A New Approach to Learning Probability in the First Statistics Course

Carolyn Keeler and Kirk Steinhorst
University of Idaho

Journal of Statistics Education Volume 9, Number 3 (2001)

Copyright © 2001 by Carolyn Keeler and Kirk Steinhorst, all rights reserved. This text may be freely shared among individuals, but it may not be republished in any medium without express written consent from the authors and advance notification of the editor.

Key Words: Cognition; Inquiry-based; Reform.

Abstract

The probability unit in a first statistics course is difficult to teach because there is not much time, the concepts and mechanics are difficult, and the students do not see the relevance of learning it. Research by Cosmides and Tooby (1996) supports our findings that instructors should avoid fractions and decimals and capitalize on students' affinity for counting things. In addition, we avoid the use of normal tables at the beginning of our discussion of continuous random variables by using uniform and triangular distributions. These ideas may be used in traditionally structured classes or in group-based and activity-based classes.

1. Introduction

The charge to statistics and mathematics educators to drastically change the way statistics courses are taught has been given by many well-known researchers in the field (see, for example, Hogg 1991; Watts 1991; Snell 1992; Cobb 1993; Garfield 1993, 1994; Snee 1993; Moore, Cobb, Garfield, and Meeker 1995; Chance 1997). One area that has received less focus in this literature is the teaching of probability. Over the last three decades much has also been written about people's inability to learn the concepts surrounding judgment under uncertainty but the research on solutions has been scarce. In a review of approaches to teaching probability and statistics, Garfield and Ahlgren (1988) made the recommendation that research into how to teach valid conceptions of probability continue. New findings by Cosmides and Tooby (1996) indicate that people do possess intuition about probability but that instruction needs to be changed to take advantage of this aptitude and to facilitate learning. Improving instruction in this area in an introductory statistics course is the focus of this paper.

In reaction to this new information about the way people learn to make judgments under uncertainty and as part of the reform efforts in statistics education, we propose changes to the way probability is taught. We will review the research by psychologists and statistics educators, summarize the extensive findings
of Cosmides and Tooby (1996), and provide exemplary material for implementing a new approach to teaching probability through learning exercises that apply Cosmides and Tooby's results and some other refinements we have developed in our classes. We then briefly review the research on authentic assessment and present assessment tasks that are consistent in content and context with the suggested learning exercises. The results of using these items in the Spring 2001 semester are then given.

2. Two Perspectives in the Literature

Psychologists doing research on how people learn probability in the 1970's and 1980's generally agreed that students' misconceptions cannot be easily overcome. In contrast, statistics and mathematics educators seem to hold out hope that students' intuitive beliefs can be used productively by teaching the appropriate links between primary intuitions and the mathematical model. In support of this second perspective, a more recent publication by Cosmides and Tooby (1996) presents empirical evidence that students do have an intuitive basis for correct probabilistic thinking when teaching approaches build on students' experience.

Statisticians, mathematicians and psychologists have been intensely researching the teaching and learning and resulting knowledge of probability for at least three decades. Throughout the 1970's and 1980's, Kahneman and Tversky's position (Kahneman and Tversky 1972, 1973, 1982; Tversky and Kahneman 1983) that people do not possess intuitions that follow a calculus of probability was widely accepted and provided the theoretical framework for the research on reasoning under uncertainty. They held that people who are statistically naive make estimates for the likelihood of events by using judgmental heuristics:

"In making predictions and judgments under uncertainty, people do not appear to follow the calculus of chance or the statistical theory of prediction. Instead, they rely on a limited number of heuristics that sometimes yield reasonable judgments and sometimes lead to severe and systematic errors." (Kahneman and Tversky 1973, p. 237)

The psychological research that resulted is replete with examples of these faulty "heuristics" which include: representativeness, the gambler's fallacy, the base-rate fallacy, availability, and the conjunction fallacy (Tversky and Kahneman 1974; Kahneman and Tversky 1982; Falk and Konold 1992; Garfield 1994). Kapadia and Borovcnik (1991) discuss the ramifications of these in their edited book on probability in education. The research presented in a book edited by Kahneman, Slovic, and Tversky (1982) is offered as evidence that inappropriate reasoning is widespread and persistent, similar at all age levels, found even among experienced researchers, and quite difficult to change. Research has identified additional misconceptions regarding conditional probability (see, for example, Pollatsek, Well, Konold, and Hardiman 1987; Falk 1988; Kelly and Zwiers 1988).

The psychology "camp" of research on stochastics teaching and learning explores how people think. A second body of research carried out primarily by mathematics and statistics educators is focused less on patterns of thinking and instead attempts to discover how statistics is learned. This research begins with different probability theories and the underlying mathematical concepts and focuses on questions of intuitive acceptance of theoretical concepts and related ideas (Borovcnik and Bentz 1991). Some of these studies provide a contrast to the psychological research and its implications (Garfield and del Mas 1989; Borovcnik 1991; Konold 1991; Garfield 1994). These authors take issue with the conclusions Kahneman and Tversky draw for reasoning under uncertainty as their results indicate that students' use of heuristics seems to vary with problem context (Garfield 1994). Other authors recommend building on students' sound intuitions (Borovcnik and Bentz 1991; Falk and Konold 1992) as "not all preconceptions are misconceptions" (Clement, Brown, and Zeitsman 1989).
Piaget's (1970) idea that intuitive notions grow with age has been the major focus of Fischbein and others' research on children's intuition and the influence of instruction (Fischbein and Gazit 1984; Fischbein, Pampu and Manzat 1970a, 1970b). In two studies with samples of 3000 and 1600 students aged seven to sixteen, Green (1982, 1988) found that facility with simple items rose with both age and ability but that with the more demanding items there was low performance that did not change with age or reasoning ability. Following this research, Green (1987, p. 8) offered "practical class based" activities to improve students' understanding. Fischbein also investigated intuition as beliefs that are adaptable and can be influenced by instruction. Fischbein et al. (1970a) were among the first to report changes in intuitions and conceptions of probability over the course of instruction. Children who received instruction on Piagetian probability ratio tasks improved in their predictions of outcomes. Fischbein (1987) believes that instruction can improve students' intuitive ideas of probability and makes a distinction between primary and secondary intuitions. Primary are the ideas and beliefs that we have before instructional intervention and secondary are restructured cognitive beliefs that we accept and use as a result of instruction or experience. He makes the recommendation that didactical situations be created that can help the student become aware of conflicts between primary and secondary intuitions. He goes on to state that this procedure has to be associated with the activity of analyzing explicitly the properties of the mathematical entities being considered. Borovcnik and Bentz (1991) also contend that conventional teaching establishes too few connections for students between primary intuitions and the mathematical model. Fischbein (1987) believes the elimination of intuitions should not be the goal, but the development of new and adequate intuitive interpretations together with the formal structures of logical reasoning. This can be done especially through practical activities and not through mere verbal explanations.

While intuitions are adaptable they are also subject to overextension. Cobb (1989), reviewing Fischbein (1987) calls this the "double edged sword" of intuitions. Intuitions can mislead and promote misconceptions of scientific reality, as well as provide help in thinking about that reality. Therefore, students have to learn that in science and mathematics not everything is intuitively understandable and that it is intuitively reasonable to suspend their primary intuitions in certain settings. McKinley (1960), Shulte (1968), and White (1974) all found that students' achievement in probability concepts increased significantly after instruction on probability. Beyth-Marom and Dekel (1983) developed a curriculum to improve probabilistic thinking, which they used with junior high students. They found that teaching students to examine their thought processes and common modes of thinking that may cause fallacies resulted in better performance on the questionnaire used for evaluation. Pfannkuch and Brown (1996) used activities to challenge students' intuitions and attempt to increase understanding of variation and probability with some success.

Barz (1970) and Shaughnessy (1977) found evidence that a practical involvement approach to students learning probability tended to result in higher student achievement than a traditional set-theoretic approach to probability. In Shaughnessy's study of how to present probability in a way that would increase student learning, he showed that students' misconceptions could be addressed through instruction. This was accomplished through hands-on experiments and activities in which students discovered counting principles and other concepts for themselves. Shaughnessy (1976, 1977) looked at the effects of instruction on stochastically naive college students' use of judgmental heuristics. In a twelve-week intensive teaching experiment, students' knowledge of probability and use of heuristics were assessed both by written responses and by taped interviews before and after instruction. Significant differences between two control classes and two experimental classes were found at the end of the instructional intervention, with greatly reduced reliance upon the heuristics in the experimental groups. Students kept a daily journal on experiments and homework problems concerning how they felt and what they learned. However, some students still did not change their responses or beliefs on those judgment tasks even after intensive instruction. "It is very difficult to replace a misconception with a normative conception, a primary intuition with a secondary intuition, or a judgmental heuristic with a
mathematical model. Beliefs and conceptions are slow to change" (Shaughnessy 1992, p. 481). Shaughnessy (1983a, 1983b) makes several suggestions for improved student performance including conducting interviews in order to explore students' thinking processes and misconceptions.

A number of the math and statistics based studies using computer modeling also have mixed results in terms of changing all students' misconceptions. Konold (1989b) used a computer modeling intervention in an attempt to influence students' misconceptions with mixed results. Some students changed their interpretation while others persisted in their erroneous positions. In their experiment with a computer program called "Coin Toss," Garfield and delMas (1989) also had mixed results. Again, some students changed their ideas about variability after using the tutorial but others persisted in their misconceptions about sample size and variation. Well, Pollatsek, and Boyce (1990) also found students had difficulty learning the effect of sample size on variability after computer aided instruction modeling sampling distributions. The evidence is that misconceptions about probability and statistics are difficult to change in some students and that computer simulations may not help (Snee 1993).

The different conclusions from the research representative of the psychology "camp" versus the statistics "camp" may represent a difference in research design. The methods of a study are a reflection of the purpose of the research and affect the way questions are posed; the influence of question format on the study's results needs to be considered as well as the findings themselves. Shaughnessy (1977) used journal writing and discussions with students to discover their thinking about probability issues. Shaughnessy (1977) and Konold (1989a) are among the few in the area of misconceptions of stochastics who have used a clinical methodology. In contrast, many of the tasks that have been administered by psychologists, such as Kahneman and Tversky, have involved forced-choice responses to particular item stems. A forced-choice method may be self-fulfilling in that alternative responses to such items do not get a chance to surface and so remain unknown. Shaughnessy's methods on the other hand provide an opportunity to explore the full range of student responses.

One task involving base rates, which has been investigated extensively in the research on judgment under uncertainty, will serve to illustrate how the researchers approach to the questions asked and the methods of a study can result in different responses and interpretations of those responses. This task has come to be known as the "Taxi Problem." It was first reported by Tversky and Kahneman (1980) and by Bar-Hillel (1980) and is stated in Shaughnessy (1992, p. 471) as follows:

"A cab was involved in a hit and run accident at night. There are two cab companies that operate in the city, a Blue Cab company, and a Green Cab company. It is known that 85% of the cabs in the city are Green and 15% are Blue. A witness at the scene identified the cab involved in the accident as a Blue Cab. This witness was tested under similar visibility conditions, and made correct color identifications in 80% of the trial instances. What is the probability that the cab involved in the accident was a Blue Cab rather than a Green one?"

The cognitive psychologists report that people ignore the base rate information (15% Blue Cabs) and place their faith in the reliability of the witness when they conclude that it was a Blue Cab. Tversky and Kahneman's earlier explanation for this is the representativeness heuristic, that people feel the single accident should be representative of the witness' 80% reliability data. An alternative explanation is causality as discussed by Ajzen (1977) and suggested by Well, Pollatsek, and Konold (1983) as the "outcome approach." Scholz and Bentrup (1984) investigate a great number of variants of the original "Taxi Problem" and conclude that responses to the problem are much more complicated than could be explained solely by the use of representativeness. Konold (1983, 1989a) asked students for their reasoning and his analysis of the reasons students gave for their choices suggested that students were not basing their responses on either the gambler's fallacy or the representativeness heuristic as claimed by Kahneman and Tversky. Again, where methods allow for a more open-ended exploration of students'
thinking, study results conflict with Kahneman and Tversky's conclusions.

Cosmides and Tooby (1996) suggest the possibility that human minds are capable of performing statistical calculations. They build on the ideas of Staddon (1988) that living organisms can be formally described as Bayesian inference machines. They also examine Gigerenzer's (1991) hypotheses that the human mind represents probabilistic information as frequencies and that some of our inductive reasoning mechanisms follow the logic of the frequentist probability school. Therefore, the idea that human brain functions include a calculus of probability along with other inductive reasoning capabilities is not so improbable as Tversky and Kahneman contend.

Cosmides and Tooby make the argument that in contrast to the standard view that human cognition does not include statistical or probabilistic thinking; that under the right conditions humans use counting and relative frequencies often and with some degree of competence. In fact, it may be the way questions are posed to students and other design issues that are accountable for demonstrated performance. When people are given frequencies as input and asked for frequencies as output there is some evidence that they are good "intuitive statisticians" (Cosmides and Tooby 1996, p. 21).

"The assumption of severe processing limitations has forestalled many researchers from seriously considering or investigating a contrasting possibility: that our minds come equipped with very sophisticated intuitive statistical competences that are well engineered solutions to the problems humans normally encountered in natural environments" (Tooby and Cosmides 1992, p. 9),

and ecologically valid input such as frequency formats may be necessary to activate these competencies. To test this hypothesis, Cosmides and Tooby conducted a series of experiments on single-event probability problems using medical students at Stanford.

They chose a problem famous in the "heuristics and biases" literature for eliciting base rate neglect; the study by Casscells, Schoenberger, and Graboys (quoted in Cosmides and Tooby 1996) on physicians and fourth-year medical students at Harvard Medical School on interpretation of clinical laboratory results. In Part I (Experiments 1-4), their purposes were to see if they could replicate Casscells et al.'s results on the original version of the medical diagnosis problem and to see whether a version of the problem using frequencies could be created that would elicit a higher percentage of correct responses by application of Bayes'rule than the original version. They presented three conditions:

1. the exact replication of the problem that Casscells et al. administered,
2. additional information to the original as well as input as frequencies and output requested as a frequency, and
3. prompts in the form of questions.

The original problem as stated by Casscells et al. follows.

" If a test to detect a disease whose prevalence is 1/1000 has a false positive rate of 5%, what is the chance that a person found to have a positive result actually has the disease, assuming that you know nothing about the person's symptoms or signs?_________%"

Cosmides and Tooby (1996) succeeded in both replicating the original research and in improving results by changing the form of the information provided. Condition 1 replicated the results of Casscells et al.:
12% of Stanford students as compared to 18% of Casscells subjects gave the correct response of "2%" and 56% of Stanford students as compared to 45% of Casscells subjects gave the incorrect response reflecting base rate neglect of "95%"; the Stanford students' performance was slightly worse on the original question. Given the information in frequencies (Condition 2) and given additional questions that prompted subjects to formulate additional frequencies (Condition 3) from the scenario, there was a preponderance of correct answers, 56% and 76%, respectively, and "base rate neglect" vanished.

In order to determine how probabilistic information is represented and processed by the human mind, Cosmides and Tooby continued their research with two more series of experiments. Their reasoning was that if non-frequency presentations of the problem could be created that would elicit high levels of performance, then they would not be justified in concluding from the previous experiments that frequency representations were responsible for correct reasoning. In Part II (Experiments 5-6), they asked:

1. Can simply clarifying the original problem elicit high levels of the correct application of Bayes' rule?
2. Are subjects who were given the original, non-frequency presentation of the problem actually good users of Bayes' rule who simply believed that the random sampling assumption had been violated?
3. If a violation of random sampling does not account for their distribution of responses, what does?

The results of Part II illustrated that neither stating the problem clearly, nor making the random sampling assumption explicit is sufficient to elicit high levels of correct responses from a problem that is not expressed in terms of frequencies.

In Part III (Experiments 7-8), they further tested the hypothesis that representations using frequencies result in the correct application of Bayes' rule by seeing whether systematically subtracting various frequency-based elements lowers performance. They did this by adding frequency representations to problems that lacked them and by subtracting frequency representations from problems that had them. In this way they addressed three questions:

1. Does asking for the answer as a frequency rather than as a single-event probability improve performance, all else being equal?
2. Does asking the subject to answer the problem with respect to an explicitly enumerated population improve performance?
3. Does presenting the problem information as frequencies, rather than as percents, improve performance?

The comparison of correct responses in these two additional experiments with those in Parts I and II shows that presenting the problem information as frequencies does in fact elicit higher levels of performance than presenting it as percents and that providing an explicitly enumerated population does not seem to make any difference. In addition, both presenting the information as frequencies as opposed to percents and asking for the answer as a frequency as opposed to a single-event probability contribute significantly to performance. In short (see the original 73 page report for more details), all predictions of the original hypothesis were confirmed (Cosmides and Tooby 1996):
1. Inductive reasoning performance differed depending on whether subjects were asked to judge a frequency or the probability of a single event.

2. Performance on frequency-based versions of problems was superior to non-frequency versions.

3. The more subjects could be mobilized to form a representation of the problem in frequencies, the better their performance.

4. Performance on frequency-based problems satisfied those constraints of a calculus of probability that we tested for (that is, Bayes' rule).

These results demonstrate the point that the way the questions are posed results in different levels of performance. More research that compares designs and methods used will add to the growing understanding of the effect of the researcher's approach on the results.

This new information and evidence of people's ability to think in probabilistic terms if the task information is represented as frequencies has led us to rethink our approach to teaching probability. The learning and assessment tasks presented below reflect this new approach. They also reflect the continuing belief that students learn better and retain more if they engage in learning activities that require them to think and process information rather than to passively listen to lectures.

3. Course Learning Materials

The introductory statistics course frequently includes a two to three week unit on the probability topics a student must understand for statistical inference. Approaches that imbed learning in activity and make deliberate use of the social and physical context are likely to produce more meaningful learning. We suggest using an inquiry-based approach that includes the strategy of having students reconcile their initial "intuitive" response with their formal working of the problem in order to confront misconceptions (delMas and Bart 1987; Garfield and delMas 1989; Konold 1991; Shaughnessy 1992). Research in support of adding small group work to this approach to teaching statistics is available elsewhere (Garfield 1992, 1993; Cobb 1993; Dietz 1993; Bilotti-Aliaga 1994; Keeler and Steinhorst 1995; Steinhorst and Keeler 1995; Zetterqvist 1997). Many researchers agree that students must be active rather than passive learners and should work on meaningful projects to answer questions of interest to them (Cobb 1991; 1993; Hogg 1992; Roberts 1992; Scheaffer 1992; Snell 1992; 1994; Snell and Finn 1992; Rossman and Short 1995; Smith 1998; West and Ogden 1998).

Whether or not you use an inquiry-based approach to teaching and learning, Cosmides and Tooby's research calls for a change in the content and format of learning exercises. Given their evidence that people can correctly figure probabilities when given problem information as frequencies, we expanded our examples that used frequencies and culled those examples set up using fractions or decimals. In addition, we tie the concept of probability to taking random samples from populations and introduce the related ideas of probability mass functions and density functions right from the start. To avoid dependence on calculus, when we introduce density functions we use rectangular and triangular densities instead of normal densities. This allows the student to concentrate on the idea of a density function rather than on the intricacies of manipulating the normal table.

Our presentation begins with identification of specific learning goals for the probability unit followed by exemplary learning exercises.

3.1. Learning Goals
This probability unit is part of an introductory course in statistics. While a semester long course may deal with a variety of other probability subjects, the topics covered in this unit are limited to those one must understand for statistical inference. In this setting, one has a relatively short period in a semester long course to cover the following topics:

**The axioms of probability:**

1. Probabilities are between zero and one
2. The probability of "all" is one
3. The addition rule for disjoint events

**Other basic rules of probability:**

4. Complement
5. General addition rule
6. Conditional probability

**Independence as a special case of 6:**

7. Independence

In a statistics course the students need to understand the connection between probability and populations. That is, they need to understand probability mass functions (pmfs) and probability density functions (pdfs).

In this unit, the overall learning goal is understanding demonstrated by correct application of the knowledge represented above (Wiggins and McTighe 1998).

### 3.2. Learning Exercises

The following exercises are based on the use of frequency representations as researched by Cosmides and Tooby (1996) and an active, inquiry-based approach to learning. It is useful if each student writes down an "intuitive" answer to start. If small groups are used, another step can involve the group developing a consensus intuitive answer from these. Finally, students work through to a formal answer alone or in the group. In this way, they must confront their original misconceptions. We have used these and similar problems as both class examples and homework.

**Learning exercise 1: University Demographics**

Current enrollments by college and by sex appear in the following table.

| College: | Ag-For | Arts-Sci | Bus-Econ | Educ | Engr | Law | Undecl | Totals |
|----------|--------|----------|----------|------|------|-----|--------|--------|
| Female   | 500    | 1500     | 400      | 1000 | 200  | 100 | 800    | 4500   |
If I select a student at random, answer the following:

a. Find $P(\text{Female or Male})$.

b. Find $P(\text{not Ag-For})$.

c. Find $P(\text{Female | BusEcon})$.

d. Find $P(\text{Male and Arts-Sci})$.

e. Are "Female" and "Educ" statistically independent? Why or why not?

This problem relies on the Cosmides and Tooby premise that humans can think probabilistically if they can determine the frequency of an event. This exercise develops the students' formal understanding of the basic rules in a context in which most students succeed.

When we introduce this example, we emphasize that probability relates to the random selection of a student. While it is true that the answers depend on calculating various proportions, probabilities are not merely proportions.

In this example, the frequencies presented are roughly those in the population. While the students know that we do not have exactly 10,000 students, they know we have 10,000 "in round numbers." This is a good example with which to start because they see the direct relationship between the frequencies and the eventual probabilities of random selection. In later examples, the frequencies reported do not necessarily relate to the numbers in a finite population. In fact the population may not be finite in size (see the following exercise).

**Learning exercise 2: Predicting Sex of Babies**

Many couples take advantage of ultrasound exams to determine the sex of their baby before it is born. Some couples prefer not to know beforehand. In any case, ultrasound examination is not always accurate. About 1 in 5 predictions are wrong. In one medical group, the proportion of girls correctly identified is 9 out of 10 and the number of boys correctly identified is 3 out of 4. The proportion of girls born is 48 out of 100.

What is the probability that a baby predicted to be a girl actually turns out to be a girl?

Think about the next 1000 births handled by this medical group. How many should be girls? How many should be boys? Of the girls, how many will the test indicate are girls? Of the boys, how many will the test indicate are girls? From these numbers you can calculate the $P(\text{girl | test says girl})$.

Kahneman and Tversky (1972, 1973, 1982) showed that students generally cannot work this problem by applying Bayes' rule to the fractions: $P(\text{test says girl | girl}) = 0.9$, $P(\text{test says girl | boy}) = 0.25$, and $P$
( girl ) = 0.48. The counting approach tied together with probing questions enables most students to work this problem.

**Learning exercise 3: Putting in Extra Trunk Lines Between [insert local names for Town A and Town B]**

Given recent flooding (or other condition more appropriate to your area) between Town A and Town B, the local telephone company is assessing the value of adding an independent trunk line between the two towns. The second line will fail independently of the first because it will depend on different equipment and routing (we assume a regional disaster is highly unlikely).

Under current conditions, the present line works 98 out of 100 times someone wishes to make a call. If the second line performs as well, what is the chance that a caller will be able to get through?

Formally,

\[
P( \text{Line 1 works}) = \frac{98}{100} \\
P( \text{Line 2 works}) = \frac{98}{100}
\]

find

\[
P( \text{Line 1 or Line 2 works}).
\]

**Learning exercise 4: Part-time Work on Campus**

A student has been offered part-time work in a laboratory. The professor says that the work will vary from week to week. The number of hours will be between 10 and 20 with a uniform probability density function, represented as follows:

![Uniform Distribution](image)

How tall is the rectangle?

What is the probability of getting less than 15 hours in a week?

Given that the student gets at least 15 hours in a week, what is the probability that more than 17.5 hours will be available?

Many beginning statistics students get lost when using the normal table to work probability problems relating to continuous random variables. If you stick to rectangles (uniform densities) and triangles (see
below), they do fine. Later, when you introduce normal densities they are confident that they know the basic principles.

**Learning exercise 5: Customer Complaints**

You are the manager of the complaint department for a large mail order company. Your data and experience indicate that the time it takes to handle a single call has the following pdf,

\[
\text{density} \\
\begin{array}{c}
\text{time} \\
0 \\
15 \\
\end{array}
\]

Show that the area under the triangle is 1.

Find the probability that a call will take longer than 10 minutes. That is, find \( P( \text{Time} > 10 ) \).

Given that the call takes at least 5 minutes, what is the probability that it will take longer than 10 minutes? That is, find \( P( \text{Time} > 10 | \text{Time} > 5 ) \).

Find \( P( \text{Time} < 10 ) \).

**Learning exercise 6: Clutch Sizes in Boreal Owl Nests**

The number of eggs in Boreal owl nests has a probability mass function with \( P(0) = 0.2 \), \( P(1) = 0.1 \), \( P(2) = 0.1 \), \( P(3) = 0.3 \), and \( P(4) = 0.3 \).

Draw the pmf correctly labeling the axes.

Show that the pmf you drew is a proper pmf.

What is the probability that a randomly chosen nest will have no eggs?

If you examine two nests and they are independent, what is the probability that neither nest will have eggs?

### 4. Assessment Tasks

The approach taken in the materials presented above for teaching a probability unit in an introductory statistics course must now be applied to the assessment of learning. New instructional techniques require new forms of assessment (Garfield 1993). We know that one result of the practice of giving students a final evaluation, usually in the form of grades, is that assessment drives learning. As psychologist Lauren Resnick said, "You get what you assess. You do not get what you do not assess" (Resnick and Resnick 1990, p. 66). What we assess therefore must be of importance and value and represent the heart of the subject or learning goals established for the course (Hogg 1992; Wiggins 1992). What do we want...
students to know and to be able to do at the end of the probability unit?

Clearly defined goals and formative assessments that measure these allow us to gather and evaluate information in order to improve teaching and learning (Garfield 1993; Keeler 1997). To do this, we must know where students are confused or unable to make the appropriate application of learning in time to reteach concepts (Keeler and Steinhorst 1995; Smith 1998).

The assessment tasks students perform must be "real world" problems, involve minimal computation and demonstrate the student's conceptual understanding. Whether used formatively or summatively, these authentic assessment tasks can be thought of as real performances or the application of knowledge in real life situations (Archbald and Newman 1988; Wiggins 1993). We strive to have tasks include an element of ambiguity to be overcome, which requires the student to follow a logical thought process. Both the content and context of assessment must match the learning exercises (Wiggins 1989; Darling-Hammond 1994). Cobb (1993), Chance (1997), and Smith (1998) have written in detail about their experiences with authentic assessment. Chance identified the following goals for using alternative assessment methods to:

a. better gauge the students' understanding of statistical ways of thinking,

b. evaluate their communication and collaboration skills,

c. measure their statistical and computer literacy, and

d. monitor their interest in statistics.

Smith's project-based assessment reflects both authentic, real-world applications and his position that "if we want students to understand and communicate statistical results, then their course grade should depend substantially on how well they do so" (Smith 1998, p. 6). Cobb (1993) in his article tying TQM to authentic assessment states that "... the most radical implication of TQM is that the entire course should be built of assessment tasks."

The following are examples of assessment tasks that match the new approach to teaching probability. In preparation for an exam, we usually give take-home questions that the students work and bring to class. The in-class portion of the test then asks questions related to or based on the take-home questions. The students are encouraged to work the take-home in groups of peers, but they may elect to do it alone. They are individually responsible for their in-class answers. If the instructor wishes to have a more traditional in-class exam, these take-home/in-class questions would be shortened and blended together.

These are the actual questions used in an introductory class of 84 students during the Spring 2001 semester. Each in-class question was worth 14 points. The points were distributed evenly for multipart questions. For example, in problem 1, parts d, e, and f are worth 5, 5, and 4 points, respectively. In problem 2 each of the two parts is worth seven points. We tallied the number of points taken off for each part of each question. These results are presented below. The first line gives the number of points missed on that question and the second line contains the frequencies of students missing that many points. The total of the frequencies in the second line is 84 in each case. The exam key contains marginal notes denoting the basis for partial points. For example in question 1 part d if they correctly worked the problem assuming independence, then only two of the five points were deducted. In problem 1 part e if they argued the case for or against mutually exclusive instead of independent and did it correctly, then they were given two of the five points. In question problem 4 part e if they correctly answered "less than 2" but based their argument on the discrete definition of the mean, then they lost two of the four points. The tests were graded by going through the entire 84 tests a problem or two at a time so that partial
credit could be fairly assigned and unbiased assessment of answers is more likely. If one grades a single test all the way through there might be a tendency to give a student who is doing well overall the benefit of the doubt on a particular question. Likewise, if a student is doing poorly and there is some confusion in his or her answer, then one might grade more severely.

**Take-home assessment question 1:**

Suppose that for every 1000 voters 400 live in urban areas, 400 live in suburban areas and 200 live in rural areas. Among those voters some favor the state spending more money to develop tourism and some do not. These numbers are captured in the table below:

| Opinion | Favor | Oppose | Totals |
|---------|-------|--------|--------|
| Urban   | 300   | 100    | 400    |
| Suburb  | 200   | 200    | 400    |
| Rural   | 50    | 150    | 200    |
| Totals  | 550   | 450    | 1000   |

If I select a voter at random, answer the following:

a. Find \( P( \text{an Urban or Suburban voter who Favors} ) \).

b. Find \( P( \text{Rural or Favor} ) \).

c. \( P( \text{Favor | Rural} ) \).

The in-class questions are:

d. Find \( P( \text{Favor and Rural}) \).

**Frequency distribution of points missed:**

| Points missed | 0  | -1 | -2 | -3 | -4 | -5 |
|---------------|----|----|----|----|----|----|
| Frequency     | 56 | 0  | 18 | 7  | 2  | 1  |

e. Are favor and rural independent? Explain your reasoning.
f. On the take-home you found $P(\text{ Favor } | \text{ Rural })$. Find $P(\text{ Favor })$. Using these results, explain why the two probabilities are different.

Frequency distribution of points missed:

| Points missed | 0 | -1 | -2 | -3 | -4 | -5 |
|---------------|---|----|----|----|----|----|
| Frequency     | 40| 3  | 4  | 13 | 2  | 22 |

This is essentially a variation of Learning Exercise 1 above. We are assessing whether or not they have mastered the basics of calculating probabilities by counting. In part e the most common wrong explanation was to confuse independence and mutually exclusive, "No, they are not because there are rural residents that do favor." In part f we are asking them to demonstrate understanding, rather than mechanics. A typical correct answer reads, "The main reason that the two probabilities are different is, when looking at $P(\text{ Favor } | \text{ Rural })$, you narrow your sample set."

Take-home assessment question 2:

One out of every 1000 Americans has disease X, that is $P(\text{ Disease X }) = 1/1000$. A test has been developed to detect when a person has disease X. Every time the test is given to a person who has the disease, the test comes out positive, that is, $P( + | \text{ Disease X } ) = 1$. But sometimes the test also comes out positive when it is given to a person who is completely healthy. Specifically, out of every 1000 people who are perfectly healthy, 50 of them test positive for the disease, that is, $P( + | \text{ No disease X } ) = 5/100$ (this is known in the medical community as a "false positive"). We want to know the probability of a patient having disease X when we get a positive test result. That is, find $P(\text{ Disease X } | + )$.

Hint: Think of 100,000 people. How many should have the disease? How many should not have the disease? Continue figuring these idealized frequencies until you get the information you need to solve the problem.

The in-class questions read:

a. Give your answer from the take-home for $P(\text{ Disease X } | + )$. 
Frequency distribution of points missed:

| Points missed | 0 | -1 | -2 | -3 | -4 | -5 | -6 | -7 |
|---------------|---|----|----|----|----|----|----|----|
| Frequency     | 61| 4  | 2  | 2  | 1  | 5  | 9  |    |

b. Using your figures from your tree diagram find \( P(\text{No disease} | -) \).

Frequency distribution of points missed:

| Points missed | 0 | -1 | -2 | -3 | -4 | -5 | -6 | -7 |
|---------------|---|----|----|----|----|----|----|----|
| Frequency     | 27| 6  | 4  | 2  | 24 | 6  | 5  | 10 |

This question illustrates several points.

1. Sometimes we ask for the answer to the take-home question to establish a baseline answer for the next part.

2. The follow-up question is particularly difficult because the correct answer is 1. Students must not only understand what they did on the take-home, but must have the courage to put down an unusual answer.

Assessment question 2 is the same question that Cosmides and Tooby used in their experiments. This version is slightly harder because these students had to work out the problem without benefit of seeing multiple-choice answers. A total of 65 out of 84 missed 0, 1, or 2 points when the question was posed as a take-home question. This is roughly the same percentage of correct answers found by Cosmides and Tooby. On the harder in-class question, 37 out of 84 did really well (missed 0, 1, or 2) and 32 out of 84 missed 3, 4, or 5. The remaining 15 missed 6 or 7 points indicating they really didn't master the logic of the question.

One student correctly worked the problem using the usual probability rules and then noted, "you could also figure this out knowing that in any negative test, the person NEVER has the disease, so if you get a negative test, 100% of the time the person will not have X."

**Take-home assessment question 3:**

The example in B6 Basics page 2 (of CyberStats 1.2, the course text) gives a probability mass function for the number of credit cards owned by students. In round numbers the pmf appears as:
If I select a student at random,

a. Find $P( X > 2 )$.

b. Define the complement of the event in part a above.

c. Find the population mean, $\mu$, of $X$.

d. Give an example of two mutually exclusive events defined using $X$ and the pmf above.

e. Are the events "$X > 2$" and "$X > 3$" independent? Show why or why not.

The in-class questions read:

f. Find $P( X \leq 4 )$, that is, find the probability that a randomly chosen student has 4 or less credit cards. Show work.

Frequency distribution of points missed:

| Points missed | 0  | -1 | -2 | -3 | -4 | -5 | -6 | -7 |
|---------------|----|----|----|----|----|----|----|----|
| Frequency     | 76 | 3  | 4  | 0  | 0  | 0  | 0  | 1  |

g. Define mutually exclusive. Are the events, "student has 0 credit cards" and "student has 3 credit cards" mutually exclusive?

Frequency distribution of points missed:
If their definition mixed the ideas of independent and mutually exclusive, they lost three of the seven points. For example, "When two things are mutually exclusive, they cannot happen simultaneously and have no effect on each other." The last phrase is more appropriate when talking about independence. A better definition is, "Mutually exclusive means two things cannot occur at once. If one occurs the other cannot."

Take-home assessment question 4:

An appliance repair firm, has recorded the time to complete service calls during the past year. They found that time varies from a low of 0 hours (no repair required) to a high of 4 hours. The pdf is given as follows:

![Graph showing the probability density function for the time to complete service calls.](image)

a. Find the height of the triangle, that is, find the height where the triangle crosses the y-axis.

If I consider a randomly selected call,

b. Find the probability that a service call will take at most an hour.

c. \( P( X > 2 \text{ hours} ) \).

d. Is the population mean, \( \mu \), less than 2 or greater than 2? Give your reasoning.

The in-class questions read,

e. Is the population mean, \( \mu \), less than 2 or greater than 2? Give your reasoning.

Frequency distribution of points missed:

| Points missed | 0  | -1 | -2 | -3 | -4 |
|---------------|----|----|----|----|----|
| Frequency     | 60 | 4  | 14 | 4  | 2  |
f. \( P( X > 3 \mid X > 2 ) \).

Frequency distribution of points missed:

| Points missed | 0  | -1 | -2 | -3 | -4 | -5 |
|---------------|----|----|----|----|----|----|
| Frequency     | 41 | 1  | 7  | 16 | 2  | 17 |




g. If the pdf is uniform (flat between 0 and 4), it will be a rectangle 0.25 high. Is \( P( X > 2 ) \) larger or smaller than the answer you found on the take-home? Draw the uniform pdf on the drawing above and shade in the area above 2 showing why your answer is reasonable.

Frequency distribution of points missed:

| Points missed | 0  | -1 | -2 | -3 | -4 | -5 |
|---------------|----|----|----|----|----|----|
| Frequency     | 53 | 5  | 3  | 3  | 5  | 14 |

In this problem, the students are working with probability density functions in simple ways to illustrate their understanding of the basic definitions, the rules of probability, and probability density functions. The first in-class question is, of course, the same as the last take-home question. Item g asks them to think about how changing shapes changes the corresponding probability calculation.

Considered as a whole, most students did well. They demonstrated mastery of basic probability concepts and calculations. Most students could do the Bayes' rule problem when they were not under the pressure of a fifty-minute exam and could talk with peers. With our method of testing, a student may be lulled into thinking he or she understands the take-home problem, but when asked to do a slightly more difficult problem in class, they cannot do as well. This is not a problem for the bulk of the students who understand the importance of using the take-home as preparation for the in-class exam. Students were able to connect probability with probability mass and density functions. Working with uniform and triangular densities gives them the confidence they need when they move on to more difficult densities like the normal density.

5. Conclusion

Historically we have used material from various texts that place too much emphasis on permutations, combinations, and the use of statistical tables. The result was students' failure to learn probabilistic thinking. As a consequence, we moved away from teaching probability in the first statistics course.
Neither situation was satisfactory. Recently, we have used the approach demonstrated above, incorporating problem information provided in tables of counts and probability given and asked for in relative frequencies, and providing probability density functions based on simple rectangles and triangles and probabilities represented as simple areas. Students are successful using these materials. They actually develop some facility with this approach and hopefully can apply their learned skills to life situations involving uncertainty after the class ends.

Some of this success may be due to the activity-based, group learning approach we use in the classroom. However, we have been using activity-based, group learning since 1990 (see Keeler and Steinhorst 1995 and Steinhorst and Keeler 1995). We first discovered and started using the Cosmides and Tooby (1996) results after a colleague who teaches our business statistics course, Ray Dacey, pointed it out to us. Spring 2001 is the third semester we have used counts to teach basic probability skills. We see no reason why instructors using traditional lecture cannot use this approach successfully.

It is important that instructors point out to students that probabilities are not just relative counts. Probability is not just the arithmetic of computing proportions. Students might become adept at computing various relative counts and still not understand the nature of randomness. This is one reason why we feel that introducing probability mass functions and probability density functions in the basic probability unit is so important. We can then talk about picking an element at random from a population that leads to a discussion of the probability of a given outcome.

The new information offered by Cosmides and Tooby (1996) about how people think about probability has been successfully applied in one introductory statistics course. The changes we suggest in the teaching of probability have direct implications for the development of curriculum and assessments. Our emphasis is on the use of frequencies within a constructivist framework to develop students' understanding of concepts such as independence and randomness. The use of simple probability density functions allows the student to focus on learning about the probabilities associated with continuous random variables without getting tied up in the mechanics of standard normal tables and $z$-scores. After students master these simple pdf's, one can introduce the normal distribution and its associated problems with calculation. Work needs to be done to investigate how students thinking about probability changes as a result of this approach.

References

Ajzen, I. (1977), "Intuitive Theories of Events and the Effects of Base-rate Information on Prediction," *Journal of Personality and Social Psychology*, 35, 303-314.

Archbald, D. A., and Newmann, F. M. (1988), *Beyond Standardized Testing: Assessing Authentic Academic Achievement in the Secondary School*, Reston, VA: National Association of Secondary School Principals.

Bar-Hillel, M. (1980), "The Base-Rate Fallacy in Probability Judgments," *Acta Psychologica*, 44, 211-233.

Barz, T. J. (1970), "A Study of Two Ways of Presenting Probability and Statistics at the College Level," unpublished dissertation, Columbia University.

Beyth-Marom, R., and Dekel, S. (1983), "A Curriculum to Improve Thinking Under Uncertainty,"
Bilotti-Aliaga, M. (1994), "Exploring Probability and Statistics Through Hands-on Experiments and Cooperative Learning," in Proceedings of the American Statistical Association Section on Statistical Education, Alexandria, VA: American Statistical Association, 150-155.

Borovcnik, M. (1991), "A Complementarity Between Intuitions and Mathematics." in Proceedings of the Third International Conference on Teaching Statistics, ed. D.Vere-Jones, Voorburg, The Netherlands: International Statistical Institute, 363-369.

Borovcnik, M., and Bentz, H. J. (1991), "Empirical Research in Understanding Probability, in Chance Encounters: Probability in Education, eds. R. Kapadia and M. Borovcnik, The Netherlands: Kluwer Academic Publishers, 73-105.

Chance, B. (1997), "Experiences with Authentic Assessment Techniques in an Introductory Statistics Course," Journal of Statistics Education [Online], 5(3). (ww2.amstat.org/publications/jse/v5n3/chance.html)

Clement, J., Brown, D. E., and Zeitsman, A. (1989), "Not All Preconceptions Are Misconceptions: Finding 'Anchoring Conceptions' for Grounding Instruction on Students' Intuitions," International Journal of Science Education, 11 (special issue), 554-565.

Cobb, G. (1991), "Teaching Statistics: More Data, Less Lecturing," Amstat News, 182(1), 4.

----- (1993), "Reconsidering Statistics Education: A National Science Foundation Conference," Journal of Statistics Education [Online], 1(1). (ww2.amstat.org/publications/jse/v1n1/cobb.html)

Cobb, P. (1989), "A Double-Edged Sword, Review of Intuition in Science and Mathematics by E. Fischbein" , Journal for Research in Mathematics Education, 20, 213-216.

Cosmides, L., and, Tooby, J. (1996), "Are Humans Good Intuitive Statisticians After All? Rethinking Some Conclusions from the Literature on Judgment Under Uncertainty," Cognition, 58, 1-73.

Darling-Hammond, L. (1994), "Performance-Based Assessment and Educational Equity," Harvard Educational Review, 64(1), 5-30.

delMas, R. C., and Bart, W. M. (1987), "The Role of an Evaluation Exercise in the Resolution of Misconception of Probability," paper presented at the Annual meeting of the American Educational Research Association, April.

Dietz, E. J. (1993), "A Cooperative Learning Activity on Methods of Selecting a Sample," The American Statistician, 47, 104-108.

Falk, R. (1988), "Conditional Probabilities: Insights and Difficulties," in Proceedings of the Second International Conference on Teaching Statistics, eds. R. Davidson and J. Swift, Victoria, BC: University of Victoria, 292-297.

Falk, R., and Konold, C. (1992), "The Psychology of Learning Probability," in Statistics for the Twenty-First Century, MAA Notes (Number 26), eds. F. Gordon and S. Gordon. Washington D.C.: The Mathematics Association of America, 151-164.
Fischbein, E. (1987), *Intuition in Science and Mathematics*, Dordrecht, Holland: D. Reidel.

Fischbein, E., and Gazit, A. (1984), "Does the Teaching of Probability Improve Probabilistic Intuitions?" *Educational Studies in Mathematics*, 15, 1–24.

Fischbein, E., Pampu, I., and Manzat, I. (1970a), "Comparison of Ratios and the Chance Concept in Children," *Child Development*, 41, 377-389.

----- (1970b), "Effects of Age and Instruction on Combinatory Ability in Children," *British Journal of Educational Psychology*, 40, 261-270.

Garfield, J. (1992), "Notes: Assessment and Teaching Statistics," *Journal of Statistics Education* [Online], 1(1).

(ww2.amstat.org/publications/jse/v1n1/cobb.supp.garfield1.html)

----- (1993). "Teaching Statistics Using Small-Group Cooperative Learning." *Journal of Statistics Education* [Online], 1(1).

(ww2.amstat.org/publications/jse/v1n1/garfield.html)

----- (1994). "How Students Learn Statistics." *Journal for Research in Mathematics Education* [Online] 19:44-63. (www.dartmouth.edu/Echance/teaching\_aids)

Garfield, J., and Ahlgren, A. (1988), "Difficulties in learning basic concepts in statistics: Implications for research," *Journal for Research in Mathematics Education*, 19, 44-63.

Garfield, J. B., and delMas, R. C. (1989), "Reasoning About Chance Events: Assessing and Changing Students' Conception of Probability," in *11th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (Vol. 2), New Brunswick: Rutgers, 189-195.

Gigerenzer, G. (1991), "How to Make Cognitive Illusions Disappear: Beyond Heuristics and Biases," *European Review of Social Psychology*, 2, 83-115.

Green, D. R. (1982), *Probability Concepts in 11-16 Year Old Pupils*, Loughborough University, UK: Centre for Advancement of Mathematical Education in Technology, 40 p.

----- (1987), "Probability Concepts: Putting Research into Practice." *Teaching Statistics*, 9(1), 8-14.

----- (1988), "Children's Understanding of Randomness: Report of a Survey of 1600 Children Aged 7-11 Years," in *The Proceedings of the Second International Conference on Teaching Statistics*, eds. R. Davidson and J. Swift. Victoria, BC: University of Victoria, 287-291.

Hogg, R. V. (1991), "Statistical Education: Improvements Are Badly Needed," *The American Statistician*, 45, 342-343.

----- (1992), "Towards Lean and Lively Courses in Statistics," in *Statistics for the Twenty-First Century, MAA Notes* Number 26, eds. F. Gordon and S. Gordon, Washington D.C.: The Mathematics Association of America, 3-13.

Kahneman, D., and Tversky, A. (1972), "Subjective Probability: A Judgment or Representativeness,"
Cognitive Psychology, 3, 430-454.

----- (1973), "On the Psychology of Prediction," Psychological Review, 80, 237-251.

----- (1982), "On the Study of Statistical Intuitions," Cognition, 11, 123-141.

Kahneman, D., Slovic, P., and Tversky, A. (1982), Judgment Under Uncertainty: Heuristics and Biases, Cambridge: Cambridge University Press.

Kapadia, R., and Borovcnik, M. (1991), "The Educational Perspective," in Chance Encounters: Probability in Education, eds. R. Kapadia and M. Borovcnik, The Netherlands: Kluwer Academic Publishers, 1-26.

Keeler, C. M., and Steinhorst, R. K. (1995), "Using Small Groups to Promote Active Learning in the Introductory Statistics Course: A Report from the Field," Journal of Statistics Education [Online], 3(2). (ww2.amstat.org/publications/jse/v3n2/Keeler.html)

Keeler, C. M. (1997), "Portfolio Assessment in Graduate Level Statistics Courses," in The Assessment Challenge in Statistics Education, eds. I. Gal and J.B. Garfield, Amsterdam: IOS Press, 165-178.

Kelly, I. W., and Zwiers, F. W. (1988), "Mutually Exclusive and Independence: Unravelling Basic Misconceptions in Probability Theory," in Proceedings of the Second International Conference on Teaching Statistics, eds. R. Davidson and J. Swift, Victoria, BC: University of Victoria, 1-26.

Konold, C. (1983), "Conceptions About Probability: Reality Between a Rock and a Hard Place," University of Massachusetts, Dissertation Abstracts International, 43, 4179b.

----- (1989a), "Informal Conceptions of Probability," Cognition and Instruction, 6, 59-98.

----- (1989b), "An Outbreak of Belief in Independence?" in Proceedings of the 11th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education, (Vol. 2), eds. G. Goldin, C. Maher, and B. Davis. Rutgers, NJ: Rutgers University Press, 203-209.

----- (1991), "Understanding Students' Beliefs About Probability," in Radical Constructivism in Mathematics Education, ed. E. von Glasersfeld, The Netherlands: Kluwer Academic Publishers, 139-156.

McKinley, J. E. (1960), "Relationship Between Selected Factors and Achievement in a Unit on Probability and Statistics for Twelfth Grade Students," Dissertation Abstracts, 21:561.

Moore, D. S., Cobb, G., Garfield, J., and Meeker, W. Q. (1995), "Statistics Education Fin de Siecle," The American Statistician, 49(3), 250-260.

Pfannkuch, M., and Brown, C. M. (1996), "Building on and Challenging Students' Intuitions About Probability: Can We Improve Undergraduate Learning?," Journal of Statistics Education [Online], 4(1). (ww2.amstat.org/publications/jse/v4n1/pfannkuch.html)

Piaget, J. (1970), Structuralism, New York: Basic Books.

Pollatsek, A., Well, A. D., Konold, C., and Hardiman, P. (1987), "Understanding Conditional
Probabilities," *Organizational Behavior and Human Decision Processes*, 40, 255-269.

Resnick, L. B., and Resnick, D. P. (1990), "Tests as Standards of Achievement in Schools," in *The Uses of Standardized Tests in American Education, Proceedings of the 1989 ETS Invitational Conference*, ed. J. Pfeiderer, Princeton: Educational Testing Service.

Roberts, H. V. (1992), "Student-Conducted Projects in Introductory Statistics Courses," in *Statistics for the Twenty-First Century*, MAA Notes, (Number 26), eds. F. Gordon and S. Gordon. Washington D.C.: The Mathematical Association of America, 109-121.

Rossman, A. J., and Short, T. H. (1995), "Conditional Probability and Education Reform: Are They Compatible?" *Journal of Statistics Education* [Online], 3(2). (ww2.amstat.org/publications/jse/v3n2/rossman.html)

Scheaffer, R. L. (1992), "Data, Discernment and Decisions: An Empirical Approach to Introductory Statistic," in *Statistics for the Twenty-First Century*, MAA Notes (Number 26), eds. F. Gordon and S. Gordon, Washington D.C.: The Mathematics Association of America, 69-82.

Scholz, R. W., and Bentrup, A. (1984), "Reconsidering the Base Rate Fallacy: New Data and Conceptual Issues, The Stochastic Thinking Project," Report No. 1, Bielefeld, FRG: University of Bielefeld.

Shaughnessy, J. M. (1976), "A Clinical Investigation of College Students' Reliance Upon the Heuristics of Availability and Representativeness in Estimated the Likelihood of Probabilistic Events," *Dissertation Abstracts International*, 37, 5662A.

----- (1977), "Misconceptions of Probability: An Experiment with a Small-Group, Activity-Based, Model-Building Approach to Introductory Probability at the College Level," *Educational Studies in Mathematics*, 8, 285-316.

----- (1983a), "Misconceptions of Probability, Systematic and Otherwise: Teaching Probability and Statistics so as to Overcome Some Misconceptions," in *Proceedings of the First International Conference on Teaching Statistics*, eds. D.R. Grey and P. Holmes, Sheffield, UK: Teaching Statistics Trust, 784-801.

----- (1983b), "The Psychology of Inference and the Teaching of Probability and Statistics: Two Sides of the Same Coin?" in *Decision Making Under Uncertainty*, ed. R. Scholz, Amsterdam: Elsevier Publishers, 325-350.

----- (1992), "Research in Probability and Statistics: Reflections and Directions," in *Handbook of Research on Mathematics Teaching and Learning: A Project of the National Council of Teachers of Mathematics*, ed. D.A. Grouws, New York: Macmillan Publishing, 465-494.

Shulte, A. P. (1968), "Effect of a Unit in Probability and Statistics on Students and Teacher of Ninth Grade General Mathematics Class," *Dissertation Abstracts*, 28, 4962A.

Smith, G. (1998), "Learning Statistics by Doing Statistics," *Journal of Statistics Education* [Online], 6 (3). (ww2.amstat.org/publications/jse/v6n3/smith.html)
Snee, R. D. (1993), "What's Missing in Statistical Education?" The American Statistician, 47, 149-154.

Snell, J. L. (1992), "CHANCE: Case Studies of Current Chance Issues," in Statistics for the Twenty-First Century, MAA Notes (Number 26), eds. F. Gordon and S. Gordon, Washington D.C.: The Mathematical Association of America, 281-297.

----- (1994), "Take a Chance on CHANCE," [Online]. (www.dartmouth.edu/~chance/course/Articles/ume.html)

Snell, J. L., and Finn, J. (1992), "A Course Called `Chance,'" [Online]. (www.dartmouth.edu/~chance/course/Articles/sf/sf.html)

Staddon, J. E. R. (1988), "Learning as Inference," in Evolution and Learning, eds. R.C. Bolles and M.D. Beecher, Hillsdale, NJ: Erlbaum.

Steinhorst, R. K., and Keeler, C. W. (1995), "Developing Material for Introductory Statistics Courses from a Conceptual, Active Learning Viewpoint," Journal of Statistics Education [Online], 3(3). (ww2.amstat.org/publications/jse/v3n3/Steinhorst.html)

Tooby, J., and Cosmides, L. (1992), "Ecological Rationality and the Multimodula Mind," Rep. No. 92-1, Santa Barbara: University of California, Center for Evolutionary Psychology.

Tversky, A., and Kahneman, D. (1974), "Judgment Under Uncertainty: Heuristics and Biases," Science, 185, 1124-1131.

----- (1980), "Causal Schemas in Judgment Under Uncertainty," in Progress in Social Psychology, ed. M. Fischbein, Hillsdale, NJ: Lawrence Erlbaum, 49-72.

----- (1983), "Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgement," Psychological Review, 90(4), 293-315.

Watts, D. G. (1991), "Why is Introductory Statistics Difficult to Learn? And What Can We Do to Make It Easier?" The American Statistician, 45, 290-291.

Well, A. D., Pollatsek, A., and Konold, C. (1983), "Probability Estimation and the Use and Neglect of Base-Rate Information," University of Massachusetts, Amherst.

Well, A. D., Pollatsek, A., and Boyce, S. (1990), "Understanding the Effects of Sample Size on the Mean," Organizational Behavior and Human Decision Processes, 47, 289-312.

West, R. W., and Ogden, R. T. (1998), "Interactive Demonstrations for Statistics Education on the World Wide Web," Journal of Statistics Education [Online], 6(3). (ww2.amstat.org/publications/jse/v6n3/west.html)

White, C. W. (1974), "A Study of the Ability of First and Eighth Grade Students to Learn Basic Concepts of Probability and the Relationship Between Achievement in Probability and Selected Factors," Dissertation Abstracts International, 35 1969A.

Wiggins, G. P. (1989), "The Futility of Trying to Teach Everything of Importance," Educational Leadership, 47(3), 44-59.
----- (1992), "Creating Tests Worth Taking," *Educational Leadership*, 49(8), 26-33.

----- (1993), *Assessing Student Performance*, San Francisco: Jossey-Bass.

Wiggins, G. P., and McTighe, J. (1998), *Understanding by Design*, Alexandria, VA: Association for Supervision and Curriculum Development.

Zetterqvist, L. (1997), "Statistics for Chemistry Students: How to Make a Statistics Course Useful by Focusing on Applications," *Journal of Statistics Education* [Online], 5(1). (ww2.amstat.org/publications/jse/v5n1/zetterqvist.html)

Carolyn Keeler
Educational Administration
University of Idaho
Moscow, ID 83844-1104
ckeeler@uidaho.edu

Kirk Steinhorst
Division of Statistics
University of Idaho
Moscow, ID 83844-1104
kirk@uidaho.edu