

**Preparation of Two-Year College
Mathematics Instructors
to Teach Statistics with GAISE**

Session on Assessment

American Association for Higher Education (AAHE)

9 Principles of Good Practice for Assessing Student Learning

<http://www.assessment.tcu.edu/assessment/aahe.pdf>

1. **The assessment of student learning begins with educational values.** Assessment is not an end in itself but a vehicle for educational improvement. Its effective practice, then, begins with and enacts a vision of the kinds of learning we most value for students and strive to help them achieve. Educational values should drive not only **what** we choose to assess but also **how** we do so. Where questions about educational mission and values are skipped over, assessment threatens to be an exercise in measuring what's easy, rather than a process of improving what we really care about.
2. **Assessment is most effective when it reflects an understanding of learning as multidimensional, integrated, and revealed in performance over time.** Learning is a complex process. It entails not only what students know but what they can do with what they know; it involves not only knowledge and abilities but values, attitudes, and habits of mind that affect both academic success and performance beyond the classroom. Assessment should reflect these understandings by employing a diverse array of methods, including those that call for actual performance, using them over time so as to reveal change, growth, and increasing degrees of integration. Such an approach aims for a more complete and accurate picture of learning, and therefore firmer bases for improving our students' educational experience.
3. **Assessment works best when the programs it seeks to improve have clear, explicitly stated purposes.** Assessment is a goal-oriented process. It entails comparing educational performance with educational purposes and expectations--these derived from the institution's mission, from faculty intentions in program and course design, and from knowledge of students' own goals. Where program purposes lack specificity or agreement, assessment as a process pushes a campus toward clarity about where to aim and what standards to apply; assessment also prompts attention to where and how program goals will be taught and learned. Clear, shared, implementable goals are the cornerstone for assessment that is focused and useful.
4. **Assessment requires attention to outcomes but also and equally to the experiences that lead to those outcomes.** Information about outcomes is of high importance; where students "end up" matters greatly. But to improve outcomes, we need to know about student experience along the way--about the curricula, teaching, and kind of student effort that lead to particular outcomes. Assessment can help us understand which students learn best under what conditions; with such knowledge comes the capacity to improve the whole of their learning.
5. **Assessment works best when it is ongoing, not episodic.** Assessment is a process whose power is cumulative. Though isolated, "one-shot" assessment can be better than none, improvement over time is best fostered when assessment entails a linked series of cohorts of students; it may mean collecting the same examples of student performance or using the same instrument semester after semester. The point is to monitor progress toward intended goals in a spirit of continuous improvement. Along the way, the assessment process itself should be evaluated and refined in light of emerging insights.

6. **Assessment fosters wider improvement when representatives from across the educational community are involved.** Student learning is a campus-wide responsibility, and assessment is a way of enacting that responsibility. Thus, while assessment efforts may start small, the aim over time is to involve people from across the educational community. Faculty play an especially important role, but assessment's questions can't be fully addressed without participation by student-affairs educators, librarians, administrators, and students. Assessment may also involve individuals from beyond the campus (alumni/ae, trustees, employers) whose experience can enrich the sense of appropriate aims and standards for learning. Thus understood, assessment is not a task for small groups of experts but a collaborative activity; its aim is wider, better-informed attention to student learning by all parties with a stake in its improvement.
7. **Assessment makes a difference when it begins with issues of use and illuminates questions that people really care about.** Assessment recognizes the value of information in the process of improvement. But to be useful, information must be connected to issues or questions that people really care about. This implies assessment approaches that produce evidence that relevant parties will find credible, suggestive, and applicable to decisions that need to be made. It means thinking in advance about how the information will be used, and by whom. The point of assessment is not to gather data and return "results"; it is a process that starts with the questions of decision-makers, that involves them in the gathering and interpreting of data, and that informs and helps guide continuous improvement.
8. **Assessment is most likely to lead to improvement when it is part of a larger set of conditions that promote change.** Assessment alone changes little. Its greatest contribution comes on campuses where the quality of teaching and learning is visibly valued and worked at. On such campuses, the push to improve educational performance is a visible and primary goal of leadership; improving the quality of undergraduate education is central to the institution's planning, budgeting, and personnel decisions. On such campuses, information about learning outcomes is seen as an integral part of decision making, and avidly sought.
9. **Through assessment, educators meet responsibilities to students and to the public.** There is a compelling public stake in education. As educators, we have a responsibility to the publics that support or depend on us to provide information about the ways in which our students meet goals and expectations. But that responsibility goes beyond the reporting of such information; our deeper obligation--to ourselves, our students, and society--is to improve. Those to whom educators are accountable have a corresponding obligation to support such attempts at improvement.

Ideas from Beth Chance (Cal Poly, San Luis Obispo, CA)

An Assessment Cycle

- **Set goals**
 - ❑ What should students know, be able to do?
 - ❑ At what point in the course?
 - ❑ Identify assessable learning outcomes that match goals
- **Select methods**
 - ❑ Identify an assessment that matches the type of learning outcome
 - ❑ Consider minute papers, article reviews, newspaper assignments, projects, short answer items, multiple choice
 - ❑ Can the assessment be built into the activity?
- **Gather evidence** (i.e., administer the assessment)
- **Draw inference**
 - ❑ Don't use results just to assign a grade
 - ❑ Consider what responses indicate about student understanding
- **Take action**
 - ❑ Provide feedback
 - ❑ What can be done to remedy a misunderstanding (an activity; extra reading; more experience with a procedure or a concept)
 - ❑ Re-examine goals and methods

Suggested Readings

Book on Assessment in Statistics Education

Gal, I. & Garfield, J. B. (1997). *The Assessment Challenge in Statistics Education*, Netherlands: IOS Press.

Articles on Alternative Forms of Assessment

Chance, B. L. (1997). Experiences with authentic assessment techniques in an introductory statistics course, *Journal of Statistics Education*, 5(3).

www.amstat.org/publications/jse/v5n3/chance.html

Fillebrown, S. (1994). Using projects in an elementary statistics course for non-science majors, *Journal of Statistics Education*, 2(2).

www.amstat.org/publications/jse/v2n2/fillebrown.html

Garfield, J. B. (1994). Beyond testing and grading: Using assessment to improve student learning, *Journal of Statistics Education*, 2(1).

www.amstat.org/publications/jse/v2n1/garfield.html

Hubbard, R. (1997). Assessment and the process of learning statistics, *Journal of Statistics Education*, 5(1).

www.amstat.org/publications/jse/v5n1/hubbard.html

Konold, C. (1995). Issues in assessing conceptual understanding in probability and statistics, *Journal of Statistics Education*, 3(1).

www.amstat.org/publications/jse/v3n1/konold.html

Articles on Statistical Literacy, Reasoning, and Thinking

Chance, B. (2002). Components of statistical thinking and implications for instruction and assessment, *Journal of Statistics Education*, 10(3).

www.amstat.org/publications/jse/v10n3/chance.html

delMas, R. C. (2002). Statistical literacy, reasoning, and thinking: A commentary, *Journal of Statistics Education*, 10(3).

www.amstat.org/publications/jse/v10n3/delmas_discussion.html

Garfield, J. (2002). The Challenge of Developing Statistical Reasoning, *Journal of Statistics Education*, 10(3).

www.amstat.org/publications/jse/v10n3/garfield.html

Rumsey, D. J., (2002) Statistical Literacy as a Goal for Introductory Statistics Courses, *Journal of Statistics Education*, 10(3).

www.amstat.org/publications/jse/v10n3/rumsey2.html



Welcome to the ARTIST website!

[Home Page](#)

[What is New?](#)

[Events](#)

[Item Database \[Assessment Builder\]](#)

[Questions and Answers](#)

[Resources](#)

[Research Instruments](#)

[Ways to participate](#)

[Feedback](#)

[ARTIST Publications and Presentations](#)

[ARTIST Glossary](#)

[About ARTIST](#)

[ARTIST Testing Page](#)

[Register to use ARTIST tests](#)

[Login to](#)

1. Request Access Code
2. Modify Access Code
3. Retrieve Reports

[National Test Results](#)

[Q&A](#)

Our goal is to help teachers assess:

- [Statistical literacy](#)
- [Statistical reasoning](#)
- [Statistical thinking](#)

in first courses of statistics.

This website provides a variety of assessment resources for teaching first courses in Statistics:

1. [Item Database \[Assessment Builder\]](#): a collection of about 1100 items, in a variety of item formats, according to statistical topic and type of learning outcome assessed. A tool called the Assessment Builder is used to access items in the database. Assessment Builder can be used to generate files that can be edited and manipulated by statistics instructors.
2. [Resources](#):
 - o Information, guidelines, and examples of alternative assessments (such as projects, article critiques, and writing assignments).
 - o Copies of articles or direct links to articles on assessment in statistics.
 - o References and links for other related assessment resources.
3. [Research Instruments](#): instruments that may be useful for research and evaluation projects that involve assessments of outcomes related to teaching and learning statistics.
4. [Implementation issues](#): questions and answers on practical issues related to designing, administering, and evaluating assessments.
5. [Presentations](#): copies of conference papers and presentations on the ARTIST project, and handouts from ARTIST workshops.
6. [Events](#): information on past and upcoming ARTIST events.
7. [Participation](#): ways to participate as a class tester for ARTIST materials.

This project is supported by the following institutions:



UNIVERSITY OF
MINNESOTA



ARTIST Assessment Builder

<http://ore.gen.umn.edu/artist/user/login.asp>

Searching the ARTIST Item Database

ARTIST Start Search Page HELP

CAUTION DO NOT use the BACK and FORWARD buttons to move from page to page ?

ARTIST Home Page

MANAGE ASSESSMENTS ?

Assessment: **Interpreting Graphs** COURSE: **Introduction to Statistics** Total Items in Assessment = 0

ITEM FORMAT	<input checked="" type="checkbox"/> Forced-Choice, Check List	<input checked="" type="checkbox"/> Open-Ended
? PERFORMANCE ITEMS:	<input type="radio"/> Include	<input checked="" type="radio"/> Exclude
	<input type="radio"/> Only performance items	
? ITEM SETS:	<input type="radio"/> Only items that meet search criteria	<input checked="" type="radio"/> ALL items in an Item Set

	LEARNING OUTCOMES		
TOPICS ?	CHECK ALL	CLEAR ALL	
CHECK ALL CLEAR ALL	<input type="checkbox"/> Literacy ?	<input type="checkbox"/> Reasoning ?	<input type="checkbox"/> Thinking ?
<input type="checkbox"/> Data Types	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/> Producing and Collecting Data	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input checked="" type="checkbox"/> Representing Data	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<input type="checkbox"/> Measures of Center	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/> Measures of Spread	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/> Comparing Groups	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/> Measures of Position	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/> Normal Distribution	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/> Bivariate Data, Quantitative	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/> Bivariate Data, Linear Regression	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/> Bivariate Data, Categorical (includes Chi-Square)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/> Probability	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Assessment Builder: Search Results

NEW SEARCH ?		MANAGE ASSESSMENTS ?
Go To Item #: <input type="text"/>	GO	NEXT SET

Assessment: **Interpreting Graphs** COURSE: **Introduction to Statistics** Total Items in Assessment = 0

ARTIST SEARCH RESULTS

Total Items Meeting Search Criteria = 152
 Items 1 to 16 are displayed on this page ?

ALL ITEMS IN AN ITEMSET ARE DISPLAYED

Check the Include box to add an item to the assessment; uncheck the include box to remove an item. ?

<p>Item ID = Q0108</p> <hr/> <p>DATA REPRESENTATION</p> <hr/> <p>REASONING</p> <hr/> <p>Last Modified March 1, 2004</p> <hr/> <p><input checked="" type="checkbox"/> Include</p> <hr/> <p>ITEM FEEDBACK</p>	<p>1. A class of students recorded the number of years their families had lived in their town. Here are two graphs that students drew to summarize the data. Which graph gives a more accurate representation of the data? Why?</p> <div style="border: 1px solid black; padding: 5px; margin-bottom: 10px;"> <p>Graph 1</p> </div> <div style="border: 1px solid black; padding: 5px;"> <p>Graph 2</p> </div>
---	--

Assessment Builder: Downloading a Test

Return to the [ARTIST Home Page](#) [Change PASSWORD](#)

REGISTERED USER: Bob delMas

ARTIST Assessment Manager

NOTE: This session will time out if you do not use the Assessment Builder for 60 minutes.

[CREATE A NEW TEST](#)

CURRENT ASSESSMENTS [HELP](#)

VIEW	DOWN LOAD	UPDATE	Assessment Name	Course Identifier	Course Name	Items	DELETE
TEST	GET	RENAME	Interpreting Graphs	GC 1454	Introduction to Statistics	7	DELETE

ARTIST Assessment Download

Return to the [ARTIST Home Page](#)

[NEW SEARCH](#) [?](#) [MANAGE ASSESSMENTS](#) [?](#)

Assessment: **Interpreting_Graphs** COURSE: **Introduction to Statistics**

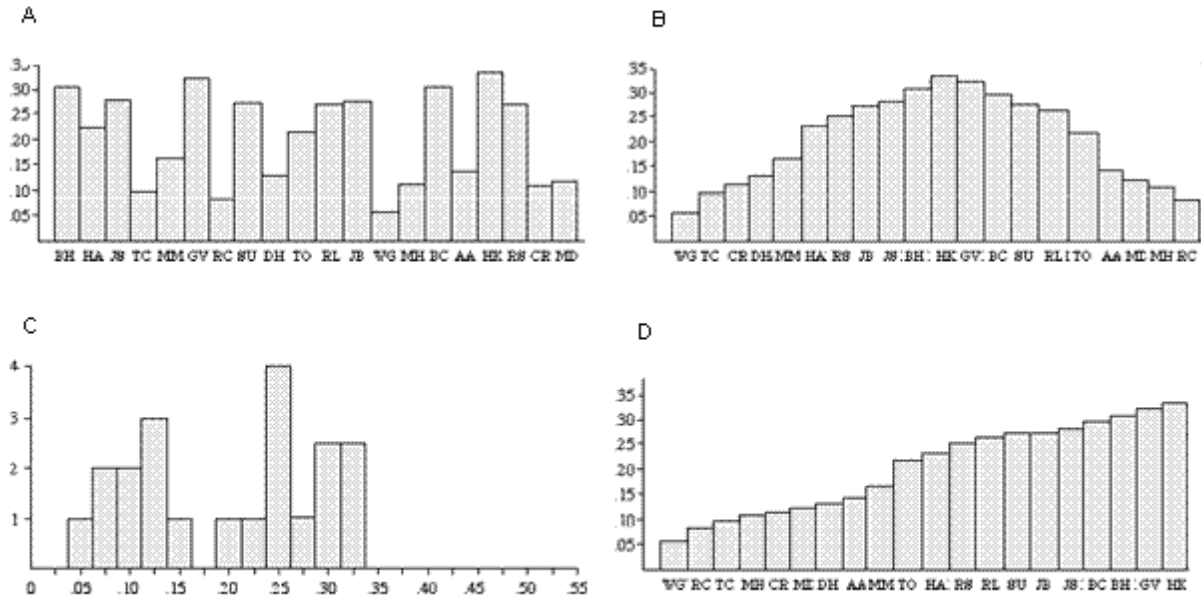
RTF File Created

[Download Interpreting_Graphs.rtf](#)

Data Representation Items

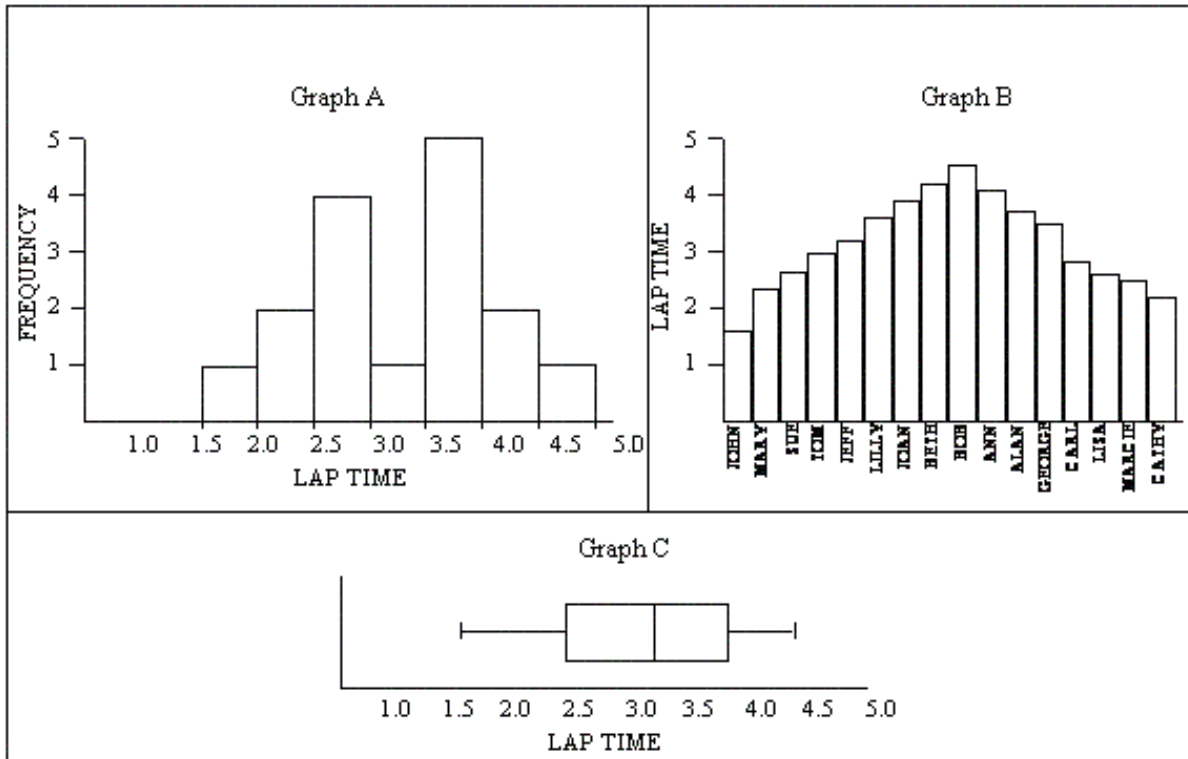
1. A baseball fan likes to keep track of statistics for the local high school baseball team. One of the statistics she recorded is the proportion of hits obtained by each player based on the number of times at bat as shown in the table below. Which of the following graphs gives the best display of the distribution of proportion of hits in that it allows the baseball fan to describe the shape, center and spread of the variable, proportion of hits?

Player	Proportion of hits	Player	Proportion of hits	Player	Proportion of hits
BH	0.305	SU	0.270	BC	0.301
HA	0.229	DH	0.136	AA	0.143
JS	0.281	TO	0.218	HK	0.341
TC	0.097	RL	0.267	RS	0.261
MM	0.167	JB	0.270	CR	0.115
GV	0.333	WG	0.054	MD	0.125
RC	0.085	MH	0.108		



RESPONSE	GUESS PERCENT	PERCENT (N = 1104)
Graph A		
Graph B		
Graph C		
Graph D		

2. A local running club has its own track and keeps accurate records of each member's individual best lap time around the track, so members can make comparisons with their peers. Here are graphs of these data.



Which of the above graphs allows you to most easily see the shape of the distribution of running times?

RESPONSE	GUESS PERCENT	PERCENT (N = 532)
A. Graph A.		
b. Graph B.		
c. Graph C.		
d. All of the above.		

Topic: Data Representation

Names of Group Members

Why do you think students are selecting the incorrect responses for each item?

Outline an instructional activity to help students develop the correct understanding.

Sampling Variability Items

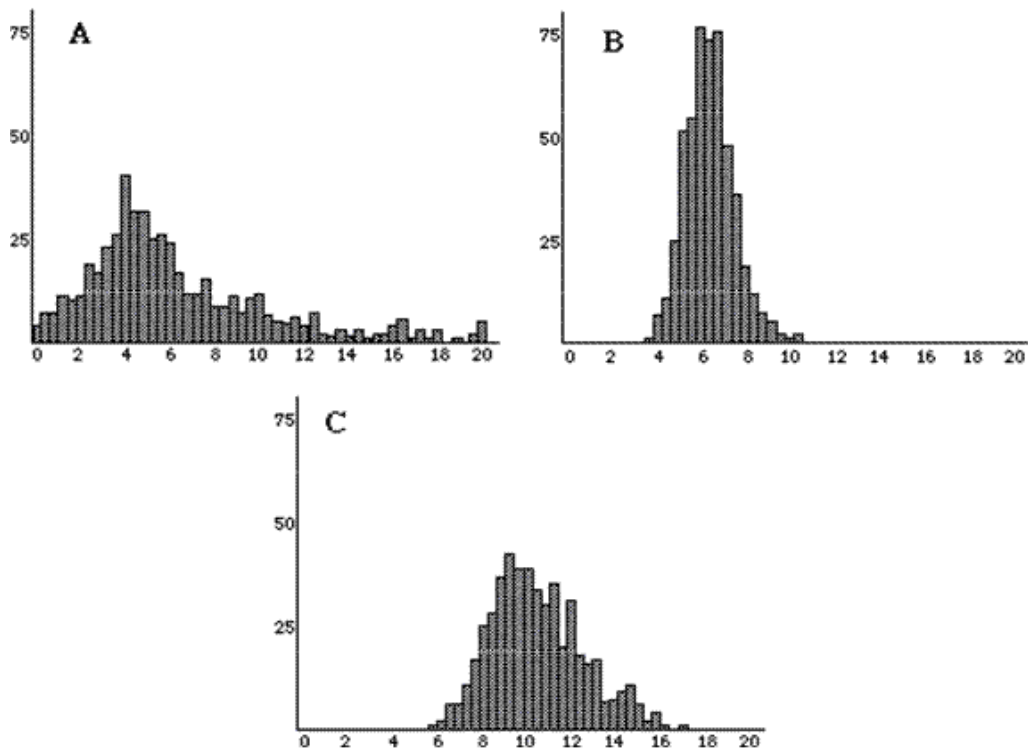
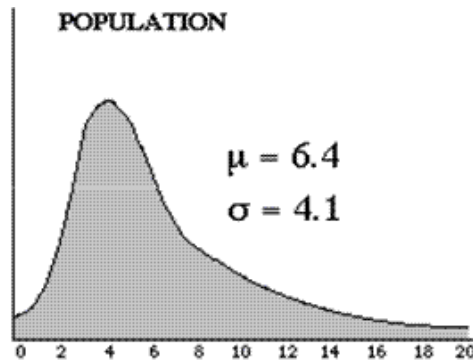
1. A certain manufacturer claims that they produce 50% brown candies. Sam plans to buy a large family size bag of these candies and Kerry plans to buy a small fun size bag. Which bag is more likely to have more than 70% brown candies?

	RESPONSE	PERCENT (N = 1107)
a.	Sam, because there are more candies, so his bag can have more brown candies.	4.0
b.	Sam, because there is more variability in the proportion of browns among larger samples.	12.5
C.	Kerry, because there is more variability in the proportion of browns among smaller samples.	30.6
d.	Kerry, because most small bags will have more than 50% brown candies.	2.5
e.	Both have the same chance because they are both random samples.	50.4

2. Consider the distribution of average number of hours that college students spend sleeping each weeknight. This distribution is very skewed to the right, with a mean of 5 and a standard deviation of 1. A researcher plans to take a simple random sample of 18 college students. If we were to imagine that we could take all possible random samples of size 18 from the population of college students, the sampling distribution of average number of hours spent sleeping will have a shape that is

	RESPONSE	PERCENT (N = 377)
a.	Exactly normal.	20.4
B.	Less skewed than the population.	35.0
c.	Just like the population (i.e., very skewed to the right).	33.7
d.	It's impossible to predict the shape of the sampling distribution.	10.9

3. The graph below in the upper left corner is a distribution for a population of test scores. The mean score is 6.4 and the standard deviation is 4.1.



Which graph do you think represents a distribution of 500 sample means from random samples each of size 9?

RESPONSE	PERCENT (N = 1106)
Graph A	25.9
Graph B	45.6
Graph C	28.5

Confidence Intervals Items

1. Suppose two researchers want to estimate the proportion of American college students who favor abolishing the penny. They both want to have about the same margin of error to estimate this proportion. However, Researcher 1 wants to estimate with 99% confidence and Researcher 2 wants to estimate with 95% confidence. Which researcher would need more students for her study in order to obtain the desired margin of error?

RESPONSE	PERCENT (N = 502)
A. Researcher 1.	48.2
b. Researcher 2.	29.3
c. Both researchers would need the same number of subjects.	10.8
d. It is impossible to obtain the same margin of error with the two different confidence levels.	11.7

2. A high school statistics class wants to estimate the average number of chocolate chips per cookie in a generic brand of chocolate chip cookies. They collect a random sample of cookies, count the chips in each cookie, and calculate a confidence interval for the average number of chips per cookie (18.6 to 21.3). Indicate if the following interpretations are valid or invalid.

Statement (N = 1101)	Valid	Invalid
We are 95% certain that each cookie for this brand has approximately 18.6 to 21.3 chocolate chips.	58.3	41.7
We expect 95% of the cookies to have between 18.6 and 21.3 chocolate chips.	48.0	52.0
If we repeated this experiment 100 times, we would expect that our sample mean would be between 18.6 and 21.3 chocolate chips about 95 times.	61.6	38.4
We are 95% certain that the confidence interval of 18.6 to 21.3 includes the true average number of chocolate chips per cookie.	72.0	28.0

Test of Significance Items

1. A researcher compares men and women on 100 different variables using a two-sample t-test. He sets the level of significance to .05 and then carries out 100 independent t-tests (one for each variable) on data from the same sample. If, in each case, the null hypothesis actually is true for every test, about how many "statistically significant" findings will this researcher report?

RESPONSE	PERCENT (N = 397)
a. 0	31.2
B. 5	37.0
c. 10	9.1
d. none of the above	22.7

2. A newspaper article claims that the average age for people who receive food stamps is 40 years. You believe that the average age is less than that. You take a random sample of 100 people who receive food stamps, and find their average age to be 39.2 years. You find that this is significantly lower than the age of 40 stated in the article ($p < .05$). What would be an appropriate interpretation of this result?

RESPONSE	PERCENT (N = 397)
a. The statistically significant result indicates that the majority of people who receive food stamps is younger than 40.	36.3
B. Although the result is statistically significant, the difference in age is not of practical importance.	39.5
c. An error must have been made. This difference is too small to be statistically significant.	17.4
d. None of the above.	6.8

3. A research article reports the results of a new drug test. The drug is to be used to decrease vision loss in people with Macular Degeneration. The article gives a p-value of .04 in the analysis section. Your friend asks you to explain what this p-value means. For each of the following statements, indicate if it is a valid interpretation of this p-value.

Statement (N = 1104)	Valid	Invalid
The probability that if the drug is tested again with a different set of people, the same result will occur.	46.5	53.5
The probability of getting results as extreme or more extreme than the ones in this study if the drug is actually not effective.	44.3	55.7
The probability that the drug is not effective.	38.2	61.8
The probability that the drug is effective.	47.5	52.5

Bivariate Quantitative Data Items

1. The number of people living on American farms has declined steadily during the last century. Data gathered on the U.S. farm population (millions of people) from 1910 to 2000 were used to generate the following regression equation: Predicted Farm Population = $1167 - .59(\text{YEAR})$. What method would you use to predict the number of people living on farms in 2050.

RESPONSE	PERCENT (N = 546)
a. Substitute the value of 2050 for YEAR in the regression equation, and compute the predicted farm population.	46.2
b. Plot the regression line on a scatterplot, locate 2050 on the horizontal axis, and read off the corresponding value of population on the vertical axis.	15.6
C. It is not appropriate to make a prediction for the year 2050 based on these data.	38.3

2. A statistics instructor wants to use the number of hours studied to predict exam scores in his class. He wants to use a linear regression model. Data from previous years shows that the average number of hours studying for a final exam in statistics is 8.5, with a standard deviation of 1.5, and the average exam score is 75, with a standard deviation of 15. The correlation is .76. Should the instructor use linear regression to predict exam scores from hours studied?

RESPONSE	PERCENT (N = 547)
a. Yes, there is a high correlation, so it is alright to use linear regression.	26.3
b. Yes, because linear regression is the statistical method used to make predictions when you have bivariate quantitative data.	27.6
C. Linear regression could be appropriate if the scatterplot shows a clear linear relationship.	40.6
d. No, because there is no way to prove that more hours of study causes higher exam scores.	5.5

Topic: _____

Names of Group Members

Why do you think students are selecting the incorrect responses for each item?

Outline an instructional activity to help students develop the correct understanding.