Statistical Literacy in the Introductory Psychology Course

Society for the Teaching of Psychology
Statistical Literacy Taskforce 2012

Corresponding Author contact information:
Susan A. Nolan, Ph.D.
Department of Psychology
Seton Hall University
400 South Orange Avenue
South Orange, NJ 07079
973-761-9484
susan.nolan@shu.edu

Statistical Literacy in the Introductory Psychology Course

Statistical Literacy Taskforce 2012 – Learning Goals

Statistical literacy is a fundamental element of psychological science in the Introductory Psychology course at the high school and college levels. All students should have a basic conceptual understanding of psychology as a science, including knowledge of variables, data, inferential statistics, and causality.

Goal One: Empiricism

1.1 Explain why psychology is an empirical science.
1.2 Explain the concept of falsifiability.
1.3 Explain the relationship between theories and hypotheses.
1.4 Explain the importance of replication in research.

Goal Two: Variables

2.1 Identify independent variables and their levels.
2.2 Identify dependent variables.
2.3 Identify or create operational definitions for variables.
2.4 Identify possible extraneous variables in a study and discuss how their effects can be minimized.

Goal Three: Data

3.1 Explain the concept and importance of standard deviation.
3.2 Identify possible threats to the reliability of a measure.
3.3 Identify possible threats to the validity of a measure.
3.4 Interpret correlation coefficients.
3.5 Construct a bar graph that clearly conveys the results of an experimental or quasi-experimental study.

Goal Four: Statistical Conclusions

4.1 Explain the purpose of inferential statistics.
4.2 Identify limitations, based on the sample used, on the generalizability of the findings of a study.
4.3 Explain how *p* values are used in behavioral research.
4.4 Explain the distinction between statistical significance and practical significance.
4.5 Explain why nothing can be “proved” in behavioral research.

Goal Five: Causality

5.1 Distinguish between experimental, quasi-experimental, and correlational research.
5.2 Provide examples to illustrate the concept that correlation does not imply causation.
5.3 Identify possible confounding variables in a study and explain how they can lead to incorrect causal assumptions.

The following guidelines are intended to clarify the scope of the statistical literacy standards for introductory psychology. Given the introductory nature of the course, the guidelines emphasize basic conceptual understanding rather than detailed technical understanding, and some concepts have been explicitly limited in scope. Although the excluded concepts are important to statistical literacy, and instructors may choose to include them, they are not essential at the introductory level.

The standards are intended to apply not only in general, but also in specific contexts. For example, for Standard 2.2, "Identify dependent variables," students should understand what a dependent variable is and also be able to identify the dependent variables within a given experiment.

Goal One: Empiricism

Students should understand that psychology is a science and that it is based on empiricism. In particular, they should understand that a scientific theory is not "just a theory," as the term is used colloquially, but rather a unifying body of knowledge incorporating empirically demonstrated facts. Although students do not need to be able to develop their own hypotheses from existing theories, they should understand that theories are strengthened or adjusted based on the results of studies with falsifiable hypotheses that were derived from those theories.

Goal Two: Variables

Students should be able to distinguish between independent and dependent variables and between qualitative and quantitative variables. For qualitative variables they should be able to distinguish between the variable itself (e.g., sex) and its levels (e.g., male and female). Students should understand that psychological variables may need operational definitions to make them explicit and measurable.

Students should understand that extraneous variables cause random error, which does not systematically affect the data but can still cause misleading results due to increased variability.

Students should be able to identify extraneous variables that could influence a particular study, and they should also be able to develop ways to limit the effects of these variables when possible. Students should understand that larger sample sizes yield more reliable results by reducing the impact of extraneous variables.

Goal Three: Data

Students should understand that standard deviation is a measure of variability and that higher levels of variability make for more ambiguous results. Students do not need to be able to calculate standard deviation, however.

Students should understand that reliability relates to the consistency of a measure and that validity relates to the accuracy of the measure. Specifically, they should understand that a valid measure is one that does in fact measure what it is intended to measure. They also should understand that a reliable measure is not necessarily valid. Their understanding should be sufficient for them to explain why validity is more important than reliability. Students do not need to be able to distinguish between different types of reliability or validity.

Students should know that the correlation coefficient *r* ranges from -1 to 1, with values further from 0 representing stronger correlations. They should also understand the distinction between positive and negative correlations.

Students should be able to construct a simple bar graph that clearly conveys the results of a study. On the *x*-axis they should be able to precisely label the independent variable and its levels, and in particular they should be explicitly aware that it should have one bar per level rather than one bar or point per participant. On the *y*-axis they should be able to precisely label the dependent variable and know to choose an appropriate range that starts at zero unless there is a specific reason to do otherwise.

Although not required, students might be expected to construct and understand other types of graphs as well, especially scatterplots to illustrate positive versus negative correlations and strong versus weak correlations. Students should understand that scatterplots are not applicable for designs with a categorical independent variable.

Goal Four: Statistical Conclusions

Students should understand that researchers want to know about populations but typically only have access to data from samples. This leads to the underlying purpose of inferential statistics, which is to use sample data to draw conclusions about populations. Students should understand that the value of using a sample to draw conclusions about a population is limited by the extent to which the sample represents the population. They should understand what sampling bias is and that it can make results invalid. They should also understand that even completely valid results may not be generalizable to other populations.

Students should understand the concept that if a predicted event does in fact occur despite being very unlikely to happen by chance, then the event probably was not coincidental. It is sufficient for students to have only an informal understanding of *p* values, so long as they can conceptualize the logic of $p < .05$ representing statistical significance.

Students should understand that although statistical significance demonstrates theoretical importance, it does not necessarily demonstrate practical importance. They should also understand that because statistics is based on probability, any statistical conclusion could be incorrect. They do not need to be introduced to a formal definition of Type I and Type II errors, however.

Goal Five: Causality

Of utmost importance to statistical literacy is the concept that correlation does not imply causation. To understand why this is, students need a good understanding that a confounding variable is one that systematically affects the results by being linked with the independent variable, meaning that any observed effects of the independent variable may in reality be effects of the confounding variable.

Students should be able to identify possible confounding variables in a study. At this point they should be most concerned with confounds resulting from lack of random assignment. For this reason they need to be able to distinguish true experimental designs from quasi-experimental or correlational designs. They should understand that because nonexperimental designs do not have random assignment, they cannot be used to establish causation. They should also understand that random assignment does not automatically provide evidence of a relationship being causal, as there still exists the possibility of procedural confounds.

[**Note:** We understand that not all Introduction to Psychology courses will be able to cover all of these goals, but we hope they provide a useful starting point for instructors designing their courses.]

The outcome of the Statistical Literacy's Subcommittee on the Introductory Psychology Course was created by the collaborative work of the following individuals:

- Kosha Bramesfeld, Maryville University, MO
- Jim Freeman, University of Virginia, VA – CHAIR
- Alexis Grosofsky, Beloit College, WI
- Brent Moore, Johnson County Community College, KS
- Susan A. Nolan, Seton Hall University, NJ – CHAIR
- Pamela Scott-Johnson, Morgan State University, MD
- Erik Wyner, Scotts Valley High School, CA

We thank the following research assistants, all graduate students at Seton Hall University, for their valuable help: Marjorie Levinstein, Katherine Moen, Jennifer Noonan, Ally Reeves, and Chelsea Reichert.