

Utah Virtual Lab: JAVA interactivity for teaching science and statistics on line

THOMAS E. MALLOY and GARY C. JENSEN
University of Utah, Salt Lake City, Utah

The Utah on-line Virtual Lab is a JAVA program run dynamically off a database. It is embedded in StatCenter (www.psych.utah.edu/learn/statsampler.html), an on-line collection of tools and text for teaching and learning statistics. Instructors author a statistical virtual reality that simulates theories and data in a specific research focus area by defining independent, predictor, and dependent variables and the relations among them. Students work in an on-line virtual environment to discover the principles of this simulated reality: They go to a library, read theoretical overviews and scientific puzzles, and then go to a lab, design a study, collect and analyze data, and write a report. Each student's design and data analysis decisions are computer-graded and recorded in a database; the written research report can be read by the instructor or by other students in peer groups simulating scientific conventions.

Early innovation (Bradley, 1978; Lehman, Starr, & Young, 1975; Rywick, 1975) and steady evolution (Bradley, 1993) have combined to lay a conceptual and practical groundwork for teachers to use computer simulations of laboratory experience as educational tools for teaching the principles of research methods. Among recent examples is the virtual lab developed by Colle and Green (1996), with graphical simulations of virtual subjects. Brown's (1999) Wildcat World allows students to design studies for doing research on human facial features. Washburn's (1999) program gives students experience in distinguishing interpretations from findings. In the Utah Virtual Lab, which is one component of StatCenter at the University of Utah, students work in an on-line virtual environment. Using on-line tools for designing research, collecting and analyzing data, and writing reports, they explore a series of interrelated research puzzles to discover the principles underlying a simulated statistical reality authored and tailored to their course by an instructor.

The Utah Virtual Lab can be used to support any empirically based content course in the social sciences in which students would benefit from discovering for themselves an accepted knowledge base in addition to learning about it in traditional ways. The Lab's authoring tool allows an instructor to create text-based research projects appropriate to any research area in which statistics such as the mean, variance, t , F , chi-square, r , and regression are appropriate. Students read the background text generated by the instructor and then use virtual tools to create experi-

mental designs and collect data that address the research questions posed in the text. The major research design and data analysis decisions made by the student are computer graded.

Malloy (2000) has argued that learning the structure (or logic) of a program is a deep and important part of what students learn through interaction with simulation software and has hypothesized that the logic of a simulation program can be used to teach students strategic thinking skills. Thus research simulation programs, to the degree that they are structured after the logic of research design principles, can—though probably to a lesser degree than actual labs—teach important research thinking skills.

Description of Virtual Lab Procedures

Public access to Virtual Lab and tutorial. The Utah Virtual Lab is a component of the University of Utah's StatCenter and can be accessed directly without login from a public access page at www.psych.utah.edu/learn/statsampler.html. The public access page has a link to an extensive tutorial on the use of the Virtual Lab. Note that the Lab itself is accessed by different links, depending on a user's preferred web browser; this is because the Virtual Lab JAVA code must be slightly different to be compatible with different browsers. Since public users do not log in to the Lab, they will not have records of data measuring their performance in research design and data analysis. Instructors and students who do not need performance data are welcome to use the Virtual Lab as public users.

Password access for educational use. The public access page provides links that allow instructors who want records of student performance to contact site managers to set up log-in procedures and database connectivity for educational groups. Site managers can also provide access to and instructions for an authoring tool that allows instructors to create their own examples and their own simulated statistical realities.

This project was supported by a Utah Higher Education Technology Initiative grant. The JAVA programming for the Utah Virtual Lab was designed and written by Gary Jensen, who also designed the Oracle database and its interaction with the JAVA applet. Correspondence regarding this article should be addressed to T. E. Malloy, University of Utah, Department of Psychology, 380 S. 1520 E., Room 502, Salt Lake City, UT 84112-0251 (e-mail: malloy@psych.utah.edu).

Example. The Virtual Lab authoring tool allows an instructor to create research contexts that differ in sophistication and are therefore appropriate for various levels of statistics courses. The example used here will be simple and straightforward, involving made-up virtual entities called “Bots,” but any content area can be authored.

It is recommended that new users, particularly students, print out and read the on-line tutorial, using its examples as a guide to navigate through Virtual Lab. The description of the Lab in this paper is not complete and is meant only to give an overview of certain educational features.

Clicking the browser-specific link to Virtual Lab downloads an applet that allows students to navigate a series of screens giving a sense of place in a virtual reality. By clicking on the front door of the Lab, students enter a virtual library. Up the stairs and on the left side of the upper level of the library are books that contain the example described here; these books are labeled “Dr. Malloy PSYCH 3000.” Clicking on the library shelves for PSYCH 3000 reveals a close-up view of several books. *Bot Res Vol 4: Explorations* is used early in a course when students have only learned basic descriptive statistics. Students select *Bot Res Vol 4: Explorations* by clicking on it and then clicking “Next.” They then see a brief overview text describing a set of hypotheses and variables. After reading this literature overview, students click “Next” to see a set of book chapters, each of which contains a research puzzle. Clicking on a puzzle piece selects a particular puzzle; for example, clicking on “The effect of language on Bot brain development” selects that puzzle piece. To read the puzzle, students click on the large blue bar that says “Read Effect of Language on Development.” That chapter proposes the following puzzle, which requires only the use of means in data analysis:

Ann, a Bot researcher, is interested in studying Bot brain development and function. You are her assistant. She hypothesizes that the number of computer languages known by Bots (they all know at least one) increases brain development in Bots. She wants you to do research to determine if training Bots to learn more languages increases Bot brain development. She would like you to randomly assign Bots to different groups who are taught different numbers of languages. Then she would like you to measure Bot brain development in each of the groups. Report the mean development for each group. Go to the Virtual Lab and do the research project she described. Make a scientific conclusion and fill out the Virtual Lab research report. Make an overall conclusion which you are willing to defend to your peers in lab next week.

After reading the research puzzle, students click “Ok” and then click “To the lab!” On the basis of course information not contained in this article or the Virtual Lab interface, students should recognize that this scientific hypothesis is causal and that they should click on the lab labeled “One IV Lab, Experimental Data.” After choosing a type of research, they click “Next.”

The graphical user interface (GUI) for the one IV experimental lab has a table with three tools; students can

pick up a tool and use it by clicking on it. The clipboard tool, which comes up automatically when the other tools are picked up, shows the experimental design and the data. The Intervener 1000 tool allows students to choose the IV and the levels of the IV. It also allows students to choose between correlated and independent groups. The DV tool allows students to select which DV they want to measure and to assign the number of participants in each experimental condition.

Selecting the independent variable. The left-hand panel of Figure 1 shows how the IV tool looks when it is picked up and indicates that in this example the instructor has given the students the choice of four IVs to study: # of Languages; Finger Skills; Learning Styles; and Music. Next to “# of Languages,” notice that the white indicator light is illuminated. This indicates that the student has chosen “# of Languages” as the IV (which is the correct IV for the current research question). Switches next to all six levels of the IV have been flipped up, indicating that the student has chosen to run a study with six groups that differ in the number (from one to six) of languages that the Bots learn. At the bottom of the left panel is a switch showing that the student wants the measurements across the groups to be independent rather than correlated.

Representing the experimental design. The panel on the right side of Figure 1 shows the clipboard GUI after it has been picked up and after the student has clicked on the “Res Design” button at the bottom. It represents the experimental design as constituted up to this point. Along the top edge of the clipboard, the IV is indicated as “# of Languages,” but the DV has not yet been chosen. In the center of the clipboard are geometric shapes representing six groups of Bots with $n = 10$ in each group. These correspond to the six levels of the IV that have been chosen.

Setting the dependent variable. The dial in the upper left corner of the left panel of Figure 2 can be rotated to point at Brain Development, Brain Electron Flow, Puzzle Solving Ability, and Brain State. Figure 2 shows that the student has turned the DV dial toward Brain Development, which is correct for the current research question. The last three DVs are distractors for the current question; but any of these distractors might be the correct choice for other research puzzles. The number of participants per group can be set with the DV tool. In the left panel of Figure 2, all six groups have been set to $n = 10$ participants (the default value); but the group size can vary from 1 to 20. Students are free to choose the size of independent groups to be equal or unequal. Along the top edge of the clipboard (Figure 2, right panel) the DV now is specified as Brain Development.

The right panel of Figure 2 shows the clipboard after the student has pressed the “Measure Bots” button. Along the bottom are the mean and standard deviation for each group. The data are pseudo-randomly generated from normal distributions whose means and sigmas have been set by the instructor. Different students therefore obtain

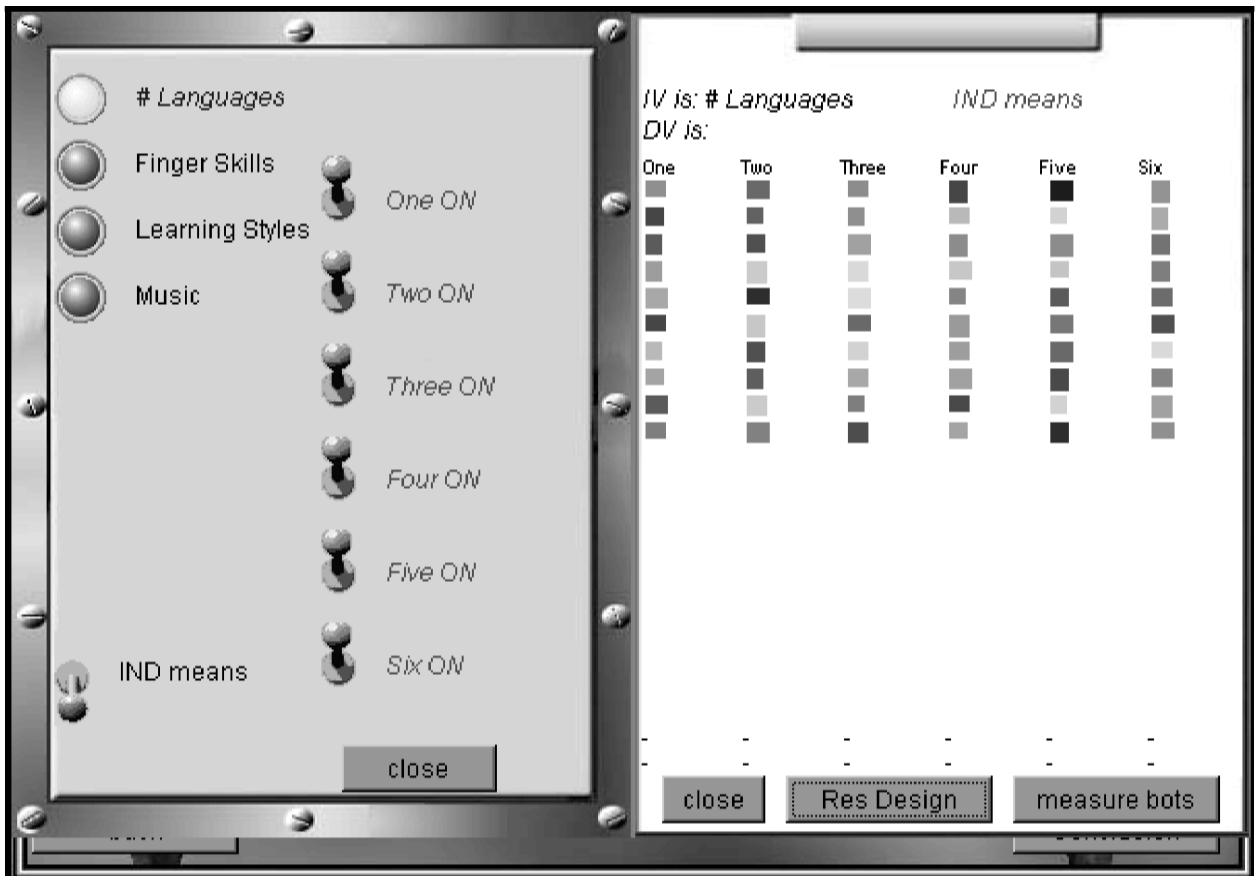


Figure 1. On the left is the tool for selecting the independent variable and its levels. On the right is the clipboard schematizing a student's research design choices.

different data. Note that, if changes are made in any aspect of the research design by further manipulation of the IV or DV tools, it is recommended that the "Res Design" button be clicked to remove old simulation data and reset the clipboard before simulating new data.

Authoring. The simulated reality that students explore is authored by instructors. Along with the text of the literature overviews and chapters, the names of the IVs and all their levels are specified by the instructor, as are the names of all the DVs. Instructors specify which IVs have causal effects on which DVs and what the effect size is. They also specify which DVs are correlated, and to what degree. Each book in the library can address different research literatures with different sets of variables and relations among variables. In this way, instructors can customize the research context so that it fits the level of sophistication that they desire. Instructors who do not have time for such effort can send their students to existing books in the library authored by other instructors. It is hoped that over time a diverse library of research areas and examples will be built up by different instructors and be made publically available in the on-line library.

Data analysis. Virtual Lab research data are automatically imported into StatTool, a simple data analysis JAVA tool built into StatCenter. On the basis of their knowledge

of the research design that they have created, students must choose the appropriate data analyses.

Conclusion and research report. When students complete their data analyses, they move on to a conclusion applet by clicking "Conclusion." They choose their class and section so that their grade is properly recorded on the database. This process will not be described in detail here, but students summarize their major research design and data analysis choices. On the basis of the simulated reality that the instructor has authored, the computer grades the students' research decisions.

After they have submitted their conclusions, students see a printable HTML page that summarizes all the information about their research, including a printout of the research question, their teacher's name, and their class section. Printing this page results in an outline of a research report that leaves spaces for students to write down observations, data summaries and analyses, and scientific conclusions. These handwritten research summaries may then be graded by the instructor.

The pedagogical goal is to have students run a series of studies that allow them to make theoretical conclusions about the nature of the virtual reality that their instructor has authored for them. This engages them in high-level scientific discovery processes. As a way to simulate sci-

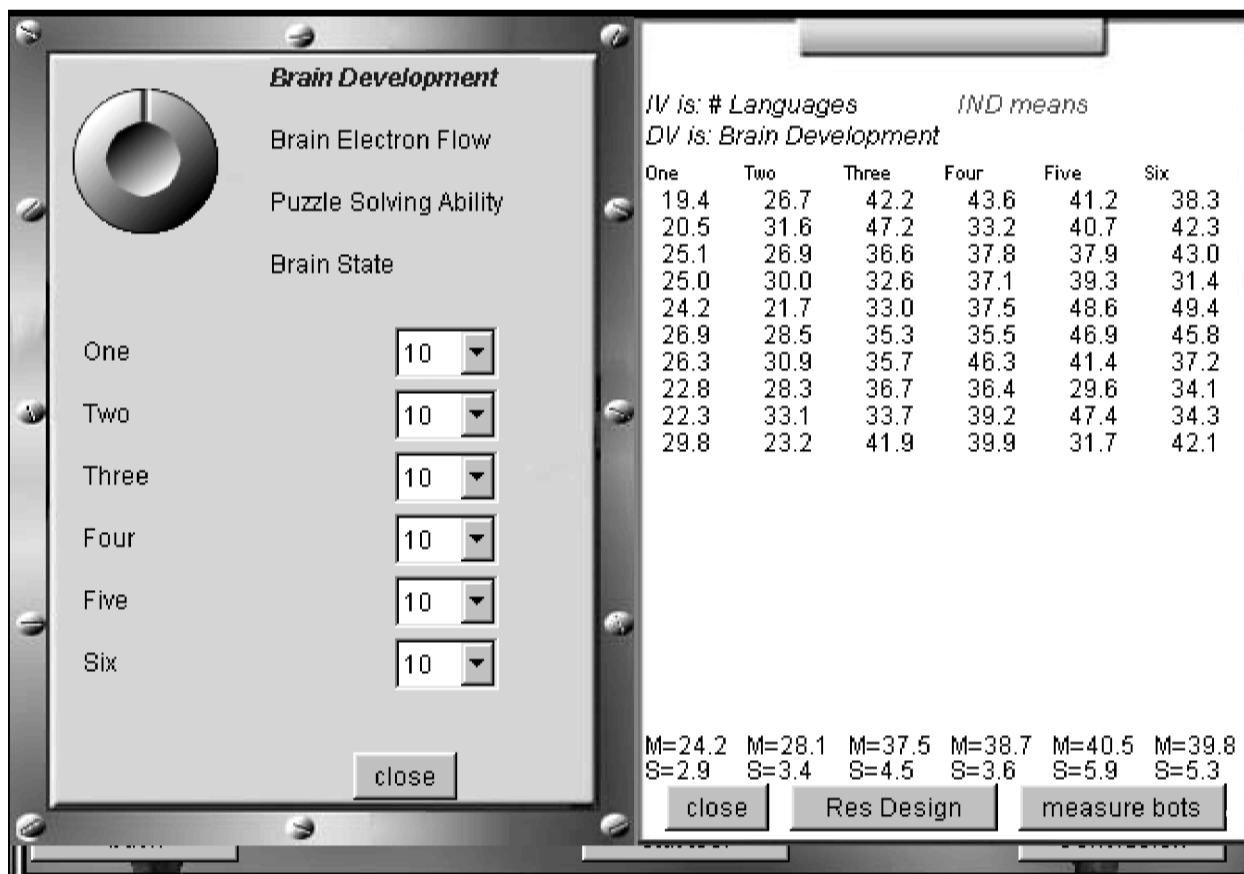


Figure 2. On the left is the tool for selecting a dependent variable. On the right is the clipboard showing the research data.

entific discussion, one useful exercise is to require students to take the series of reports generated by their research program and to work with other students in collaborative groups. Since the Virtual Lab GUI is open-ended and allows students much latitude in choosing research designs, no two students are required to do exactly the same research. Moreover, no two students, even if they do the same experiment, will get the same results, since the data are randomly generated from probability distributions authored by the instructor. In addition, the instructor can make a particular treatment effect so small that some students will find a significant effect and other students not. Peer group discussion of such diverse findings simulates the processes of the scientific community, including disagreements that may require replication.

Discussion

Utah Virtual Lab allows instructors to generate simulated realities that are based on statistical equations. These equations generate data uniquely for each student. Instructors can simulate principles and accepted research findings that they want students to learn. Like real researchers, no two students will get exactly the same results even when they do the same study. The Virtual Lab engages students in scientific discovery processes by giv-

ing them the compressed experience of making many systematic research strategy decisions with the goal of discovering the principles of a complex virtual reality. The data simulations of the Utah Virtual Lab are not as comprehensive or as sophisticated as those developed by Bradley (1993), but they are well tailored to an introductory course and are available on line.

Rice Virtual Lab (Lane, 1999) is a compilation of valuable tools for teaching and learning statistics on line. Some of these tools are data simulations and others not. The Rice Virtual Lab includes an on-line textbook (HyperStat Online) and many interactive applets. It also has case studies based on data from published studies. As a compilation of resources, Rice Virtual Lab is at the same level as and more directly comparable to WISE (Aberson, Berger, Emerson, & Romero, 1997) or to StatCenter. Much like the Rice Virtual Lab, StatCenter is a suite of free and freely available on-line resources for learning and teaching introductory statistics and research methods. It includes a full on-line statistics text in printable HTML format, a homework applet with database connectivity so that students can do homework on line and be computer graded on line, and printable written homework assignments with answer keys in PDF format. StatCenter also includes many interactive applets and simulations, such as those for the

normal and binomial distributions, detecting difference (effect size), sampling distributions, ANOVAs, and power and beta. One component of StatCenter is the Virtual Lab described in this paper.

Many educational virtual labs (e.g., Colle & Green, 1996; Hachette, Zivian, Zivian, & Okada, 1999; Keenan & Keller, 1980) are built around set research examples. The Utah Virtual Lab differs from these educationally valuable virtual labs in that it allows instructors to create custom research examples. In that regard, Virtual Lab is similar to the work of Eamon (1982). It differs from that case in its graphical virtual reality and its availability on the Web.

Whatever technology and approach motivate a virtual lab with data simulation capabilities, the active experience of students doing research, albeit virtual, adds value over what a teacher can accomplish through personal lectures because a teacher cannot give this kind of independent experience even in a one-on-one tutorial. Moreover, through new media interactivity (Malloy, 2000), virtual laboratory experiences can teach students the discovery-based thinking skills that underlie empirical research methodology.

REFERENCES

- ABERSON, C. L., BERGER, D. E., EMERSON, E. P., & ROMERO, V. L. (1997). WISE: Web Interface for Statistics Education. *Behavior Research Methods, Instruments, & Computers*, **29**, 217-221.
- BRADLEY, D. R. (1978). An interactive data-generating and answer-correcting system for problems in statistics. *Behavior Research Methods & Instrumentation*, **10**, 218-227.
- BRADLEY, D. R. (1993). Multivariate simulations with DATASIM: The Mihal and Barrett study. *Behavior Research Methods, Instruments, & Computers*, **25**, 148-163.
- BROWN, M. F. (1999). Wildcat World: Simulation programs for teaching basic concepts in psychological science. *Behavior Research Methods, Instruments, & Computers*, **31**, 14-18.
- COLLE, H. A., & GREEN, R. F. (1996). Introductory psychology laboratories using graphic simulations of virtual subjects. *Behavior Research Methods, Instruments, & Computers*, **28**, 331-335.
- EAMON, D. B. (1982). CEDATS: A cognitive experimental design and testing system. *Behavior Research Methods, Instruments, & Computers*, **14**, 142-145.
- HACHETTE, V., ZIVIAN, A. R., ZIVIAN, M. T., & OKADA, R. (1999). STAZ: Interactive software for undergraduate statistics. *Behavior Research Methods, Instruments, & Computers*, **31**, 19-23.
- KEENAN, J. M., & KELLER, R. A. (1980). Teaching cognitive processes: Software for laboratory instruction in memory and cognition. *Behavior Research Methods & Instrumentation*, **12**, 103-110.
- LANE, D. M. (1999). The Rice Virtual Lab in Statistics. *Behavior Research Methods, Instruments, & Computers*, **31**, 24-33.
- LEHMAN, R. S., STARR, B. J., & YOUNG, K. C. (1975). Computer aids in teaching statistics and methodology. *Behavior Research Methods & Instrumentation*, **7**, 93-102.
- MALLOY, T. E. (2000). Teaching deductive, inductive, and inferential logic through interactive online computer simulation. *Journal of Information Technology in Medicine* [On line]. Available: <http://www.J-ITM.com>.
- RYWICK, T. (1975). Increasing student interest by the use of interactive computer simulations. *Behavior Research Methods & Instrumentation*, **7**, 103-104.
- WASHBURN, D. A. (1999). Distinguishing interpretation from fact (DIFF): A computerized drill for methodology courses. *Behavior Research Methods, Instruments, & Computers*, **31**, 3-6.

(Manuscript received November 1, 2000;
revision accepted for publication April 23, 2001.)