics results tends to be limited to stating immediate features such as comparing means rather than providing justifications for how their analysis was appropriate to the question (Pfannkuch, Budgett, Parsonage, & Horring, 2004).

Intensive meaning with regard to EVORV indicates the concept of what the expected value signifies in the context of given statistics problems. In addition, it includes the notion associated with the mathematical roles that $x \cdot f(x)$ and $x_i \cdot f(x_i)$ play, respectively, in the expressions

$\int_{-\infty}^{\infty} xf(x)\, dx$ and $\sum_{i=1}^{n} x_i\, f(x_i)$ (as described in Table 2) in yielding the weighted average of the values that X takes. Understanding such mathematical ideas would help students understand what the expected value signifies (the average number of the variable outcomes in the long run) in a given context. Therefore, students with such a mathematical understanding of EVORV would be able to interpret the results of statistical processes involving ideas of EVORV and explain how their analysis of statistics information supports their conclusions.

Methods

This study was conducted with eight students enrolled in a statistics course intended for juniors majoring in science and engineering at a mid-sized public university in the U.S. The course is calculus-based in the sense that two calculus courses, *Calculus I* and *II*, are prerequisites. These students volunteered from a class in which approximately 70 students were enrolled at the time data were collected. Among the eight students, only one had taken a statistics course before (AP statistics). The curriculum of the course was typical of a reform-oriented classroom in that the instructor emphasized statistical reasoning and thinking. However, the instruction was traditional in the sense that the class content was delivered primarily via lecture.

The first author of this paper was the class's teaching assistant (TA). It was his third time as TA for the same course and same instructor. The class met twice a week for one semester, and the author was responsible for an additional recitation class once a week. In his first semester as TA for the class, the author had voluntarily observed most of the classes. The course introduced the design of controlled experiments and the collection and analysis of scientific data. It covered graphical data analysis, statistical process control, regression, correlation, multifactor experimental designs, confidence intervals and hypothesis testing. Students were expected to learn various probability distributions. Use of a statistical software package was an integral part of the course. The instructor emphasized drawing conclusions from software (SAS) and interpreting them. Knowledge of calculus was essential for the course in that the discussion of the major concepts such as expected value, standard deviation and probability density function required integration.

The instructor often provided the meaning and interpretation of statistical concepts when introducing their mathematical expressions, but generally did not discuss the statistical concepts' mathematical foundations in detail. For example, EVORV was introduced both in the context of mathematics as $\int_{-\infty}^{\infty} xf(x)\, dx$ or $\sum_{i=1}^{n} x_i\, f(x_i)$ and in the context of statistics as the long-run average value expected in indefinite repetition. However, this was the entire treatment of the link between the mathematical expression and the statistical interpretation.

The degree to which the class conformed to the recommendations of the *GAISE College Reports* (Aliaga et al., 2005; Carver et al., 2016) was limited. For introductory statistics, the reports recommend developing statistical literacy/thinking, using actual data, promoting conceptual understanding, fostering active learning and using technology. In particular, the reports suggest teachers use formulas that enhance the understanding of concepts (as opposed to formulas that do not) to improve students' understanding of the role of statistical concepts in context. The class did not follow certain GAISE recommendations in that (1) explanations of how mathematical definitions or expressions are linked to the meanings they represent in the statistical context were rare; and (2) computations of statistical processes rarely depended on technology in either class lessons or homework/exams. However, it is important to note that the guidelines provide general guidance for teaching statistics at the *introductory* college level. This may limit the range of statistics classes in which the GAISE guidelines can be fully adopted. Because the goal of the course was to teach experimental design to junior engineering and science majors, there was no reason to expect that it would have completely adopted the GAISE recommendations. The class did conform to the recommendations in the sense that it often generated and used actual data and incorporated the use of software to derive statistical outcomes.

For data collection, the researchers administered a ten-item survey assessment to the participants and then conducted follow-up interviews. We created the assessment items to measure students' understanding of EVORV and other relevant topics such as random variable, variance, estimator and bias. The survey items focused on assessing three aspects of learning: (1) ostensive (the ability to explicitly state the definitions of statistical terms and symbols), (2) actuative (the ability to perform computations accurately), and (3) intensive (the understanding of the underlying mathematics notions of the statistics topics). Some items dealt with a single topic, while others dealt with overlapping topics. The greatest emphasis was placed on the concept of expected value. In the interviews, students were asked to explain their survey answers. All interviews were conducted by the first author. The interviewer initially asked the subjects to explain their answers to the survey items. Following questions then naturally arose, focusing on the reasons why they answered the questions the way they did. For example, following the participants' initial responses, the interviewer might ask questions such as "what does that mean?"; "can you explain why you…?"; "how do you know…?" and "why do you think…?" When remarks were unclear, the interviewer attempted to clarify them by following up with questions such as "did you mean…or…?" All interviews were audio-recorded and transcribed for the qualitative analysis. We assessed each utterance for the information it gave about a student's understanding of the expected value of a random variable, and then categorized the utterances by the type of understanding students showed within the utterance. We then read within and across categories to find students' learning patterns.

We coded the categories that emerged from the data using a three-level system to indicate the topic, the nature of the knowledge (i.e., the three aspects of learning), and the students' understanding. First, to denote the topics, we used RV for *random variables*, EV for *expected value* and *variance*, and BS for *bias*. This study includes students' understanding of variance in our consideration of their understanding of EVORV. This is because, while the concepts' symbolic expressions are similar, taking the function-like forms $\mathrm{E}(X)$ and $\mathrm{Var}(X)$, their mathematical bases are different: $\mathrm{E}(X)$ has linear features while $\mathrm{Var}(X)$ does not. We noted contrasts in a student's sense-making of expected value and variance and used the differences to describe how that student understood expected value. Second, we used numbers to represent the nature of the knowledge: 1 for ostensive meaning, 2 for actuative, and 3 for intensive. In measuring students' actuative meaning, we focused on the accuracy of their computations, that is, arriving at the correct answer through the correct usage and manipulation of symbols. Lastly, the symbols + or – indicated whether the relevant student understanding was "proper" or

Table 3. *Code List*

Type of meaning	Random variable	Expected value and variance	Bias
Ostensive meaning	RV1+ / RV1–	EV1+ / EV1–	BS1+ / BS1–
Actuative meaning	RV2+ / RV2–	EV2+ / EV2–	BS2+ / BS2–
Intensive meaning	RV3+ / RV3–	EV3+ / EV3–	BS3+ / BS3–

"improper," in the sense of conforming to a statistical expert's way of thinking. For example, RV1+ indicates: "the student properly states the definition of 'random variable' or makes sense of the symbol X that represents a random variable"; EV2– indicates "the student has not fully developed an actuative meaning of (i.e., shows lack of accuracy in the computations for finding) the expected value or the variance of a random variable." The codes are listed in Table 3. The use of RV, EV and BS in conjunction with 1, 2 and 3 and +/– in the coding allowed us to compare students' understanding of EVORV with their understanding of the related topics (*random variables* and *bias*) and organize our findings by the types of meaning for EVORV.

The ten survey items were created to measure students' sense-making of the three types of meaning for the three aforementioned topics. The items addressed the topics and knowledge types as shown in Table 4. As a final step of coding, we evaluated the items (or sub-items) in terms of the severity of any misconceptions. We scaled this evaluation to range from 0 to 2, with 0 indicating that the students did well in general (i.e., most students answered the questions in these items well) and 2 indicating that the students needed improvement in general (i.e., more than half the students held misconceptions).

Table 4. *Survey Items: Topics and Types of Meaning*

Items	1	2	3	4	5	6	7	8	9	10
RV	RV3		RV1/2/3	RV1/2/3						RV1
EV		EV3	EV1/2/3	EV1/2/3	EV3	EV3	EV2	EV3	EV1/3	
BS						BS1/3	BS2	BS2		

Findings

The findings of this study are presented in two parts. The first part provides a summary of the overall survey results. We report the students' challenges and strengths in regard to the topics and aspects of knowledge we investigated. Further, we discuss the nature of the students' work, focusing on the items that revealed student challenges. The second part focuses on the students' challenges in understanding with regard to the lexical features of EVORV, semiotic features of $\mathrm{E}(X)$ and mathematical notions underlying the concept of EVORV, in that order. In the second part, qualitative evidence from the analysis of the interviews as well as the surveys is provided.

Summary of the Survey Results

The survey outcomes are summarized in Table 5. The participants' names are all pseudonyms. The first column of Table 5 lists the survey items by number. Items 3, 4, 7 and 9 had sub-questions. These items are sub-divided in this column into more than one row if different sub-questions addressed different topics or aspects of learning. For items 7 and 9, some sub-questions are grouped together because they addressed the same aspects of learning. The "scale" column shows the ratings used to assess student performance for each question or group of sub-questions. Items 1 and 5 have numerical scales ranging from 0 to 3 and 0 to 7, respectively. This is because item 1 had three parts that are alike and item 5 had seven parts that are alike. We gave points for these two items depending on the number of parts each student answered correctly. Other items have either "× or ○" or "× • ○" in the scale column. In the rows with "× or ○" (items 2 to 4b), × means the answer is incorrect while ○ means the answer is correct. These items are either multiple choice questions or short-answer type questions. For items 6 to 10, × means that the answer is incorrect, • means the answer is partly correct, and ○ means the answer is correct. The last column shows the severity of the students' misconceptions involved in each item, scaled to range from 0 to 2, with 0 indicating that the students did well in general and 2 indicating that the students need improvement in general. The summary shows that the student performances on items 4, 9 and 10 were not as good as their performances on the other items. In what follows, we discuss the nature of the students' work on these items.

Items 3 and 4. While the questions in items 3 and 4 are similar, the student performances showed differences. The two items follow:

Table 5. *Survey Outcomes Summarized by Item*

Items	Scale	Performance of eight participants								Topics			Severity
		Debora	Simon	Sandy	John	Norma	Kara	Renata	Aaron	RV	EV	BS	0–2
1	0–3	1	3	1	3	3	0	1	2	3			1
2	× or ○	○	○	○	×	○	×	○	×		3		1
3a	× or ○	×	○	○	○	×	×	×	×	1/2			1
3b	× or ○	○	×	○	○	×	○	×	×		1/2		1
4a	× or ○	×	×	×	○	×	×	○	×	1/2/3			2
4b	× or ○	×	×	×	○	×	×	×	×		1/2/3		2
5	0–7	4	4	6	5	6	5	7	4		3		1
6	× • ○	•	•	○	○	•	○	○	•		3	1/3	1
7ab	× • ○	○	○	○	○	○	○	○	○		Symbol Questions		0
7c	× • ○	○	○	○	○	○	•	○	•		2	2	0
8	× • ○	○	○	○	○	○	○	•	○		3	2	0
9a-d	× • ○	•	×	×	×	×	×	×	×		1/3		2
9ef	× • ○	×	×	×	×	×	×	×	×		3		2
10	× • ○	×	○	×	×	×	×	×	×	1			2

3. Let X denote a variable that takes on any of the values -1, 0, and 1 with respective probabilities $P(X=-1)=.2$, $P(X=0)=.5$, and $P(X=1)=.3$.
a. Is X a random variable? Why or why not?
b. Compute $\mathrm{E}(X)$.

4. Let X denote a random variable that takes on any of the values 0, 1, and 2 with respective probabilities $P(X=0)=.1$, $P(X=1)=.3$, and $P(X=2)=.6$.
a. Is X^2 a random variable? Why or why not?
b. Compute $\mathrm{E}(X^2)$.

Sub-questions 3*a* and 4*a* examine students' ability to use the criteria of random variables in determining the validity of a random variable. Sub-questions 3*b* and 4*b* address students' computational ability to find the expected value of a given expression. The rate of correct answers for items 3*a* and 3*b* combined is 7/16, whereas the rate for items 4*a* and 4*b* combined is 3/16 (Table 6).

Table 6. *Student Performance on Items 3 and 4*

Items	3a	3b	4a	4b
Success rate	3/8	4/8	2/8	1/8
Combined success rate	7/16		3/16	

Items 3 and 4 are similar in that part *a* addresses identifying a random variable, for which students are led to use its definition (intensive meaning) and perform the computations (actuative meaning), and part b addresses finding the expected value, for which, again, students are led to use its definition (intensive meaning) and perform the computations (actuative meaning). However, item 4 is different from item 3 in the sense that item 4 requires a deeper conceptual understanding (i.e., a stronger grasp of the intensive meaning) of random variables (part a) and expected value (part b). For item 4, the students needed to manipulate X^2 as opposed to X, using the definitions of random variables (part a) and expected values (part b), to determine whether X^2 is a random variable and to find the expected value. Although the students had encountered the expected value of mathematical expressions involving X^2 in class, the students had done more computations manipulating X than X^2. Hence, they were less familiar with the expected value of X^2 than of X. Therefore, they were able to set up $\sum P(x)$ (to ensure that the sum equals 1) and compute $\mathrm{E}(X)$ by habitual routine to answer item 3. But they would have needed a firm understanding of the definition of a random variable (that the sum of all values equals 1) and of the concept of the expected value of a random variable to set up $\sum P(x^2)$ and compute $\mathrm{E}(X^2)$ to answer item 4. Thus, the difference between these two items measured how well the students understood the intensive meaning of EVORV. The fact that the students did not do as well on item 4 (3/16 correct) as on item 3 (7/16 correct) suggests that an understanding of actuative meanings does not lead a learner to develop an understanding of intensive meaning. Only students who had developed a firm understanding of the definition and concepts of expected value would be able to do computations for $\mathrm{E}(X^2)$ successfully.

Items 9 and 10. Item 9 tested students' understanding of two types of meaning (ostensive and intensive) of the expected value and variance of a random variable. Item 10 examined their understanding of the ostensive meaning of a random variable by asking students to explicitly state the definition of a random variable. The two items follow:

9. Suppose that X is a random variable that ranges from a to b.

a. How is the expected value $\mathrm{E}(X)$ defined in terms of integrals?
b. How is the variance $\mathrm{Var}(X)$ defined in terms of expressions involving expected values?
c. How is the variance $\mathrm{Var}(X)$ defined in terms of integrals?
d. How are $\mathrm{E}(X^2)$ and $\mathrm{E}(X^3)$ defined in terms of integrals?
e. Is $\mathrm{E}(X)$ a function of X? Why?
f. Is $\mathrm{Var}(X)$ a function of X? Why?

10. What does it mean to you that a variable is a random variable?

Parts a, b and c of item 9 inquire about the ostensive meaning of the expected value and variance of a random variable: parts a and c involve stating the definitions (in mathematical expressions) for the expected values (and variances) of a random variable X; and part b requires recalling the formula $\mathrm{Var}(X) = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2$. As such, all three sub-items ask questions about ostensive meaning. In contrast, part d asks for the definitions (in mathematical expressions) for the expected values of random variables X^2 and X^3 as opposed to X. Although the students had seen mathematical expressions that involved X^2 and X^3 in class, such expected values had been discussed significantly less than the expected value of X. Because in part *d*, the students have to manipulate X^2 and X^3 as opposed to X, the contrast between the students' answers to part a and part d suggests the degree to which the students understood the *intensive meaning* of EVORV. If the students answered part d correctly, it could mean that their understanding of the definition of expected value was based on the underlying mathematical notion; that is, they had developed the intensive meaning as well as the ostensive meaning. Parts e and f of the same item intended to find out if the students misconceived the symbols $\mathrm{E}(X)$ and $\mathrm{Var}(X)$ as functions of a variable x (as in a variable in mathematics). Both $\mathrm{E}(X)$ and $\mathrm{Var}(X)$ are functions (or, more specifically, linear operators) of the class of random variables, but they should not be considered functions of a variable x.

Finally, item 10 intended to assess the students' ability to explicitly state their understanding of the definition or defining characteristics of a random variable. The students' class used the textbook *Engineering Statistics* (Montgomery, Runger, & Hubele, 2009). The textbook's description of a random variable is summarized in Table 7.

Table 7. *Summary of the Description of a Random Variable from the Textbook* Engineering Statistics *(Montgomery, Runger, & Hubele, 2009)*

Definition	A random variable is a numerical variable whose measured value can change from one replicate of the experiment to another. (p. 54)
Properties	1. $P(X \in R) = 1$, where R is the set of real numbers. 2. $0 \leq P(X \in E) \leq 1$ for any set E. 3. If $E_1, E_2, \ldots, E_k$ are mutually exclusive, $P(X \in E_1 \cup E_2 \cup \cdots \cup E_k) = P(X \in E_1) + \cdots + P(X \in E_k)$. (p. 56)

The three properties in Table 7 can be rephrased as: (1) the sum of the probabilities associated with all elements equals 1, (2) the probability of any event with respect to the variable is non-negative and less than or equal to 1, and (3) the probability of the union of mutually exclusive events equals the sum of the probabilities of the events. The properties serve as criteria for examining the validity of a random variable. In the class, the concept of *random variable* was introduced with the definition summarized in the table. However, only the first two properties in the table were used as criteria to determine the validity of a random variable. That is, the class almost treated these two properties as defining characteristics of a random variable. Therefore, in their responses to item 10, the students were expected to state these two properties either instead of or in addition to the textbook's formal definition. In this study, we refer to the formal definition and these two properties (as shown in the table) together as the *defining criteria* of the term. While item 10 assesses students' grasp of the ostensive meaning of the phrase "random variable," like part *a* of items 3 and 4, there is a difference in that item 10 targeted students' ability to state the definition's two critical features explicitly, while items 3 and 4 targeted the use of these two features in their determination of a random variable. The results of the sub-items in items 9 and 10 are summarized in Table 8.

Table 8. *Student Performance on Items 9 and 10*

Items or sub-items	9*a*	9*b*	9*c*	9*d*	9*e*	9*f*	10
Success rate	1/8	0/8	0/8	0/8	0/8	0/8	1/8

As shown in the table, the participants' performance on items 9 and 10 was poor in general. The poor performance on *a*, *b*, *c* and *d* of item 9 suggests that the students understood neither the ostensive meaning nor the intensive meaning. Only one student (Debora) answered 9*a* correctly, and she still gave incorrect answers for the other three parts. This indicates that for this student, the ostensive meaning of E(X) had not led to the intensive meaning. Also, the poor performance of the students on *e* and *f* of item 9 may suggest that the students had misconceptions about the symbols $\mathrm{E}(X)$ and $\mathrm{Var}(X)$: they viewed them as functions of a variable x (as in a variable in mathematics). The students' performance on item 10 in comparison with their performance on items 3 and 4 indicates that the students' ability to use critical features of the definition in their determination of a random variable did not mean they had the ability to explicitly state the definition.

In sum, the students in this study had apparently not developed knowledge of the ostensive meaning of *expected value of a random variable* and *random variable*. Although some students (4 out of 8 on 3*b* and 1 out of 8 on 4*b*) used the definition of the expected value when they needed it for computation, they were unable to provide the mathematical definition when they were asked to state it explicitly in 9*a*, 9*c* and 9*d*. Also, in sub-items 3*a* and 4*a*, many students (5 out of 8 on 3*a* and 6 out of 8 on 4*a*) were unable to use the definition of a random variable. Moreover, on item 10, only one student stated the definition of a random variable correctly. The students' poor performances on items 4, 9 and 10 and relatively better performance on item 3 suggest student challenges in making sense of the intensive meaning.

The Nature of Student Challenges

In this section, we shed light on the nature of the challenges students face in learning EVORV by analyzing the participants' comments made during the interviews and their written responses to the survey. We give a detailed explanation of the challenges arising from the lexical and semiotic features of EVORV (or $\mathrm{E}(X)$) in connection with the underlying mathematical notions and in comparison with its computational aspect.

Lexical Features of EVORV. The student responses on items 3, 4 and 10 and their comments on these items provide evidence that the students wrongly associate the statistics term "random" in "random variable" with either its meaning as in "random sampling" or the everyday meaning of the term. Excerpt 1 shows an example from Debora's interview.

Excerpt 1

Interviewer: Okay, number 3. 'Let X denote a variable that takes on any of the values -1, 0, 1 with respective probabilities $P(X=-1) = .2$, $P(X=0) = .5$, and $P(X=1) = .3$.' (Is X a random variable?)

Debora: I said yes.

Interviewer: Why do you think it is a random variable?

Debora: *Because X can be any of the three values*.

Debora's final comment (italicized) in Excerpt 1 shows that she associated "random" in "random variable" with the everyday meaning of the term (i.e., "without definite aim, direction, rule, or method"; *Merriam-Webster Online*, 2011). This is a sign of an undeveloped ostensive meaning of the phrase "random variable." Signs of an undeveloped ostensive meaning are further shown in the students' performances on item 10, which asked the students to explicitly state the definition of a random variable. Excerpt 2, from the interview conversation with Aaron, shows his misunderstanding of the term.

Excerpt 2

Interviewer: … what does it mean to you that a variable is a random variable? And you wrote 'a sample was taken randomly and unbiased and therefore the outcome of an experiment is not affected by the one conducting the experiment.' Okay.

Aaron: So, I am the person giving the survey to other people. *I design the survey in such a way that I am not affecting [inaudible] (the sampling design)*. So it's random.

In Aaron's comment (italicized), he associated the word "random" in "random variable" with the meaning of the word in "random sampling." These comments from Aaron and Debora indicate the challenge students face in developing the ostensive meaning of the statistics terms.

The interview and survey data further provide evidence that being unable to explicitly and correctly state the definition or defining criteria of "random variable" did not mean the students were unable to use the defining criteria. For example, John's survey responses to sub-items 3*a* and 4*a* contrasted with his answer to item 10. For part *a* of items 3 and 4, he wrote “yes, because they add to 1” and “yes. They add to 1,” respectively. Excerpt 3 shows the claims he made during the interview about sub-item 4*a*.

Excerpt 3

Interviewer: So X is a variable that takes on any of the values -1, 0, (and) 1. And probabilities are given (as shown). So is X a random variable?

John: My answer is yes. Um, *I know there are two conditions that it must satisfy. I honestly can’t remember the first condition, but I know the probabilities must equal one*. And in this case they do. So I would say yes.

John’s comment (italicized) shows that he knows the defining criteria for determining random variables (the first two properties in Table 7). However, when asked to state what “random variable” meant in item 10, he did not mention the criteria on the survey or during the interview. His interview comments are shown in Excerpt 4.

Excerpt 4

Interviewer: What does it mean to you that a variable is a random variable?

John: …it means *it’s an unbiased selection of a population*. I wrote sample, but I mean population. So it’s going to be biased in some way, so if you have a population of students. And you are trying to take random samples of students, you can’t just take all the students that receive A’s in the class, for example. …So random variable is taken from random students.

John’s comment (italicized) shows that, like Aaron, he associated “random variable” with the meaning of the word “random” in “random sampling.” Although John knew the defining criteria of “random variable” (he used them in his verbal description for sub-item 4*a*), he was unable to state the criteria explicitly when asked to do so. Taken together, John’s responses imply that a seemingly undeveloped understanding of ostensive meaning (as shown by the inability to utter the explicit definition) does not necessarily result in an inability to apply that ostensive meaning in context; that is, to use it as a criterion for determining random variables (i.e., showing understanding of the intensive meaning) or to do computations with examples (i.e., showing understanding of the actuative meaning).

Semiotic Features of $E(X)$. As mentioned earlier, the results for the sub-items 9*e* and 9*f* in the survey show that all eight students failed to make proper sense of the symbols $\mathrm{E}(X)$ and $\mathrm{Var}(X)$ by incorrectly viewing the symbols as functions of a variable X. Excerpt 5 shows what Debora wrote as answers to sub-items 9*e*, “Is $\mathrm{E}(X)$ a function of X? Why?” and 9*f*, “Is $\mathrm{Var}(X)$ a function of X? Why?” To minimize loss of information, we reproduce her handwritten responses.

Excerpt 5

Item	Debora’s Answer
9*e*:	yes. x is the independent variable and E(x) is the dependent variable that depends on x.
9*f*:	yes. x is the independent variable and Var(x) depends on x.

Debora’s written response to sub-item 9*e* implies that she knows the definition of a function. Discussing 9*e* in the interview, Debora gave the explanation in Excerpt 6.

Excerpt 6

Because X is…I don’t know how to explain. Because *if X is changing, the expected value of X is going to change as well*$_{(1)}$ because *the expected value of X at one value won’t be the same if X is a different value*$_{(2)}$, I guess.

Her comment (italicized, subscript (1)) in Excerpt 6 further shows that Debora holds a proper understanding of a function. However, her following comment (italicized, subscript (2)) indicates that she fails to make proper symbolic sense of $\mathrm{E}(X)$ and $\mathrm{Var}(X)$ by viewing X as a variable and the two symbols as functions of X. Because both $\mathrm{E}(X)$ and $\mathrm{Var}(X)$ are fixed constants determined by the parameters (such as μ and σ) of a population, they cannot be viewed as functions of X.

John's verbal response was similar to Debora's. In the survey, for sub-items 9*e* and 9*f*, John wrote, "yes, I don't have a good explanation" and "Yes," respectively. Excerpt 7 is from John's interview. His last remark in Excerpt 7 (italicized) indicates that, as in Debora's case, John viewed the symbols $\mathrm{E}(X)$ and $\mathrm{Var}(X)$ as functions of a random variable, where the variable takes on various values.

Excerpt 7

John:	… if I give values of X, we could get the same variances of X.
Interviewer:	What does it mean to you to say that you have different values of X here in statistics?
John:	Um, that would be a different random variable, different part of, different piece of the same [inaudible].
Interviewer:	Different data points? Or are you thinking like different random variables?
John:	No, *I was thinking different points with the same variable*.

Sub-items 9*e* and 9*f* intended to address the students' understanding of the underlying mathematical and structural notions; that is, the intensive meaning. The written and verbal descriptions in Excerpts 5 through 7 along with the student outcomes in Table 8 show that John, Debora and all the other students incorrectly viewed $\mathrm{E}(X)$ and $\mathrm{Var}(X)$ as functions of a variable. In particular, five of the eight participants, including John and Debora, believed that X would vary and that $\mathrm{E}(X)$ is a function of X depending on the value X takes. One of the other three (Simon) saw $\mathrm{E}(X)$ as changing depending on the sample taken. Two of the participants were unable to clearly explain their ideas. The students' performance on these two sub-items contrasts with their survey responses to items 3 and 4, which intended to address the students' understanding of the ostensive and actuative meanings of $\mathrm{E}(X)$. On these two items, some students successfully found the values of $\mathrm{E}(X)$ (3*b*) and $\mathrm{E}(X^2)$ (4*b*) (with success rates of 4 out of 8 for 3*b* and 1 out of 8 for 4*b*, as shown in Table 6). This contrast shows that while the students in general had failed to develop an understanding of the intensive meaning of these symbols, some students had developed an understanding of the ostensive and actuative meanings.

Mathematical Notions of EVORV. In addition to revealing how the students' sense-making in regard to EVORV was related with lexical and semiotic features, the qualitative analysis sheds light on how the students understood the underlying mathematical notions. We provide evidence of the students' computational accuracy when doing calculations for finding the expected value of (some expressions of) a random variable, which is a necessary procedure to determine the bias of estimators. This was shown in the students' answers to item 8 and their explanations of how they did the problem.

8. Let $x_1, x_2, \ldots, x_{10}$ be a sample of size 10 from a population with mean μ and variance σ^2. Consider the following expressions: $\hat{\theta}_1 = \frac{x_1 - x_2 + 3x_3}{3}$ and $\hat{\theta}_2 = \frac{x_1 + 2x_2 + 3x_3}{3}$.
a. Is $\hat{\theta}_1$ an unbiased estimator of μ?
b. Is $\hat{\theta}_2$ an unbiased estimator of μ?

Debora's answer to 8*a* was "yes." Excerpt 8 is from her interview conversation on this question. It also contains the work she did during the interview to show her understanding.

Excerpt 8

Interviewer: Can you explain why (you wrote yes to part *a*)?

Debora: Because I did this formula. I did it like that…
(writing) $E(\mu) = \frac{E(\mu) - E(\bar{\mu}) + 3E(\bar{\mu})}{3}$

Interviewer: What is it you are trying to show to be able to say yes for part *a*?

Debora: I am trying to show that the expected value of $\hat{\theta}_1$ is equal to the expected value of μ.

Interviewer: Okay. Then what is the expected value of $\hat{\theta}_1$?

Debora: $\bar{X}$, I think.

Interviewer: Okay

Debora: No. The expected value of $\hat{\theta}_1$ is…I think it is supposed to be μ's because it is the proportion that you are looking at. Since you get 3 over 3 that goes to 1…

Interviewer: Can you approach the problem by writing the expected value of $\hat{\theta}_1$ first?…

Debora: (writing) $E(\hat{\theta}_1) =$
I don't know what it means.

As shown in Table 5, seven of the eight participants did the calculation properly and gave correct answers to item 8. However, as illustrated in the conversation in Excerpt 8, when questioned about the meaning, none of the students were able to explain the process of the calculation for obtaining the expected value in connection with the meaning of *biased* or other relevant implications of the expected value. This indicates the contrast between the students' knowledge of actuative meaning and their knowledge of intensive meaning. The data show that most of the students possessed the actuative meaning of $\mathrm{E}(X)$, which they needed for the computations to find the expected values of a given *statistic* (an estimate, based on a sample of observed data, of a population parameter). However, they had failed to develop, and thus they failed to explain, knowledge of the intensive meaning of how the idea of weighted average yields a value or how the expected value of a statistic connects to the meaning of *biased.*

Further evidence of undeveloped intensive meaning comes from the students' verbal responses regarding item 5. Item 5 tested students' ability to manipulate symbols to find expected values and variances, which is an essential part of doing computations properly.

5. Suppose X is a random variable. Circle the statement if it is true in general.

a. $\mathrm{E}(aX) = a \cdot \mathrm{E}(X)$
b. $\mathrm{E}(X + b) = \mathrm{E}(X) + b$
c. $\mathrm{E}(X^2) = [\mathrm{E}(X)]^2$
d. $\mathrm{E}(aX) = a^2 \cdot \mathrm{E}(X)$
e. $\mathrm{Var}(aX) = a \cdot \mathrm{Var}(X)$
f. $\mathrm{Var}(X + b) = \mathrm{Var}(X) + b$
g. $\mathrm{Var}(aX) = a^2 \cdot \mathrm{Var}(X)$

Table 9 shows the detailed results for each of the seven statements of item 5 from the eight participants. In general, the students correctly determined which statements were valid, with success rates ranging from 4/7 to 7/7. This is evidence that the students in general possessed the foundational knowledge needed for developing actuative meaning. However, when the interviewer asked the students to provide the reasoning for their answers, they were generally unable to do so. In particular, while 7 of 8 student participants correctly determined the validity of statement 5*a*, no one was able to explain the underlying meaning of the formula. For example, John determined the validity of 5*a* correctly in the survey, but was unable to give his reasoning. To the interviewer's question, "Why do you think (the statement) *a* ($\mathrm{E}(aX) = a \cdot \mathrm{E}(X)$) is true?" John answered, "…because, um, if I am not mistaking it, the distributive property…I can remove

Table 9. *Student Performance on Item 5*

Participant / Statement	Debora	Simon	Sandy	John	Norma	Kara	Renata	Aaron	Success rate
a		×							7/8
b	×	×	×	×		×		×	2/8
c				×	×	×			5/8
d		×							7/8
e	×							×	6/8
f									8/8
g	×							×	6/8
Success rate	4/7	4/7	6/7	5/7	6/7	5/7	7/7	4/7	73.21%

Note: × indicates that the student determined the statement's validity incorrectly.

any number from that and still get the same number." In his answer, he does not explain how he understands the statement or why the statement is true. He simply paraphrases the statement $\mathrm{E}(aX) = a \cdot \mathrm{E}(X)$, making an association with a mathematics rule. The same pattern appeared again when he was asked to explain his answer to statement 5*c*, as shown in Excerpt 9.

Excerpt 9

Interviewer: …What about part *c* (of item 5)?

John: That is definitely not right now that [inaudible] when I look at it after I explained what my rationale for giving the other answer that doesn't match.

Interviewer: Can you tell me what the answers are [again]? …Can you explain why you think differently for *c*? You said based on what you've done…

John: Um, because if *the expected value of X squared is, um, different from the expected value of X, and if I take the expected value of X to square it, it's going to be a different number than the expected value of x squared random variable*.

Interviewer: How do you know this now? You didn't know this before?

John: [laughter] Um, I think it's by looking back at how I calculated the expected value.

Interviewer: Okay. All right. And you circled (statement) *g*. How do you know that that is true?

John: Um, *I don't know how I know, but I think I have seen that some time or been told that some time. But I think it might be explained with the…if you look at the integrals, you define the integrals a certain way.*

John's comments (italicized) in Excerpt 9 along with his other comments prior to the excerpt show his inability to explicitly explain how he knew whether the statements in item 5 were valid. Like John, the other students were not able to explicitly state how they knew the validity of the statements in item 5. That is, while many students are able to distinguish valid statements from invalid ones (evidence of developed actuative meaning), they often do not understand the mathematical notions that underlie the computational properties. This is evidence of undeveloped intensive meaning regarding EVORV and variance.

Discussion

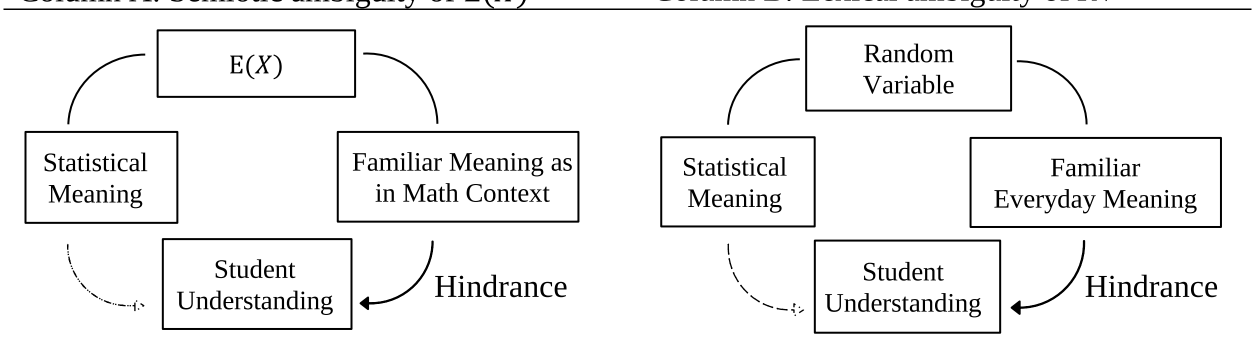
In an attempt to identify the challenges that calculus statistics students face in learning the statistics concept of EVORV, this study's research questions addressed (1) how students make sense of the phrase "the expected value of a random variable" and the symbol $\mathrm{E}(X)$ in conjunction with their lexical and semiotic features, and (2) how students connect the concept of EVORV to the underlying mathematical notions. Regarding the first research question, the study findings showed that the majority of the students, at the end of their calculus statistics course, were unable to make ostensive sense of the symbols $\mathrm{E}(X)$ and $\mathrm{Var}(X)$ and of the phrase "expected value of a random variable." Concerning this phrase, student comments such as "I design the survey in such a way that I am not affecting (the sampling design)" (Aaron, Excerpt 2) and "Because X can be any of the three values" (Debora, Excerpt 1) indicate that the student challenges in developing the ostensive meaning of "random variable" are associated with the lexical ambiguity in the phrase. This finding is consistent with Kaplan, Fisher and Rogness's (2010) claim that the lexical ambiguity of "random sample" hinders students' sense-making of the phrase. Concerning the symbol $\mathrm{E}(X)$, the results of the study showed that students generally had developed its ostensive meaning (i.e., being able to describe $\mathrm{E}(X)$ as the sum of the products of each possible value and its associated probability) and actuative meaning (i.e., successfully finishing the actual computation of the sum). However, the student responses to sub-items 9*a* and 9*d* provided evidence that most of the students failed to make ostensive sense of the symbol.

Regarding the second research question, the findings of this study suggest that most (if not all) of the students had failed to develop an understanding of the statistics concept in relation to its mathematical expression and underlying notions. That is, the students in general had not developed a proper understanding of the intensive meaning of EVORV. Evidence for this claim comes from (1) the responses to sub-items 9*a* and 9*d*, (2) the responses to item 9*e*, which reveal a misconception of viewing $\mathrm{E}(X)$ as a function of X, and (3) student comments after they found the value of $\mathrm{E}(X)$; for example, "I don't know what it ($\mathrm{E}(\hat{\theta}_1)$) means" (Debora; Excerpt 8). Further, we saw the widespread student misconception of $\mathrm{E}(X)$ as a function of X, as revealed by the student responses in item 9*e*. However, many students, without understanding the ostensive and intensive meanings, had grasped the actuative meaning (computational accuracy in finding the expected values), enabling them to answer computational questions about expected value.

The reasons for this contrast between the students' grasp of ostensive–intensive and actuative meanings seem likely to be related to the ambiguity inherent in the symbols and the phrase. The symbol $\mathrm{E}(X)$ involves semiotic ambiguity, so students develop misconceptions as they develop their understanding of the expected value of a random variable in relation to the symbol. The word "random" in the concept's name involves lexical ambiguity, as it is used differently in other statistics situations (as in random sampling) and in everyday life. These two types of student misconceptions – one of the symbol $\mathrm{E}(X)$, in which $\mathrm{E}(X)$ is viewed as a function of X, and the other of the phrase "random variable," in which it is associated with the everyday meaning of the word "random" – seem to share a common nature. The misconception of "random variable" is caused by its lexical ambiguity, and the likely cause of the misconception

Figure 1. *Students' Sense-making for the Symbol $E(X)$ and the Phrase "Random Variable" (RV)*

Column A: Semiotic ambiguity of $\mathrm{E}(X)$ — Column B: Lexical ambiguity of *RV*



of $\mathrm{E}(X)$ is the semiotic ambiguity inherent in the symbol. We present the similar nature of these lexical and semiotic ambiguities by juxtaposing them in Figure 1.

The symbol $\mathrm{E}(X)$ represents a linear operator of the set of all possible random variables, and is a fixed value once a random variable is determined. However, students' habitual perception of the symbol $f(x)$ as a function (or as a variable depending on the value of x) in mathematics contexts may have caused them to think of $\mathrm{E}(X)$ as a variable depending on the actual values taken by the random variable X. That is, the students' familiarity with the meaning of $f(x)$ in mathematics may hinder their understanding of $\mathrm{E}(X)$ in statistics due to the similar appearance of the two symbols (see Column A of Table 10). The student misconception regarding random variables (or random samples) is similar in that it is caused by lexical ambiguity: the familiar everyday meaning of the word "random" hinders student understanding of the phrases "random sample" (as shown by Kaplan et al., 2010) and "random variable" (as implied by this study) (see Column B of Table 10).

The findings of this study have implications for developing statistical literacy and reasoning at the calculus statistics level. According to Garfield (1999) and Rumsey (2002), developing statistical literacy at the elementary statistics level requires students to understand statistics terms and symbols. At the calculus statistics level, students are expected to go beyond this by making sense of statistics symbols and phrases in connection with the relevant mathematical notions. However, the study's findings suggest that students commonly develop misconceptions when making such connections with mathematics. The findings also provide evidence that such misconceptions might arise from the semiotic features of statistics symbols and lexical features of statistics phrases. Therefore, this study implies that instructors of calculus statistics should acknowledge the semiotic and lexical ambiguities of statistics symbols and terms. Furthermore, the study shows students' lack of statistical reasoning ability. Ben-Zvi and Garfield (2004) defined statistical reasoning at the elementary statistics level as an action of explaining statistical processes and the thinking process of interpreting statistical results. This study showed that the students lacked knowledge of the intensive meaning of EVORV: the participating students were unable to explain EVORV in connection with its underlying mathematical concepts. This finding implies the prevalence of a hindrance to developing statistical reasoning: insufficient development of the knowledge needed to understand and explain statistical processes and interpret statistical results based on mathematical concepts at the calculus statistics level.

In addition, the study's findings lay the foundation for further studies related to developing students' ostensive and intensive meanings in calculus statistics, and point to several other directions for future research. One avenue emerges from the premise that the concept of expected value of a random variable is critical for students to develop an understanding of bias. Along these lines, it is worth conducting a larger scale study to investigate whether students' understanding of expected value, including the particular misconceptions they hold due to ambiguous lexical and semiotic features, supports or hinders their understanding of bias. Another avenue of research emerges from the need to identify and document the relationship between students' understanding of foundational mathematical concepts such as functions and definite integrations and their attitudes toward calculus statistics or statistics in general. Such research could be based on the larger existing body of research in statistics education, but focus on a fine-grained description of how students' understanding of mathematical notions is associated with particular ways of conceiving statistics.

Finally, this study adopted a qualitative methodology, collecting data from a small number of students in a single class, in order to gain in-depth understanding and provide a detailed description of students' perceptions of EVORV. Although qualitative approaches can provide rich descriptions and in-depth explorations (Myers, 2000), they can also limit the extent to which a study's implications can be generally applied. To generalize the findings in this study to a larger statistics education community, it is essential to conduct larger scale research using more quantitative methodologies.

References

Aliaga, M., Cobb, G., Cuff, C., Garfield, J., Gould, R., Moore, T., Rossman, A., Stephenson, B., Utts, J., Velleman, P., & Witmer, J. (2005). *Guidelines for assessment and instruction in statistics education (GAISE) college report 2005*. Alexandria, VA: American Statistical Association. [Online: *http://www.amstat.org/education/gaise/GaiseCollege_full.pdf*]

Ben-Zvi, D., and Garfield, J. (2004). *The challenge of developing statistical literacy, reasoning and thinking*. Dordrecht, Netherlands: Kluwer Academic.

Carver, R., Everson, M., Gabrosek, J., Horton, N., Lock, R., Mocko, M., Rossman, A., Rowell, G. H., Velleman, P., Witmer, J., & Wood, B. (2016). *Guidelines for assessment and instruction in statistics education (GAISE) college report 2016*. Alexandria, VA: American Statistical Association. [Online: www.amstat.org/education/gaise]

Garfield, J. (1999). *Thinking about statistical reasoning, thinking, and literacy.* Paper presented at the First Annual Roundtable on Statistical Thinking, Reasoning, and Literacy (STRL-1). Kibbutz Be'eri, Israel.

Garfield, J., & Ben-Zvi, D. (2008). *Developing students' statistical reasoning: Connecting Research and Teaching Practice.* New York: Springer.

Garfield, J., & Chance, B. (2000). Assessment in statistics education: Issues and challenges. *Mathematical Thinking and Learning, 2*(1–2), 99–125.

Godino, J., & Batanero, C. (2003), "Semiotic functions in teaching and learning mathematics," in M. Godino, J., and Batanero, C. (2003). Semiotic functions in teaching and learning mathematics. In M. Anderson, A. Saenz-Ludlow, S. Zellweger, and V. Cifarelli (Eds.), *Educational perspectives on mathematics as semiosis: From thinking to interpreting to knowing* (pp. 149–168). Ottawa, ON: Legas.

Jung, S., & Hwang, J. (2015). Students' understanding of statistical terms having lexical ambiguity. In Dani Ben-Zvi and Katie Maker (Eds.), *The Teaching and Learning of Statistics: International Perspectives*. *Proceedings of the 12th International Congress on Mathematical Education. Proceedings of the 12th International Congress on Mathematical Education* (pp. 155–164). Seoul, Korea: Springer.

Kaplan, J., Fisher, D., & Rogness, N. (2010). Lexical ambiguity in statistics: How students use and define the words: association, average, confidence, random and spread. *Journal of Statistics Education*, *18*(1). [Online: http://www.amstat.org/publications/jse/v18n2/kaplan.pdf]

Kim, H., & Fukawa-Connelly, T. (2015). Challenges faced by a mathematically strong student in transferring his success in mathematics to statistics: A case study. *Journal of the Korean Society of Mathematics Education–Series A: The Mathematics Education, 54*(3), 223–240.

Kim, H., Fukawa-Connelly, T., & Cook, S. (2016). Student understanding of symbols in introductory statistics courses. In D. Ben-Zvi, & K. Makar (Eds.), *The Teaching and Learning of Statistics: International Perspectives* (pp. 163-174). Cham, Switzerland: Springer.

Lavy, I., & Mashiach-Eizenberg, M. (2009). The interplay between spoken language and informal definitions of statistical concepts. *Journal of Statistics Education*, *17*(1). [Online: http://www.amstat.org/publications/jse /v17n1/lavy.html

Mayén, S., Díaz, C., & Batanero, C. (2009). Students' semiotic conflicts in the concept of median. *Statistics Education Research Journal, 8*(2), 74–93.

Merriam-Webster Online. (2011). Random. Accessed October 10, 2014. Available at http://www.merriam-webster.com/dictionary/random

Mokros, J., & Russell, S. J. (1995). Children's concepts of average and representativeness. *Journal for Research in Mathematics Education*, 20-39.

Montgomery, D., Runger, G., & Hubele, F. (2009). *Engineering statistics*. New York: John Wiley & Sons.

Montiel, M. (2005). *The process of integration and the concept of integral: How does success with applications and comprehension of underlying notions such as accumulation relate to students' mathematical fluency*? (Unpublished doctoral dissertation). University of New Hampshire, Durham.

National Council of Teachers of Mathematics (NCTM) (2000). *Principles and standards for school mathematics*. Reston, VA: NCTM.

Pfannkuch, M., Budgett, S., Parsonage, R., & Horring, J. (2004). *Comparison of data plots: Building a pedagogical framework.* Paper presented at the Tenth International Congress on Mathematics Education (ICME-10), Copenhagen, Denmark, 4–11 July.

Rumsey, D. J. (2002). Statistical literacy as a goal for introductory statistics courses. Journal of Statistics Education, *10*(3). Retrieved July 15, 2014 from [Online: http://www.amstat.org/publications/jse/v10n3/rumsey2.html]

Simpson, G. B. (1981). Meaning dominance and semantic context in the processing of lexical ambiguity. *Journal of verbal learning and verbal behavior, 20*(1), 120–136.

Walpole, R., Myers, R., Myers, S., & Ye, K. (2014), *Probability and statistics for engineers and scientists*, (9th ed.) Boston, MA: Pearson Education.

Watkins, J. C. (2010). On a calculus-based statistics course for life science students. *CBE-Life Sciences Education*, *9*(3), 298-310.

Wild, C. & Pfannkuch, M. (1999). Statistical thinking in empirical inquiry. *International Statistical Review 67*(3), 223–265.

Yates, D., Moore, D., & Starnes, D. (2008). *The practice of statistics* (3rd ed.). New York: W. H. Freeman and Company.