A Case Study on Teaching the Topic "Experimental Unit" and How it is Presented in Advanced Placement Statistics Textbooks

Jamis J. Perrett
Texas A&M University

Key Words: Unit of analysis; Teacher preparation; College Board.

Abstract

This article demonstrates how textbooks differ in their description of the term *experimental unit*. Advanced Placement Statistics teachers and students are often limited in their statistical knowledge by the information presented in their classroom textbook. Definitions and descriptions differ among textbooks as well as among different editions of the same textbook. Furthermore, many schools use older editions of textbooks rather than current editions that contain updated information and thus lose the benefit of improved discussions and clarifications. Advanced Placement Statistics teachers should be aware of this issue and seek additional training through workshops, additional textbooks, and webinars to increase and strengthen their knowledge and understanding of key statistical concepts. Textbook authors should be aware of teachers' dependence on the authors' presentation of topics and ensure that key topics like *experimental unit* are covered thoroughly. This article considers three prior Advanced Placement Statistics exam questions to illustrate how different Advanced Placement Statistics textbooks may have influenced students' answers based on the textbooks' authors' treatment of *experimental unit*.

1. Introduction

An important aspect of the analysis of experiments and studies is the identification of the experimental unit, which in a basic single-level analysis would also be considered the "unit of analysis" (see Smolkowski, 2012 for further discussion on the nomenclature and an extensive list of reference papers). The experimental unit may be an individual or a group of individuals. What constitutes the experimental unit is determined by how the treatments are assigned. For

example, consider a simplistic study in which two different training methods are compared. One at a time, company engineers sit at a computer. The computer program provides one of two demonstrations, selected at random, as to how a particular task can be accomplished. Of the 40 total engineers, 20 engineers are taught individually using method A and the other 20 are taught individually using method B. Because the teaching occurs with each individual, independent of each other individual, the instruction to each individual is considered independent and individual engineer is considered the experimental unit for a total of 40 experimental units. Consider another simplistic example in which 20 engineers meet together in a room and method A is demonstrated using a projection system to the group all at once. Next, the other 20 engineers meet together in the room and method B is demonstrated using the projection system to the group all at once. In that case, the training that takes place among engineers within one of the groups is considered non-independent as the engineers were together when the training was received. Consequently, the group of 20 engineers is considered the experimental unit for a total of two experimental units. The two studies are characteristically different and consequently would be analyzed differently.

Misidentification of the experimental unit can lead to grossly inflated Type-1 error rates in hypothesis testing (Blair, 1983). It is not an uncommon mistake. A systematic review of medical journals discovered a misidentification error rate of 44% (Calhoun, Guyatt, Cabana, Lu, Turner, Valentine, and Randolph 2008). The editor-in-chief of *Research Quarterly for Exercise and Sport* met with reviewers and section editors for their journal and determined that misidentification of the unit of analysis was enough of a problem to warrant an editorial in the journal (Silverman, 2004). Also, the editorial board of the *Journal of Teaching in Physical Education* invited a researcher to present a treatise on the unit of analysis issues as well as to suggest appropriate ways to confront them by describing related topics such as treatments, random assignment, etc. (Silverman and Solmon, 1998). Mundfrom (2009) also provided a treatise on the subject of experimental unit, its identification and misidentification. Of course, in some studies the unit of analysis can be the experimental unit, but in other physical education studies a multi-level design may have more than one unit of measurement. However, such complexities go beyond what is required for Advanced Placement (AP) Statistics.

A student's understanding of statistical terms comes from several sources, including the teacher's explanation and the way the term is defined and explained in the student's textbook. If a teacher is unfamiliar with a term from prior training, the teacher as well may derive the explanation from the class textbook. This is particularly true of AP Statistics teachers, who may not have received any or sufficient formal training in statistics. This places great importance on how a term is presented in a textbook. As there are several different textbooks available for use in an AP Statistics class, if a statistical term is presented in different ways in different textbooks, students may derive different understandings of that term.

2. The AP Statistics Exam

Each year in the spring the AP Statistics exam is administered to students across the United States as well as in other countries. The AP Statistics exam is a standardized exam intended to test the student's proficiency in topics that would be covered in an introductory statistics course in college. In many cases, colleges will accept a passing grade on the AP Statistics exam as a

replacement for taking the equivalent course at the college. The College Board handles the production of the exam and enlists the aid of high school teachers and college faculty for the development of the exam. How a term is expected to be used in an AP Statistics exam question may not match how it is presented in different textbooks used for AP Statistics. The College Board does create a list of topics that the exam may cover (College Board, 2010a), including the term *experimental unit*, but does not provide a universal curriculum for learning these topics. Consequently, it is conceivable that some students may miss a question on the AP Statistics exam because of the way a statistical term was taught to them, as presented in the textbook used in their class.

3. Defining Experimental Unit

The experimental unit may be defined as the object independently treated in an experiment. There are different aspects of the explanation of the term experimental unit. The experimental unit may be an animate object like a person or an inanimate object like a building. The experimental unit may consist of one object or several (recall the examples of the two training methods and the 40 engineers from the Introduction). The way experimental unit is taught and learned affects a student's conceptual understanding of the unit that is actually being experimented on. The word "independently" may be missing from the experimental unit definitions in textbooks but is nonetheless an essential part of how an experimental unit is defined. Independence among experimental units is an essential feature of an experiment aimed at proving cause and effect. The word "independently" also aids in the identification of the experimental unit as treatments applied independently to units in a group make the units the experimental units, but treatments applied to a group of units together makes the entire group a single experimental unit.

In addition to the term experimental unit, there are other terms related to experiments that can be difficult for students to understand. A misunderstanding of one of these terms can result in a misunderstanding of related terms. For example, misidentification of experimental unit could lead a student to use the wrong degrees of freedom. Misidentification of the treatments of an experiment or how they are applied could lead a student to identify the multiple individuals that make up a single experimental unit as experimental units themselves. These additional terms related to experiments are not explicitly addressed in this article, but include treatments, control groups, randomization, and replication.

3.1 How Degrees of Freedom are Computed

The appropriate calculation of degrees of freedom for an inferential procedure depends on an understanding of the experimental unit. Consider the example of the engineers being trained, described in the Introduction. Assume there is interest in using a two-sample independent equal variance t-test to compare the two methods. The general formula for degrees of freedom in this case is $df = c(g - 1)$, where c represents the number of conditions being tested (two conditions in this case, the two training methods) and g represents the number of experimental units within each condition (20 engineers for each training method). When the individual engineers are considered the experimental units, there are $2(20 - 1) = 38$ degrees of freedom for the test. However, when the groups of 20 engineers are considered the experimental units, there are

$2(1 − 1) = 0$ degrees of freedom for the test. There isn't a traditional t distribution with zero degrees of freedom. Consequently, a critical value or a P value for such a test cannot be computed. As a consequence, the proposed t-test cannot be performed and an alternate form of analysis, if any, must be used.

3.2 Five Popular AP Statistics Textbooks

Consider the ways the following five AP Statistics textbooks define experimental unit:

Bock, Velleman, and De Veaux (2010), hereafter referred to as BVD, defines experimental unit as "the individuals on whom or which we experiment" (p. 294). It also discusses the need to "assign our participants to [their] treatments" (p. 294), indicating that participant is simply another name for an experimental unit. This definition and description are consistent across the first three editions of BVD (2004, 2006, 2010), the third being the current edition.

The BVD book also includes a section on replication but does not address the unit of analysis issue, which essentially determines whether the experimental unit is the unit on which observations are recorded or whether the experimental unit may consist of a group of such units.

Peck, Olsen, and DeVore (2012), hereafter referred to as POD, takes a more traditional and technical approach to the definition of an experimental unit defining it as "the smallest unit to which a treatment is applied" (p. 68). A further explanation is given regarding the definition of replication with respect to the experimental unit. This definition and further explanation are consistent across the different editions of POD (2000, 2005, 2010, and 2012).

POD gives a warning that researchers ensure replication in any experiment. Then an example is given in which two classes of students are used in a study in which one class is assigned teaching method 1 and the other class is assigned teaching method 2 (p. 68). POD explains that in such a design, the *classes* are the experimental units rather than the *students*. The idea that a group of individuals (students) may constitute the experimental unit rather than an individual (student) may not be well understood by all AP Statistics students. This concept is not discussed in BVD in either the section on experimental unit or the section on replication.

The Practice of Statistics, hereafter referred to as TPS, has recently released an extensively-revised fourth edition (Starnes, Yates, and Moore 2012). In the third edition (Yates, Moore, and Starnes 2008), the experimental unit is defined as "the individuals on which the experiment is done" (p. 354). Similar to BVD, the experimental unit is associated with an individual. However, in the fourth edition of TPS, the definition of an experimental unit is presented similarly to the definition in POD: "the smallest collection of individuals to which treatments are applied" (p. 233) in a section on Experiments. The following page (p. 234) includes an example that identifies middle schools as experimental units, thus addressing the unit of analysis issue.

Agresti and Franklin (2009), hereafter referred to as AF, defines the experimental units as "the subjects-the people, animals, or other objects to which the treatments are applied" (p. 180) with mention of schools or stores as being possible experimental units (p. 175), but without further

clarification on the unit of analysis issue. The definition and discussion in the first edition of AF (2006) is essentially the same as in the second/current edition.

Watkins, Scheaffer, and Cobb (2008), hereafter referred to as WSC, defines the experimental unit as "the people, animals, families, classrooms, and so on to which treatments are randomly assigned" (p. 252). The WSC book, like the POD book, discusses the situation in which a classroom may be considered the experimental unit rather than individual students (p. 252). The definition and discussion in the first edition of WSC (2004) is nearly the same as in the second/current edition.

The following table summarizes the definitions for the different textbooks:

Table 1: The definitions of Experimental Unit for Various AP Statistics Textbooks, and Related Comments

Book	Definition	Comment
BVD 2010	"the individuals on whom or which we experiment" (p. 294).	Focus is on the individual, not a group of individuals
POD 2012	"the smallest unit to which a treatment is applied" (p. 68)	Does not focus on either individual or a group
TPS 2012	"the smallest collection of individuals to which treatments are applied" (p. 233)	Mention is made of both individuals and groups
TPS 2008	"the individuals on which the experiment is done" (p. 354)	Focus is on the individual, not a group of individuals
AF 2009	"the subjects-the people, animals, or other objects to which the treatments are applied" (p. 180)	Focus is on the individual, not a group of individuals
WSC 2008	"the people, animals, families, classrooms, and so on to which treatments are randomly assigned" (p. 252)	Mention is made of both individuals and groups

Regardless of how experimental unit is defined and discussed in the edition or version of a textbook used by an AP Statistics class, it is important for the teacher to understand experimental unit and how it is used in practice. It is also important for the teacher to teach it in such a way that students receive a comprehensive and proper understanding of the unit of analysis issue.

3.3 Old Textbook Editions Still in Classrooms

The reason it is relevant to discuss multiple editions of the textbooks is that there will likely always be schools using older editions of AP Statistics textbooks. One reason for using older editions is the effort that it takes teachers to change to a new edition of the text when that new edition may have very few improvements over the previous edition. Another reason for using older editions is due to state contracts with the publishers as well as lack of funds to purchase new sets of books.

Incomplete sales and contract data for W. H. Freeman and Company, publishers of the TPS books, show eight outstanding state contracts for second and third editions of the TPS book that do not expire until 2012 or later. These records also show that during the first part of 2010, five school districts purchased copies of the first edition of TPS (Yates, Moore, and McCabe 1999), 47 school districts purchased copies of the second edition of TPS (Yates, Moore, and Starnes 2003) (though one of those school districts has pre-ordered the new fourth edition of TPS), and 253 school districts purchased copies of the third edition of TPS. In all, this represents 9,779 old editions of TPS purchased by school districts from the publisher during just the first part of 2010 (C. Bleyer - W.H. Freeman and Company, personal communication, November 1, 2010). These sales figures do not include the additional purchases by retail and wholesale merchants.

In addition to those school districts purchasing additional copies of the old editions of TPS, the AP Statistics Listserv (College Board 2010c, search="books") shows additional requests by individual teachers for copies of old editions of AP Statistics textbooks. Between the months of March and September 2010 there were six requests by AP Statistics teachers for a total 134 copies of the second edition of TPS to supplement their existing copies for their growing class sizes in Fall 2010. Some teachers who are unable to purchase new editions of books for their students due to budget constraints ask other teachers who are no longer using their old editions to send such editions to them for the cost of the shipping. There were also two requests for 40 to 50 total copies of the first edition of BVD, a request for an unspecified number of copies of the second edition of POD, and even a request for twelve copies of the first edition of TPS, which was published more than ten years ago.

4. AP Statistics Free-response Exam Questions

From time to time the high school AP Statistics exam includes a free-response question that involves the understanding of the statistical term experimental unit. Three such questions have been identified.

4.1 AP Statistics Exam 2010 Free-Response Question 1: European Starlings

The first free-response question on the 2010 AP Statistics exam (College Board, 2010b) hereafter referred to as "APSQ1," is as follows:

> "Agricultural experts are trying to develop a bird deterrent to reduce costly damage to crops in the United States. An experiment is to be conducted using garlic oil to study its effectiveness as a nontoxic, environmentally safe bird repellant. The experiment will use European starlings, a bird species that causes considerable damage annually to the corn crop in the United States. Food granules made from corn are to be infused with garlic oil in each of five concentrations of garlic—0 percent, 2 percent, 10 percent, 25 percent, and 50 percent. The researchers will determine the adverse reaction of the birds to the repellant by measuring the number of food granules consumed during a two-hour period following overnight food deprivation. There are forty birds available for the experiment, and the researchers will use eight birds for each concentration of garlic. Each bird will be kept in a separate cage and provided with the same number of food granules."

Students are asked to identify the experimental unit.

Several answers were given on APSQ1, including "bird," "the 40 birds," "food granules," "5 groups of 8 birds," and "cages." The College Board's rubric gives the following solution: "The experimental units are the birds (starlings), each placed in an individual cage." Also, a follow-up note indicates that "it is acceptable to identify the experimental units as the cages" (College Board, 2010d, pp. 1-2). So, credit was given to students who recognized there were 40 experimental units rather than five and that the experimental units were referring to the bird or the cage, but not the food granules. So, answers like "bird," "the 40 birds," and "cages" received credit, but answers like "food granules" and "5 groups of 8 birds" did not. The grading of the AP Statistics free-response questions is holistic in the sense that a limited number of points are available for each question. These points are divided up among the different sections of the question, but there is often not the same number of sections as there are possible points.

Identifying the experimental unit in a study can be quite challenging, even for a student who is well trained. How experimental unit is presented in a textbook may either help or hinder a student's understanding of experimental unit.

All the textbooks mentioned in this study could lead students to the correct answer of "bird" or "the 40 birds." However, a misunderstanding of any of the textbooks could also lead to the acceptable but less desirable answer "cages" or the incorrect answer "food granules." This is so because words like "individuals" in the BVD and third edition of TPS descriptions strongly support the understanding of living beings such as birds as experimental units. Perhaps it would seem unnatural to identify cages, food granules, or even groups (i.e., "group" vs. "individual") of birds as experimental units with such a word. The fourth edition of TPS changes the wording from "individuals" to "collection of individuals," continuing the idea of living beings but now allowing for "a group of individuals" to be thought of as experimental units, perhaps leading some students to mistakenly identify "5 groups of 8 birds" as the experimental units.

The word "unit" used by POD and further descriptions such as "the subjects-the people, animals, or other objects to which the treatments are applied" from AF and WSC allow for experimental units to be inanimate objects, perhaps leading students to identify things such as "cages" and "food granules" as experimental units. "Food granules" might be considered because one might reason that the different concentrations of garlic oil, the treatments, are being "applied" (wording used in most of the textbooks' definitions of experimental unit) to the food granules as they actually "physically" are being mixed into the food granules. However, in fact, the granules are merely a method of delivery of the treatments to the experimental units, the birds.

Unfortunately, in APSQ1 garlic concentrations may be considered by students as being applied to garlic oil, garlic oil with different concentrations of garlic as being applied to food granules, and food granules infused with different concentrations of garlic as being applied to each cage containing each bird. It can be quite confusing. That is why conceptually the details and purpose of the study must be considered. The study is about repelling birds with the smell and taste of garlic. Cages and food granules are involved, but they are mediums used only to accommodate the study. Again, the study is about repelling birds with garlic. So, it is the different garlic

concentrations (the treatments) that are being served (independently) to the individual birds (the experimental units).

4.2 AP Statistics Exam 2006 Free-Response Question 5: Tiger Shrimps

In some experiments, the experimental units are in fact groups rather than individuals. One specific example is used in the fifth free-response question on the 2006 AP Statistics Exam (College Board, 2006a) hereafter referred to as "APSQ5."

> "A biologist is interested in studying the effect of growth-enhancing nutrients and different salinity (salt) levels in water on the growth of shrimps. The biologist has ordered a large shipment of young tiger shrimps from a supply house for use in the study. The experiment is to be conducted in a laboratory where 10 tiger shrimps are placed randomly into each of 12 similar tanks in a controlled environment. The biologist is planning to use 3 different growth-enhancing nutrients (A, B, and C) and two different salinity levels (low and high)."

Because the treatments, salinity levels and growth-enhancing nutrients, were added to the tank, the 10 shrimps placed in each tank were treated collectively rather than individually (independently). Consequently, the experimental unit was the group of 10 shrimps in each of the 12 tanks (or the experimental units could be referred to as the "12 tanks"—analogous to the "cages" in APSQ1) rather than the 120 individual shrimps, resulting in 12 total experimental units rather than 120. In APSQ5, students were asked to describe the experimental design, which should involve identification of the experimental units. The scoring rubric (College Board, 2006b) indicates that if students discuss randomization of treatments applied to the shrimps rather than to the tanks, that portion of the problem is considered incorrect.

Because the books BVD, TPS (first three editions), and AF focus on the experimental unit being an individual, it would seem unusual for students using one of these books to correctly identify a tank of 10 shrimps as the experimental unit. The wording of these books seems to support the identification of a single shrimp as the experimental unit.

The discussion of POD and TPS (fourth edition) regarding the possibility of the experimental unit being a group of individuals supports students using one of these books correctly identifying a tank of 10 shrimps as the experimental unit. Multiple shrimps in a tank can be seen as similar to multiple students in a classroom.

Again, it is unclear how students using the WSC book would respond.

4.3 AP Statistics Exam 2006 Form B Free-Response Question 5: Tractor Plow Hitches

In the fifth free-response question on Form B of the 2006 AP Statistics Exam (College Board, 2006c) hereafter referred to as "APSQ5b," a comparison of two different hitches used to connect a tractor to a plow was employed to determine if the new hitch would help reduce draft more than the standard hitch.

"When a tractor pulls a plow through an agricultural field, the energy needed to pull that plow is called the draft. The draft is affected by environmental conditions such as soil type, terrain, and moisture.

"A study was conducted to determine whether a newly developed hitch would be able to reduce draft compared to the standard hitch. (A hitch is used to connect the plow to the tractor.) Two large plots of land were used in this study. It was randomly determined which plot was to be plowed using the standard hitch. As the tractor plowed that plot, a measurement device on the tractor automatically recorded the draft at 25 randomly selected points in the plot.

"After the plot was plowed, the hitch was changed from the standard one to the new one, a process that takes substantial amount of time. Then the second plot was plowed using the new hitch. Twenty-five measurements of draft were also recorded at randomly selected points in this plot."

Part A of that question specifically required students to identify the experimental unit for the study. Part C of that question required students to determine if replication was used properly in the study. For both parts, students needed to understand the concept of experimental unit and the unit of analysis issue.

For Part A, the scoring rubric (College Board, 2006d) indicates that not identifying the experimental units as the two plots will result in that part being counted as only "partially correct." As well, Part C will be considered "incorrect" if the student "incorrectly argues that the 25 measurements taken on each experimental unit (plot) provide proper replication."

There are three specific aspects to this free-response problem that contribute to complexity with respect to identifying the experimental unit. The first aspect is the fact that there is subsampling. At each of 25 different locations on each plot, measurements were taken, leading some to erroneously believe that the 25 different locations represent 25 experimental units. The second (related) aspect is the fact that this experiment is unreplicated. There is only one experimental unit, plot of land, for each treatment level, hitch (new vs. old). Students may expect any experiment described in a free-response question to have replication. Third, students may be looking for an animate object or something close to it, like a tractor or a hitch, to be the experimental unit, not an inanimate plot of land.

Only the POD book does not use any of the words "individual," "people," or "subjects" in the definition of experimental unit. These words appear to imply an animate object or something close to it like a tractor or a plow hitch. Also, the books that only discuss experimental units as individuals, like BVD, TPS (first three editions), and AF, rather than the possibility of experimental units as groups, may be more likely to consider the individual measurement units, the 25 locations, as the experimental units.

5. Survey Results

A small-scale survey was conducted to get an initial idea as to how textbook wording might influence the responses to questions aimed at identifying the experimental unit. Ninety-seven undergraduate students in an introductory statistics course at a major research university were each provided with photocopies of the pages of an AP Statistics textbook that related to identifying the experimental unit and replication. Each student was also provided with all three of the above AP Statistics Free Response questions: APSQ1, APSQ5, and APSQ5b. In addition to the question stem, two questions were given:

1. Identify the experimental units.
2. How many experimental units are there in this example?

The first question was already a part of the AP Free Response questions. The second question was added to clarify potentially vague answers provided for the first question. For example, giving "shrimp" as the answer to APSQ5 was not enough information to determine whether the student knew the answer was the 12 groups of 10 shrimp or if the student incorrectly assumed each individual shrimp was an experimental unit. So, if a student answered "shrimp" to Question 1 and "12" to Question 2, it was assumed the student understood the correct answer. However, answering "shrimp" to Question 1 and "120" to Question 2 led to the assumption that the student thought that one individual shrimp was an experimental unit. The order in which the Free Response questions appeared was varied across and within surveys for each textbook. There were 18 surveys for each textbook. The surveys were passed out to students in such a way as to ensure an equal number of surveys for each textbook were completed.

The first table, Table 2, was created to show the proportion of correct answers (based upon answers to questions 1 and 2) across textbooks. Proportions appear as percentages for clarity.

Table 2: Proportion of Students who Answered Correctly.

Question	Book					
	AF (n=15)	BVD3 (n=14)	POD4 (n=17)	TPS3 (n=16)	TPS4 (n=18)	WSC (n=14)
APSQ1 (Bird)	73.3%	80.0%	35.3%	87.5%	50.0%	21.4%
APSQ5 (Shrimp)	0.0%	0.0%	47.1%	6.3%	22.2%	42.9%
Total	36.7%	40.0%	41.2%	46.9%	36.1%	32.1%

As expected, for the textbooks that give a thorough explanation of replication and the potential for groups being the unit of analysis (POD4, TPS4, and WSC), scores are higher for APSQ5, which deals with the experimental unit being a group of individuals. As well, textbooks that give little discussion regarding the possibility of the experimental unit being a group of individuals (AF, BVD3, and TPS3) have higher scores for APSQ1, which deals with experimental units that are individuals. It is not surprising that students reading AF, BVD3, and TPS3 did not perform well on APSQ5 as those textbooks did not include discussion on the unit of analysis potentially consisting of a group of individuals as is the case in APSQ5.

The next table, Table 3, categorizes responses as choosing a "group" of individuals as the experimental unit or "individuals" as the experimental unit. All responses that did not fall into one of those two categories were excluded.

Table 3: Proportion of Students for Each Question who Answered "Individual" (vs. "Group") for Each Book and Each Question.

Question	Book					
	AF (n=15)	BVD3 (n=14)	POD4 (n=17)	TPS3 (n=16)	TPS4 (n=18)	WSC (n=14)
APSQ1 (Bird)	100%	92.3%	50%	100%	60%	33.3%
APSQ5 (Shrimp)	60%	35.7%	0%	57.1%	17.6%	11.1%

In Table 3, for the Bird question a percentage close to 100% would be ideal and for the Shrimp question a percentage close to 0%. Because of the low numbers of surveys used for the above tables, it is perhaps best to consider percentages above 50% as supportive of individuals as experimental units and percentages below 50% as supportive of groups of individuals as experimental units. Using that criteria, BVD3 and TPS4 seem to perform well. AF and TPS3 appear to lean too far toward individuals, and WSC and POD appear to lean too far toward groups.

Since BVD3 only identifies "individuals" as experimental units, it was assumed that for APSQ5b students might be less apt to provide an inanimate object "plot" as an experimental unit as opposed to an object that might seem more animate like "tractor." However, more (53.8%) of the BVD3 students gave "plot" as the experimental unit than those who gave "tractor" as the experimental unit. Also, there were other textbooks with more "tractor" answers than BVD3. So, the assumption that students might be less apt to provide an inanimate object as an answer to a question given their textbook only identified "individuals" as experimental units was not supported by survey results. However, as both "plot" and "tractor" are inanimate objects, APSQ5b was not an ideal question for determining whether or not students would consider the word "individuals" to represent animate objects.

6. Conclusions and Recommendations

As illustrated in this study, it is very important for students of an AP Statistics class to understand the concept of experimental unit because their understanding can impact their performance on the AP Statistics exam and their ability to use statistics appropriately in their futures. The way the concept of experimental unit is defined and used also has impact on the way data are analyzed and consequent results are explained. As different textbooks and even different editions of the same textbook may explain statistical concepts differently, it is important that AP Statistics teachers gain a comprehensive understanding of the statistical topics they teach and use resources in addition to their class textbook, which can include additional textbooks, other teachers, workshops, and webinars.

Textbook authors should provide ample discussion of unit of analysis with examples of experiments for which experimental units are individuals and groups to help teachers and students gain a more comprehensive understanding of experimental unit.

The importance of students understanding statistical concepts is not limited to that of experimental units. It is generally expected that AP Statistics teachers, or any statistics teacher for that matter, should gain a comprehensive understanding of all the statistical concepts covered in their courses. Textbook authors should be comprehensive in their presentations of statistical concepts. However, there are some key topics like experimental unit that tend to be more difficult for students to adequately understand. Such topics merit special attention to enable students to excel on their AP Statistics exams as well as in their future statistical endeavors.

References

Agresti, A. and Franklin, C. (2006), *Statistics The Art and Science of Learning from Data*, Upper Saddle River, NJ: Pearson Education Inc.

Agresti, A. and Franklin, C. (2009), *Statistics The Art and Science of Learning from Data* (2nd ed.), Upper Saddle River, NJ: Pearson Education Inc.

Blair, R. C., Higgins, J. J., Topping, M. E. H., and Mortimer, A. L. (1983), "An Investigation of the Robustness of the t Test to Unit of Analysis Violations," *Educational and Psychological Measurement*, 43: 69.

Bock, D. E., Velleman, P. F., and De Veaux, R. D. (2004), *Stats: Modeling the World*, Boston, MA: Pearson Education Inc.

Bock, D. E., Velleman, P. F., and DeVeaux, R. D. (2006), *Stats: Modeling the World* (2nd ed.), Boston, MA: Pearson Education Inc.

Bock, D. E., Velleman, P. F., and DeVeaux, R. D. (2010), *Stats: Modeling the World* (3rd ed.), Boston, MA: Pearson Education Inc.

Calhoun, A. W., Guyatt, G. H., Cabana, M. D., Lu, D., Turner, D. A., Valentine, S., and Randolph, A. G. (2008), "Addressing the Unit of Analysis in Medical Care Studies: A Systematic Review," *Medical Care*, (46) 6: 635-643.

College Board (2006a), "AP Statistics 2006 Free-Response Questions." Available at http://apcentral.collegeboard.com/apc/public/repository/_ap06_frq_statistics_51653.pdf. [Accessed 28-May-2012]

College Board (2006b), "AP Statistics 2006 Scoring Guidelines." Available at http://apcentral.collegeboard.com/apc/public/repository/_ap06_statistics_sg_revised.pdf. [Accessed 28-May-2012]

College Board (2006c), "AP Statistics 2006 Free-Response Questions: Form B." Available at http://apcentral.collegeboard.com/apc/public/repository/_ap06_frq_statistics__51654.pdf. [Accessed 28-May-2012]

College Board (2006d), "AP Statistics 2006 Scoring Guidelines (Form B)." Available at http://apcentral.collegeboard.com/apc/public/repository/_ap06_statistics_sg_formb.pdf. [Accessed 28-May-2012]

College Board (2010a), "Statistics Course Description: Effective Fall 2010." Available at http://apcentral.collegeboard.com/apc/public/repository/ap-statistics-course-description.pdf. [Accessed 28-February-2012.]

College Board (2010b), "AP Statistics 2010 Free-Response Questions." Available at http://apcentral.collegeboard.com/apc/public/repository/ap10_frq_statistics.pdf. [Accessed 28-May-2012]

College Board (2010c), "AP Statistics Listserv." Available at http://lyris.collegeboard.com/read/?forum=ap-stat. [Accessed 3-November-2010.]

College Board (2010d), "AP Statistics 2010 Scoring Guidelines." Available at http://apcentral.collegeboard.com/apc/public/repository/ap10_statistics_q1.pdf. [Accessed 28-February-2012.]

Mundfrom, D. J. (2009), Chapter 21: Can We Make a Silk Purse from a Sow's Ear?, *Quality Research in Literacy and Science Education: International Perspectives and Gold Standards*, (Shelley II, M. C., Yore, L. D., and Hand, B., eds.), New York, NY: Springer, pp. 430.

Peck, R., Olsen, C., and DeVore, J. (2000), *Introduction to Statistics and Data Analysis*, Belmont, CA: Brooks/Cole.

Peck, R., Olsen, C., and DeVore, J. (2005), *Introduction to Statistics and Data Analysis* (2nd ed.), Belmont, CA: Brooks/Cole.

Peck, R., Olsen, C., and DeVore, J. (2010), *Introduction to Statistics and Data Analysis* (3rd ed.), Belmont, CA: Brooks/Cole.

Peck, R., Olsen, C., and DeVore, J. (2012), *Introduction to Statistics and Data Analysis* (4th ed.), Belmont, CA: Brooks/Cole.

Silverman, S. and Solmon, M. (1998), "The Unit of Analysis in Field Research: Issues and Approaches to Design and Analysis," *Journal of Teaching Physical Education*, 17: 270-284.

Silverman, S. (2004), "Analyzing Data From Field Research: The Unit of Analysis Issue," *Research Quarterly for Exercise and Sport*, (75) 2: iii-iv.

Smolkowski, K. (2012) "Experimental Unit," *Oregon Research Institute website*, http://www.ori.org/~keiths/Files/Tips/Stats_Unit.html. [Accessed 6/6/2012].

Starnes, D. S., Yates, D. S., and Moore, D. S. (2012), *The Practice of Statistics* (4th ed.), New York: W. H. Freeman and Company.

Watkins, A. E., Scheaffer, R. L., and Cobb, G. W. (2004), *Statistics in Action: Understanding a World of Data*, Emeryville, CA: Key Curriculum Press.

Watkins, A. E., Scheaffer, R. L., and Cobb, G. W. (2008), *Statistics in Action: Understanding a World of Data* (2nd ed.), Emeryville, CA: Key Curriculum Press.

Yates, D. S., and Moore, D. S., and McCabe, G. P. (1999), *The Practice of Statistics*, New York: W. H. Freeman and Company.

Yates, D. S., Moore, D. S., and Starnes, D. S. (2003), *The Practice of Statistics* (2nd ed.), New York: W. H. Freeman and Company.

Yates, D. S., Moore, D. S., and Starnes, D. S. (2008), *The Practice of Statistics* (3rd ed.), New York: W. H. Freeman and Company.

Jamis J. Perrett
Department of Statistics
Texas A&M University
3143 TAMU
College Station, TX 77843-3143
Email: jamis@stat.tamu.edu