2004-03-17

Using Graphics, Animations, and Data-Driven Animations to Teach the Principles of Simple Linear Regression to Graduate Students

Daniel Taylor Rowe
Brigham Young University - Provo

www.manaraa.com

USING GRAPHICS, ANIMATIONS, AND DATA-DRIVEN ANIMATIONS

TO TEACH THE PRINCIPLES OF SIMPLE LINEAR REGRESSION

TO GRADUATE STUDENTS

by

Daniel T. Rowe

A master's project submitted to the faculty of

Brigham Young University

in partial fulfillment of the requirements for the degree of

Master of Science

Department of Instructional Psychology and Technology

Brigham Young University

April 2004

BRIGHAM YOUNG UNIVERSITY

GRADUATE COMMITTEE APPROVAL

of a project submitted by

Daniel T. Rowe

This project has been read by each member of the following graduate committee and by majority vote has been found to be satisfactory.

_____   _____
Date   Paul Merrill, Chair

_____   _____
Date   Stephen Yanchar

_____   _____
Date   Stephanie Allen

BRIGHAM YOUNG UNIVERSITY

As chair of the candidate's graduate committee, I have read the project of Daniel T. Rowe in its
final form and have found that (1) its format, citations, and bibliographical style are consistent
and acceptable to fulfill university and department style requirements; (2) its illustrative
materials including figures, tables, and charts are in place; and (3) the final manuscript is
satisfactory to the graduate committee and is ready for submission to the university library.

_____   _____
Date   Paul Merrill
                                                             Chair, Graduate Committee

Accepted for the Department   _____
                                                             Andrew S. Gibbons
                                                             Department Chair

Accepted for the College   _____
                                                             Richard Young
                                                             Dean, David O. McKay School of Education

ABSTRACT

USING GRAPHICS, ANIMATIONS, AND DATA-DRIVEN ANIMATIONS

TO TEACH THE PRINCIPLES OF SIMPLE LINEAR REGRESSION

TO GRADUATE STUDENTS

Daniel T. Rowe

Department of Instructional Psychology and Technology

Master of Science

This report describes the design, development, and evaluation of the Simple Linear

Regression Lesson (SLRL), a web-based lesson that uses visual strategies to teach graduate

students the principles of simple linear regression. The report includes a literature review on the

use of graphics, animations, and data-driven animations in statistics pedagogy and instruction in

general. The literature review also summarizes the pertinent instructional design and

development theories that informed the creation of the lesson. Following the literature review is

a description the SLRL and the methodologies used to develop it. The evaluation section of the

report details the methods used during the formative and summative evaluation stages, including

results from a small-group implementation of the SLRL. The report concludes with a review of the product's strengths and weaknesses and the process' strengths and weaknesses.

ACKNOWLEDGEMENTS

I would like to express special thanks to my wife, Melissa, for her unwavering support and love. She truly eased my burdens during the long hours and late nights of this project. I would also like to thank my parents and siblings for their help and support as I completed my master's degree.

Sincere thanks and appreciation are extended to my committee: Dr. Paul Merrill, Dr. Stephen Yanchar, and Dr. Stephanie Allen. Special thanks goes to Dr. Yanchar, who acted as the subject matter expert and project sponsor. His patience and willingness to devote time and energy to this project made it a successful endeavor.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

In order to succeed in academics and industry, graduate students must have, at the very least, a basic knowledge of descriptive and inferential statistics (Giesbrecht, 1996). Most graduate programs require students to enroll in statistics courses; unfortunately, these courses are notorious for their difficulty (Dallal, 2000). Most students struggle to grasp the meaning of statistical concepts and the idiosyncrasies of statistical methodologies. Because of this problem, teachers and professors of statistics are constantly searching for better ways to help students understand concepts that are sometimes very difficult. The sheer number of periodicals and organizations whose primary purpose is to improve the teaching of statistics is evidence of this point. One statistics pedagogy journal, for example, seeks to "disseminate knowledge for the improvement of statistics education at all levels, including elementary, secondary, post-secondary, post-graduate, continuing, and workplace education" (JSE Mission Statement, n.d.). This desire to improve statistics teaching wherever and whenever it occurs is common among statistics pedagogy organizations and periodicals. The journals and organizations provide instructors with research and resources to enhance their students' learning experience; research topics range from using humor to teach statistics (Friedman, Friedman, & Amoo, 2002) to specific classroom projects that help students apply statistical understanding in real life situations (Moore, 1997).

One strategy on which statistics pedagogy research is beginning to focus is the use of static graphics, animations, and data-driven animations to enhance learning. Research has shown that, when used appropriately, graphics and animations can enhance instruction and improve learning outcomes (Anglin, Towers, & Levie, 1996; Merrill & Bunderson, 1981; Reiber, 1990). Graphics and animations are capable of fulfilling a variety of instructional purposes, including

attracting learner attention and representing abstract concepts, both of which are important in teaching statistics (Emery, 1993). In addition, new software developments are continually making animations and graphics easier and less expensive to produce (Misanchuk, Schwier, & Boling, 2000). The instructional effectiveness and relative ease of production make the use of graphics and animations an important strategy to consider when designing and developing instruction.

Statement of Purpose

The purpose of the project described in this report was to provide a web-delivered statistics lesson that employs static graphics, animations, and data-driven animations in order to help graduate students better comprehend the major concepts underlying statistical methods. The project was developed for the Instructional Psychology and Technology (IPT) Department at Brigham Young University (BYU). The scope of the project was limited to one lesson in the IPT 550 "Empirical Inquiry and Statistics" course: Simple Linear Regression. This particular lesson was selected based on the results of a needs analysis conducted during the initial stages of the project (a description of the analysis can be found in the Design and Development Process section of this report).

The IPT department at BYU requires that all students pursuing advanced degrees complete the aforementioned "Empirical Inquiry and Statistics" course. The IPT 550 course syllabus gives the following description of the course's purpose: "The course introduces students to the practice of educational and behavioral science research. . . . We will focus on elementary statistical analysis, which includes descriptive statistics, regression, correlation, probability, and statistical inference" (Yanchar, 2004, Course Description).

The overarching instructional objective associated with the Simple Linear Regression Lesson (SLRL) was as follows: "As a result of [using the SLRL], students should be able to . . . understand elementary statistical theory and procedures [of simple linear regression], select the proper statistical analysis for a quantitative study, be able to read the statistical output and interpret data . . . [related to simple linear regression]" (Yanchar, 2004, Course Objectives). General instructional objectives for the SLRL, including specific learning outcomes, were developed with the assistance of statistics professors in the IPT department. A detailed list of these objectives and outcomes can be found in Appendix A.

Several instructional, evaluative, and assessment materials were developed in conjunction with the SLRL. Deliverables associated with the final product included low-fidelity, mid-fidelity, and high-fidelity prototypes of the SLRL used in the design and development process, a simple linear regression test, and a follow-up questionnaire. These materials were created to guide the design and development processes, to provide data to answer the evaluation questions of stakeholders, and to establish a base for future research, design, and development in graphics and animations used in statistics pedagogy.

<center>Description of Target Audience</center>

The target audience for the SLRL was graduate students enrolled in IPT 550. The students in this course generally range from 23 to 55 years of age and have a wide variety of experiences and backgrounds (S. Yanchar, Personal Communication, September 4, 2003). They come from different areas of undergraduate study. Some students enrolled in this course have prior experience with statistics at an undergraduate level, but for other students, this course is their first exposure to statistical concepts (S. Yanchar, Personal Communication, September 4, 2003).

In addition to the target audience, three stakeholder groups were impacted by and associated with the design and development of the SLRL: professors of statistics, instructional designers and developers overseeing the project, and professors and employers who expect IPT 550 students to know and apply the content from the course. Stakeholders were surveyed and interviewed before design and development of the SLRL began. Their input facilitated the creation of lists of evaluative questions, criteria, and standards that guided design and development and were used to formatively evaluate the lesson.

Literature Review

The analysis phase of the project consisted of a formal literature review following the methods recommended by Osguthorpe (1985). Osguthorpe recommends dividing the literature review into three fundamental parts: an instructional materials review, a content review, and an instructional theory review. A summary of the information pertaining to each section of the literature review is included below.

Instructional Materials Search Review

Osguthorpe (1985) writes that in part one of the literature review "designers should attempt to build on existing instructional materials by gaining a broad understanding of presently available [instructional] materials" (p. 20). The materials search thus focused on web-delivered statistics instruction available to students and teachers of statistics. A list of websites was compiled through conventional web searching, a review of statistics pedagogy literature, and interviews with statistics professors. Each site was reviewed and analyzed based on the following five questions posed by the project's stakeholders: (a) are the instructional materials appropriate for the target audience? (b) do the instructional materials stand-alone? (c) do the instructional materials provide explanations through graphics and animations? (d) do the instructional

materials give users the opportunity to manipulate statistical data? and (e) do the instructional materials provide lesson content on simple linear regression? The following paragraphs provide an overview of the information found during the instructional materials search review.

A conventional web search revealed hundreds of websites containing statistics instruction. The overwhelming majority of these websites simply display the contents of statistics textbooks. While these sites do fit the materials search criteria in their stand-alone functionality, they do not contain any graphics, animations, or data-driven animations that go beyond the contents of a traditional statistics textbook. The purpose of this project was to provide materials that go above and beyond the traditional, static textbook instruction that is common in statistics pedagogy.

West and Ogden (1998) provide an overview of possibilities for using web-delivered "interactive demonstrations," or data-driven animations, in statistics instruction. They describe products that, in general, meet the instructional objectives associated with the SLRL. West and Ogden include in their report a list of uniform resource locators (URLs) that access exemplary materials of this sort. Unfortunately, of the three websites listed, only one is still functioning as of January 2004. The functioning site contains a series of graphical demonstrations that change based on user input, or data-driven animations. Therefore, these demonstrations meet the criteria in that they give users the opportunity to manipulate statistical data. However, the data-driven animations are meant for teacher presentation and, as a result, are not stand-alone. In addition, the materials addressing simple linear regression are not sufficiently thorough to meet the objectives of the SLRL. Nonetheless, these data-driven animations provided an example on which to build and expand during the development of the SLRL.

In 1998, the Center for Instructional Design (CID) at BYU developed a series of animated presentations available on the WWW to students and instructors of the University's Statistics 221 courses (J. McDonald, Personal Communication, September 5, 2003). These lessons contain content similar to the kind of instructional material that this project was intended to provide. Each lesson contains useful graphics and animations that help the student visualize the statistical concepts they are attempting to master. They also contain sophisticated data-driven animations that are dependent on user input. However, the lessons developed by the CID fall short of the criteria given above in that they are meant to be presentation materials and not stand-alone study materials. In addition, the lessons do not meet the needs of the specified stakeholders and target audience for this project because of differing instructional objectives and lesson content. Despite this drawback, the Statistics 221 materials did provide a foundation of principles and functionality on which to rely as the design and development of the SLRL began.

In summary, considering the vast number of people committed to improving statistics pedagogy, it is surprising that so few products are currently available that meet the criteria specified by the designers and stakeholders of this project. West and Ogden (1998) make the following point about statistics graphics found on the WWW: "The information contained at most WWW locations pertaining to statistics education is not significantly different from what one encounters in a classical printed medium. While such information is certainly useful and easily accessible, it is static in that it cannot respond to user input" (para. 3). The instructional materials search revealed a number of products that fit some of the criteria, but none that fit all of the criteria. Generally, the found materials fit into two categories: (a) materials that were stand-alone but did not contain animations and data-driven animations, and (b) materials that contained animations and data-driven animations but were not stand-alone. The instructional materials

search revealed that there is a need for the type of lesson described by the project's stakeholders. It also provided the designer with numerous examples of the types of products that are available to statistics teachers and students. These examples served as the foundation on which the SLRL was built.

Instructional Content Review

The second portion of the literature review outlined by Osguthorpe is meant to provide a survey of "current research and theory related to [the project's] content" (1985, p. 20). In accordance with this guideline, the instructional content review for this project focused on the following areas of research: (a) graphics, animations, and data-driven animations in statistics pedagogy, and (b) methods and strategies for teaching simple linear regression.

Graphics, Animations, and Data-Driven Animations in Statistics Pedagogy

In this section it is necessary to define what is meant by "static graphics," "animations," and "data-driven animations." Misanchuk et al. define graphics broadly as "anything that isn't text" (2000, Graphics, Definition, p. 2). Given the context of this project and the nature of its content, this definition is too ambiguous. Traditional statistical symbols and equations would fall under this definition, and the purpose of the SLRL was to go beyond the use of these traditional symbols in instruction. For the purpose of this report, "graphics" will refer to pictures, diagrams, and data graphics (charts and graphs).

Statistics instructors and textbook authors have used static graphics for over a hundred years. In fact, some statistical graphics are almost universally accepted and used in pedagogy. These graphics include histograms, box plots, Venn diagrams, and, of course, the ubiquitous bell curve. These and other statistical graphics give students conceptual models of real data that improve their recall of conceptual information and increase their ability to find creative solutions to transfer problems (Mayer, 1989). Hong and O'Neil (1992) demonstrate that extensive

graphical representation facilitates students' problem solving and understanding of statistical concepts. In addition, Hong et al. (1992) suggest that repeated use of graphical models in statistics pedagogy makes students' knowledge more accessible, functional, and improvable.

To animate a graphic is to give it "movement or life" (Misanchuk et al., 2000). This is accomplished by stringing together a sequence of still graphics or pictures and showing them in rapid succession in order to give the illusion of real movement. This definition is widely accepted among designers and researchers (Misanchuk et al., 2000; Reiber, 1990; Velleman & Moore, 1996). Misanchuk et al. make the following distinction about instructional animations: "Digital animation for instruction can be classified as fixed-path animation (pre-programmed by the software developer), or data-driven (controlled by the constantly changing data drawn from user input)." For the purpose of this report, "animation" will refer to fixed-path animation that is pre-programmed by the developer. Data-driven animations will be discussed below.

There is relatively little research concerning the use of animation in statistics instruction. Thisted and Velleman (1992) envision computer animations that allow students to observe lively experimentation and rapid replication of results (as cited in Mathieson, Doane, & Tracy, 1995). In addition, Velleman and Moore (1996) posit that carefully designed animations synchronized with text in the form of spoken narration can be a very effective way of illustrating statistical and mathematical concepts and demonstrating problem solving procedures. In summary, while there is very little research pertaining to animations and statistics pedagogy, there is a large body of research addressing general principles of instructional animation. This research is summarized below in the "Instructional Theory Review" section.

Finally, "data-driven animations" are graphical sequences that are manipulated "by the constantly changing data drawn from user input" (Misanchuk et al., 2000, Animation,

Introduction, p. 5). Because this project seeks to provide statistics students with interactive graphics that respond to user input, the term "data-driven animations" will be used throughout this report.

Many researchers have discussed the potential of data-driven animations to increase the power of statistical pedagogy (including classroom presentation) and to aid students' understanding of statistical concepts (West & Ogden, 1998; Aberson, Romero, Berger, & Healy, 2000; Velleman & Moore, 1996; Mathieson, Doane, & Tracy, 1995). Velleman and Moore (1996) write, "the most effective use of [statistical] animations is to construct objects that a student can manipulate to learn about a concept" (p. 220). Data-driven animations are now being developed with the hope that they will enhance statistics pedagogy (West & Ogden, 1998; Aberson et al., 2000; Velleman & Moore, 1996; Mathieson et al., 1995). These animations are primarily scatter plots and distribution graphs that users can manipulate by entering and modifying specific numbers (West & Ogden, 1998; Mathieson et al., 1995). Students who use these animations report that they are beneficial to their learning, but very little formal research has been conducted regarding the data-driven animations' effects on learning outcomes (West & Ogden, 1998; Aberson, Romero, Berger, & Healy, 2000). This project therefore seeks to provide information about this topic. A review of the general theories and research in the areas of instructional graphics, animations, and data-driven animations is contained below in the "Instructional Theory Review" section.

Methods and Strategies for Teaching Simple Linear Regression

Several documented methods and strategies have been shown to enhance students' understanding of simple linear regression. The most widely recognized graphical strategy is the ubiquitous use of the Venn diagram to help students perceive the variance in one variable that is

accounted for by the variance in other variables (Ip, 2001). Introduced by Venn (1880), the Venn diagram has been used consistently in statistics texts and by statistics instructors. Research has demonstrated that using Venn diagrams while teaching certain statistical concepts, including regression, increases students' understanding (Ip, 2001; Kennedy, 2002; Warner, Pendergraft, & Webb, 1998).

Another visual strategy for teaching simple linear regression is described by Velleman and Moore (1991): using a "toy" (data-driven animation) developed by Rubin and Bruce (1991, p. 220). According to Velleman and Moore (1991), "the student can quite easily find the least squares line by repositing the [regression] line on the scatter plot" (p. 220). This simple data-driven animation allows students to visualize certain assumptions about the regression line and the standard error of the estimate.

West and Ogden (1998) created a visual method to help students understand how individual points can affect the regression line and which outliers exhibit maximum effect. In this data-driven animation, students are able to move data points to different locations on the scatter plot. Each time a data point is moved, the computer redraws the regression line. Students are thus able to comprehend the principle that the regression line comes as close as possible to each point in the data set.

These three methods and several basic statistics texts were reviewed and analyzed in order to inform the design of the presentation, content, and sequence of instruction in the SLRL. The reviewed texts included (a) *Research Methods and Statistics: An Integrated Approach* (Furlong Lovelace, & Lovelace, 2000), and (b) *Fundamental Statistics for the Behavioral Sciences* (Howell, 1999). Two subject matter experts were also interviewed about their use of teaching strategies specific to simple linear regression. These interviews revealed a few

instructional preferences of statistics instructors, including specific graphical strategies and concept label methods. These strategies and methods played an important role in the design and development of the SLRL; they provided examples that the designer used in the creation of instructional methods and sequences.

Instructional Theory Review

The third portion of the literature review as outlined by Osguthorpe is the instructional theory review. The purpose of the theory review is to "provide perspective advice to the designer" (1985, p. 21). Osguthorpe recommends that the theory review focus on literature that addresses one or more of the following topics: (a) general principles of instructional design, (b) theory and research related to a specific category of learning, and (c) principles associated with a particular delivery system.

Instructional Design and Development Process

The design and development process used for the SLRL was the widely recognized model introduced by Dick and Carey: the "systems approach model for designing instruction" (Dick, Carey, & Carey, 1996, p. 30). The Dick and Carey model proposes eight stages of design and development: (a) determining instructional goals, (b) conducting instructional analyses, (c) determining entry behaviors, (d) creating performance objectives, (e) deciding on instructional strategies, (f) creating instructional materials, (g) conducting formative evaluations, and (h) conducting a summative evaluation. This model was slightly adapted to fit the circumstances of this instructional project. For example, the formative evaluation phase of the project was conducted with the use of a series of rapid-prototypes, an evaluative strategy described by Rosson and Carroll (2002).

The SLRL project began with the idea that carefully designed graphics, animations, and data-driven animations would enhance students' understanding of entry-level statistics material. This idea was confirmed in instructional design and statistics pedagogy literature. Therefore, the project and its content were negotiated with specific instructional strategies already in mind, even though selection of instructional strategies is actually stage six of the Dick and Carey model (Dick et al., 1996). Once the subject matter expert (SME) and the designer (the author) agreed that using graphics, animations, and data-driven animations would be an appropriate instructional strategy, they conducted a needs analysis. This analysis included: (a) a survey of students who had previously taken the IPT 550, and (b) personal interviews with statistics professors in the IPT department. Individuals who participated in the surveys and interviews were asked to identify specific statistical principles whose presentation would be enhanced with the use of graphics, animations, and data-driven animations.

Based on the data collected from these sources, the SME and the designer concluded that "simple linear regression" would be an appropriate statistical topic for the project. They then created a list of instructional goals and project objectives (stage one: instructional goals) and the designer conducted task and audience analyses in order to reveal the nature of the learning task and the characteristics of learners in the target audience (stage two: instructional analysis; and stage three: entry behaviors and learner characteristics). Data collected for the task analysis came from three sources: (a) a review of introductory statistics texts, (b) a review of the SME's presentation notes for the simple linear regression unit in IPT 550, and (c) a personal interview with the SME. Using the information provided by these sources, the designer created an outline of principles and concepts to be covered in the SLRL. For the audience analysis, the designer interviewed the SME concerning the entry behaviors and characteristics of students in IPT 550.

The SME had previous experience teaching the course and was therefore familiar with the characteristics of the target audience. In addition, the designer had taken the course previously and was able to offer insight from a student's perspective.

Once these analyses were complete, the designer created a series of instructional objectives for the SLRL (stage four: performance objectives). In alignment with the work of Gronlund (1991), the instructional objectives were stated at two different levels of specificity: general instructional objectives (GIOs) and specific learner outcomes (SLOs). The objectives were reviewed by the SME and modified by the designer in a series of iterations. A final version of the list of GIOs and SLOs can be found in Appendix A. Using these instructional objectives, the designer developed a test on simple linear regression (stage five: criterion-referenced test items). This test was later reviewed and approved by the SME. A final version of this test is located in Appendix B.

Stages seven (development of instructional materials) and eight (formative evaluation) were conducted simultaneously, in accordance with the rapid-prototyping approach to development and evaluation (Rosson & Carroll, 2002). The summative evaluation (stage nine) of the SLRL occurred during an implementation of the product in the IPT 550 course taught in fall semester of 2003.

Principles of Instructional Design

The general principles of instructional design used in the SLRL were based on the writings of Gagne, Briggs, and Wager (1992). In *Principles of Instructional Design,* they outline nine events of instruction based on the assumption that "instruction consists of a set of events external to the learner designed to support the internal processes of learning" (Gagne et al., 1992, p. 189). These "nine events" are widely known and recognized by students and practitioners of

instructional design: (a) gaining attention, (b) informing the learner of the objective, (c) stimulating recall and prerequisite learning, (d) presenting stimulus material, (e) providing learner guidance, (f) eliciting performance, (g) providing feedback, (h) assessing performance, and (i) enhancing retention and transfer. Important aspects of the nine events are summarized below.

Gagne et al. (1992) suggest several ways of gaining and maintaining learner's attention (event one), including presenting a stimulus change and appealing to the specific interests of learners. Visually engaging graphics and animations can also be this stimulus change and can focus the attention of learners (Misanchuk et al., 2000).

The second event in the instructional sequence is informing the learner of the objective. According to Gagne et al. (1992), learners should understand "what kind of performance will be used as an indication that learning has, in fact, occurred" (p. 191). An important principle of informing learners of the objectives is putting objectives in terms that they can readily understand. Depending on the nature of the instructional materials and the target audience, learners may need to be reminded of specific instructional objectives after the instruction has started.

The recall of prerequisite capabilities (event three) is "critical and essential" to successful learning (Gagne et al., 1992, p. 192). This is based on the assumption that most "new learning" is simply the combining of ideas. It may be best accomplished by asking recall or recognition questions (Gagne et al., 1992).

The presentation of stimulus material (event four) is the crux of the instructional content within the nine events. Here learners should view explanations, examples, and exercises, and should participate in a variety of strategic instructional activities (Gagne et al., 1992). Careful

analysis of specific learning tasks should guide the selection of presentation strategies used in this stage of instruction. Gagne et al. (1992) state that the presentation of the material must engage and maintain the attention of learners.

According to Gagne et al. (1992), the instructional materials should provide learners with appropriate guidance (event five). Learner guidance should be designed to suggest a "line of thought" that will lead learners to make appropriate conclusions about the principles and concepts being taught (Gagne et al., 1992, p. 192). This type of guidance helps learners to "keep on the track," and thus contributes to the efficiency of learning (Gagne et al., 1992, p. 192).

The sixth event (eliciting performance) in the instructional sequence is the "show me" or "do it" portion of the instruction (Gagne et al., 1992, p. 192). Learners should be required to perform the necessary tasks to the degree prescribed by the instructional objectives. This event serves as an opportunity for learners to become "convinced" that they have acquired new knowledge or skills (Gagne et al., 1992, p. 192).

Once performance is elicited, instructional materials should provide learners with feedback on the correctness of their performance (event seven). In the case of web-based instruction, feedback should be immediate and should appear on the same screen that elicited the performance (Clark & Mayer, 2003).

During the eighth event of instruction, learners' performance should be assessed according to the instructional objectives. Care should be taken to ensure that performance assessments are valid and reliable (Gagne et al., 1992).

The final event in the instructional sequence is "enhancing retention and transfer" (Gagne et al., 1992, p. 198). Knowledge retention can be enhanced by practicing data recall, providing a

network relationship for newly learned material, and creating stimulus clues (Gagne et al., 1992, p. 198)

Although the final version of the SLRL did not follow the nine events of instruction *exactly* as prescribed by Gagne, Briggs, and Wager, the designer did view this outline of events as a guide on how to accomplish instructional goals and sequence instructional events in the SLRL. The actual implementation of the events in the SLRL is described under "Description of the Product."

Theory and Research Related to Graphics, Animations, and Data-Driven Animations

Graphics. Research studies in education and psychology have established the effectiveness and importance of graphics in instruction (Anglin et al., 1996; Clark & Mayer, 2003; Merrill & Bunderson, 1981; Misanchuk et al., 2000). Emery (1993) asserts that graphics serve the following purposes in instruction: "attracting learner attention, directing learner action, illustrating quantifiable constructs, representing concrete objects, representing abstract concepts, and simplifying complex information" (p. 21). Several studies indicate that students who learn with instructional graphics score higher on tests of comprehension (Emery, 1993; Peck, 1987). In addition, when graphics are used, students understand concepts more clearly (Alesandri, 1984; Emery, 1993; Levie, 1984).

While there are many advantages to using graphics and visuals in instruction, the following caution is repeated throughout the literature: poor use of graphics and visuals is distracting and interferes with learning (Anglin et al., 1996; Emery, 1993; Merrill & Bunderson, 1981; Misanchuk et al., 2000; Clark & Mayer, 2003). Researchers suggest that the most frequent misuses of graphics in instruction include misleading graphics, non-contextual graphics, graphics whose significance is overly redundant, and gratuitously used graphics (Anglin et al., 1996;

Emery, 1993; Merrill & Bunderson, 1981; Misanchuk et al., 2000). Given the instructional objectives of the SLRL, carefully designed and properly executed graphics were essential to the product's instructional success.

Animations. Instructional animations, like instructional graphics, can be powerful instructional tools when used appropriately. Research indicates that there are several benefits to using animations in instruction. Most scholars agree that one of its most powerful effects is that it focuses learners' attention (Misanchuk et al., 2000; Reiber, 1990). In addition, animations can illustrate concepts and processes that static graphics simply cannot illustrate (Misanchuk et al., 2000). Others suggest that properly executed animations increase learners' incentives to participate in instruction, and raise the credibility of the instruction [from the participants' perspectives] (Xiaoquan & Jones, 1995).

Reiber (1990) emphasizes that animation should be used in instruction only when it is appropriate. He identifies three things to consider when employing animation in instructional products: (a) animation should only be used when it is consistent with the learning task, (b) animation should be used to illustrate concepts and processes that learners cannot visualize on their own, and (c) animation should serve the dual purpose of attracting attention and illustrating content material (Reiber, 1990). Similar to graphics, animations used inappropriately are distracting and annoying to users (Spool, Scanlong, Schroeder, Snyder, & DeAngelo, 1997). This includes animations that are excessive and frivolous, and those contain repetitive motion (Misanchuk et al., 2000).

Data-driven animations. Data-driven animations constitute a relatively new area of research and theory in instructional design. Current literature suggests that data-driven animations can be powerful learning tools (Misanchuk et al., 2000; Reiber, 1990; Velleman et

al., 1996). For example, Reiber (1990) suggests that animation's greatest contribution to instruction may be in the area of interactive graphic applications. Misanchuk et al. posit that instructional designers should "produce animations that allow learners to manipulate elements and see the results of their interventions" (2000, Animation, Pros and Cons, p. 3). Other significant research in the area of data-driven animations is summarized above in the "Instructional Content Review" section.

Principles Associated with Web-Based Tutorial Instruction

In an attempt to make the SLRL an effective resource for students and as instructionally sound as possible, several principles of web-based instruction were strategically implemented in its design and development. These principles were derived from Clark and Mayer's book, *E-Learning and the Science of Instruction*. Those that were especially pertinent to the SLRL include (a) the contiguity principle, (b) the modality principle, (c) the redundancy principle, (d) the personalization principle, and (e) the practice principle (Clark & Mayer, 2003).

According to Clark and Mayer (2003), the principle of contiguity suggests the following practical guidelines concerning web-based instruction: (a) corresponding words and graphics should be placed near each other, (b) feedback should appear on the same screen as questions, and (c) activity instructions should appear on the same page as the exercise to be completed.

The modality principle states that words should be presented in the form of audio narration rather than on-screen text so that learners can focus on graphics and animations (Clark & Mayer, 2003). Research indicates that students are unable to attend to an animation that utilizes text-based explanations. Text should be used for information that learners may need to refer to, such as instructions for practice exercises (Clark & Mayer, 2003).

The redundancy principle is that the presentation of instructional content in the form of written text *and* audio narration can actually hinder learning (Clark & Mayer, 2003). As stated about the modality principle, graphics and animations should be described by words presented in the form of audio narration alone, not by audio narration and redundant text (Clark & Mayer, 2003).

Clark and Mayer (2003) also emphasize that appropriate implementation of the personalization principle can be a powerful strategy in instructional products. They recommend that "virtual coaches" or "agents" introduce the instructional content (p. 131). The following principles apply to the use of agents in instruction: (a) agents may be visually realistic or line art, (b) agents need not be on screen throughout the instructional sequence, (c) agents' dialogue should be presented via audio narration, (d) agents' narration should be natural and conversational, and (e) agents should serve a valid instructional purpose (Clark & Mayer, 2003).

Finally, according to Clark and Mayer (2003), learners should have frequent opportunities to test their knowledge with practice exercises. The principle of practice implies that: (a) learners should have several realistic practice exercises per topic; (b) practice situations should be distributed throughout the lesson, not just in a single location; (c) each practice exercise should include clear instructions; and (d) feedback should be provided when appropriate (Clark & Mayer, 2003).

In summary, the literature review confirmed the initial hypothesis of the designer and the SME: that carefully designed graphics, animations, and data-driven animations can enhance learning. The instructional materials search proved to be an important part of the design process because it provided the designer with numerous examples of animations and data-driven animations. Furthermore, the review of pertinent instructional theory provided a framework for

the instructional strategies implemented in the final version of the SLRL. The next section contains an explanation of the graphics, animations, data-driven animations, and instructional methods used in the SLRL.

Description of the Product

Application Architecture

The SLRL was developed primarily with the tools of Macromedia Flash MX. Flash's animation tools (e.g., motion tweening) and programming language (Actionscript) provided the ideal means for creating the SLRL's animations and data-driven animations. Each section of the product (see below) was produced as a separate Flash movie and exported from the Flash tools in "Shockwave Flash" (SWF) format. These movies contained the entirety of the lesson's content, including graphics, animations, data-driven animations, audio files, and navigation tools. The SWF files were embedded in HTML documents and saved to the IPT department server, thus enabling students and instructors to access the SLRL using a standard web browser.

SLRL Modules

The SLRL is divided into six modules, the designation and content of which were based on the results of the "Instructional Content Review," interviews with the SME, and the lesson's instructional objectives. The modules were designed in such a way that students can either proceed through lesson material in the prescribed order, or review selected sections in any order. The SLRL recommends that students proceed in the prescribed order if it is their first time using the product. During subsequent sessions, students are encouraged to review the material in any order. The SLRL contains the following modules: (a) introduction to simple linear regression, (b) review of correlation, (c) the regression line, (d) the assumptions of simple linear regression, and (e) prediction error and the standard error of the estimate. An additional section called "the

coefficient of determination" was designed and partially developed, but was not included in the small group implementation due to time constraints. The SLRL also contains a welcome page with an animated agent that provides introductory learner guidance, and a conclusion designed to increase retention and transfer. The SLRL modules feature audio narration by the agent, which presents lesson content and provides learner guidance when necessary.

Module Templates

Each module of the SLRL has five parts and follows the same basic instructional template based loosely on the nine events of instruction (Gagne et al., 1992). The five parts of each module are (a) module objectives, (b) presentation of module content (agent-narrated graphics and animations), (c) data-driven animation (and instructions for proper use), (d) content questions and feedback, and (e) module summary and conclusion.

The navigation tool shown in figure one allows users to navigate through each part of the modules. With these tools, users can move forward ("next" button) and backward ("back" button) through the lesson as they desire. When the current part of the lesson has been completed, the "next" button flashes yellow to inform users that they can continue when they are ready. The navigation tools serve a strategic purpose in allowing users to proceed through the lesson at their own pace. For example, part two of the module template (presentation of the content) often contains many minutes of audio narration accompanied by graphics and animations. This large amount of material is divided into manageable sections of no more than two minutes. Small circles on the navigation bar represent these sections of material (the section that the student is currently viewing is indicated with a yellow circle). The circles are active buttons that move the users directly to a specific section of the module.

<u>Figure 1.</u> Detail of the SLRL navigation controls.

These navigation features allow users to proceed through lesson materials at a pace that suits them, quickly browse the content of the module, and review specific sections as often as needed. The functionality of the navigation tools was refined during the several rounds of usability and prototype testing described in the evaluation section.

Module Descriptions

A brief description of each lesson module, and the welcome and conclusion sections, is provided below. An explanation of the frame-by-frame details of the animations is not included in the product description; the entirety of the animated sequence can be viewed in the SLRL itself.

Welcome screen and menu. Using the appropriate URL, users are taken directly to the "welcome" section of the SLRL. The first screen is a sound check screen that tells users to put on their headphones before continuing. Users can press the "sound check" button to test the functionality and volume of their headphones. The "begin lesson" button takes users to a brief lesson introduction presented by an animated agent. Figure two contains a representative screen shot from this animation during which the agent welcomes users to the lesson and advises them on how to proceed through the content. The agent animation, though brief, serves as a means to gain attention and to provide a visual reference for the narrator's voice (Clark & Mayer, 2003). Users who have already seen the welcome sequence can bypass it by clicking "skip intro," or by using a different URL that goes directly to the menu. Figure three depicts the main menu of module topics that appears as the agent's presentation concludes. Users are able to begin the

lesson content by selecting a menu item. The entire script as narrated by the agent can be found

in Appendix C.

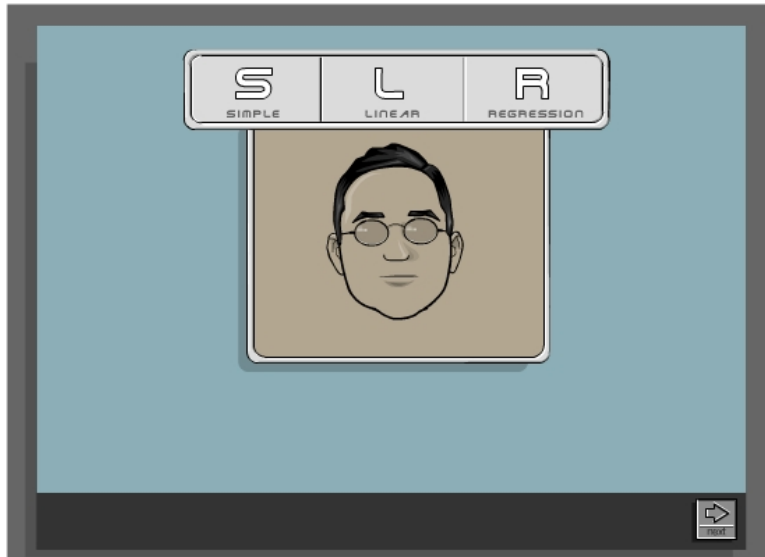Figure 2. Screen shot of the animated agent during the welcome and introduction to the lesson.
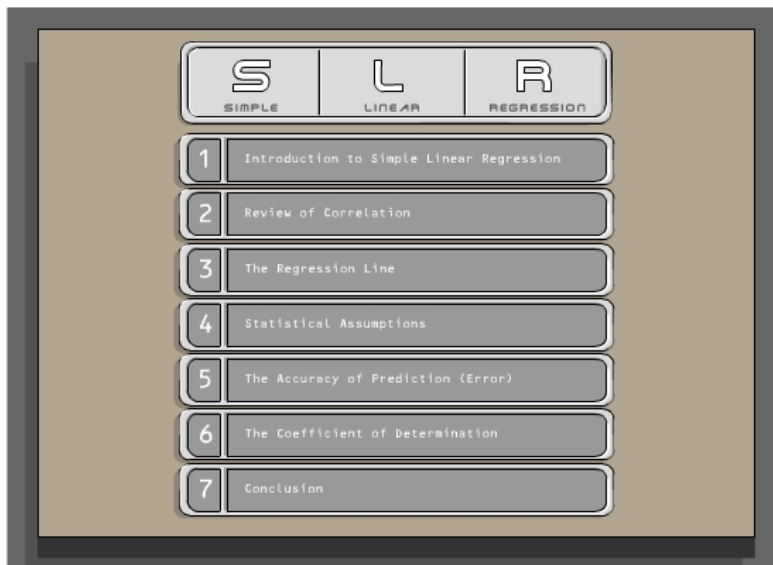
Figure 3. Screen shot of the SLRL main menu.

Module one: introduction to simple linear regression. Each module begins with a

statement of its objectives. These module objectives reflect the GIOs and SLOs (contained in

Appendix A), but are written in broad terms. This was done primarily to make the objectives

quickly accessible to statistics students. The objective for module one is to help learners "understand the basic principles and purposes of simple linear regression."

The next two sections of module one contain the presentation of the instructional content (the purposes and functions of simple linear regression) through a practical example: predicting students' graduate grade point averages (GPAs) using information about their Graduate Records Exam (GRE) scores. The explanation of the example is accompanied by a series of static and animated scatter plots.

Figure four is a screen shot of the data-driven animation activity that users complete after the presentation of the content. Given a specific GRE score ("2000" in figure four), users are asked to predict the graduate GPA by clicking in the appropriate location on the Y-axis. The animation provides simple positive and negative feedback based on their choice.
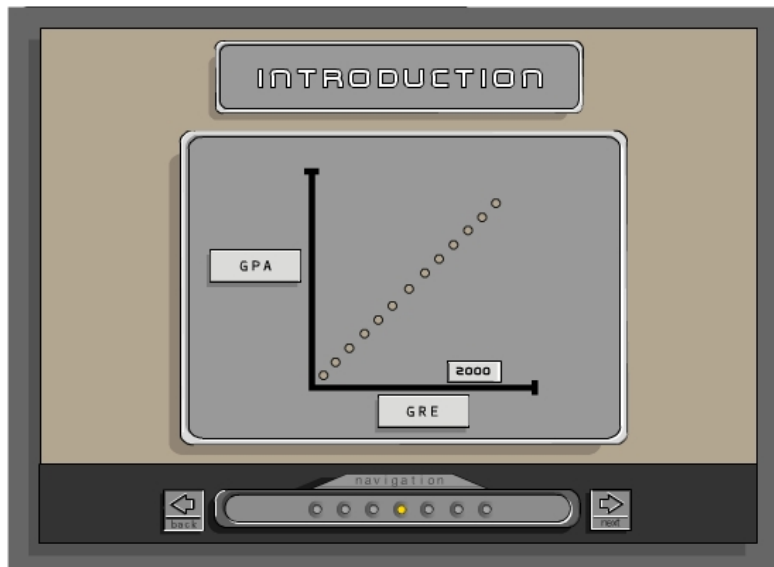
Figure 4. Screen shot of the "GPA prediction" data-driven animation in module one.

Next, students are asked a question about the module's content. This question is meant to help users consider what they have learned and identify areas that they may need to review (Gagne et al., 1992; Clark et al., 2003). The question in module one is multiple-choice and asks

what kind of research situation would require the use of simple linear regression. This specific question allows users to demonstrate their knowledge of the purposes of simple linear regression, as well as begin to understand what types of research situations require regression analysis. Users receive feedback on their answers that explains why their answer was right or wrong.

The final screen summarizes the module's content with animated bullets synchronized with the agent's narration. The conclusion and summary sections were designed to increase users' retention and recall of the material (Gagne et al., 1992)

Module two: review of correlation. The primary purpose of module two is to "stimulate the recall of prerequisite information" (Gagne et al., 1992, p. 167). Prior to beginning the SLRL, users should be familiar with the material covered in this module. On the opening screen, users are informed of the module's objective: to prompt users to recall what they have already learned about the correlation between two variables and the correlation coefficient.

The next three sections of module two contain explanations of the purpose of correlation and the nature of the correlation coefficient (including magnitude and direction). Examples of correlation studies are also included to make the module content clear and applicable. Animated information graphs, scatter plots, scales, and bulleted lists accompany these explanations.

The data-driven animation in module two is a correlation slider activity and is depicted in figure five. In this animation, users are instructed to "click and drag" the slider (the red knob in figure five) across the range of correlation coefficient values. As users drag the slider, the data points in the scatter plot animate to create a representative plot based on the indicated correlation coefficient. The data-driven animation is in real time so that users can watch the data points move as they modify the correlation coefficient. Users can "play" with the correlation slider as

much as they desire. Once they decide to continue, the agent reviews the concepts taught by the data-driven animation and informs them of important details that they should have noticed.
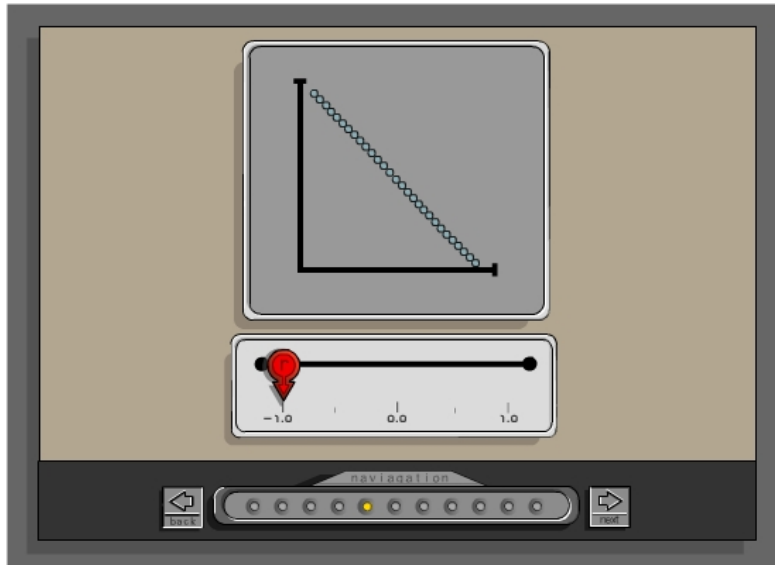
Figure 5. Screen shot of the correlation slider data-driven animation in module two.

Figure six depicts module two's review questions, in which users are asked to estimate the correlation coefficients of a few example data sets. After selecting a correlation coefficient from three options, users are given explanatory feedback (positive or negative) based on their selection. The final screen summarizes the content of module two with an animated, bulleted list synchronized with the agent's narration. The summary includes an explanation of the conceptual and practical differences between correlation and simple linear regression.

Module three: the regression line. Module three begins with the presentation of the following objectives: (a) understand the relationship between simple linear regression and correlation, (b) understand the principle of the best-fitting line, and (c) understand how to make predictions by calculating the regression line.
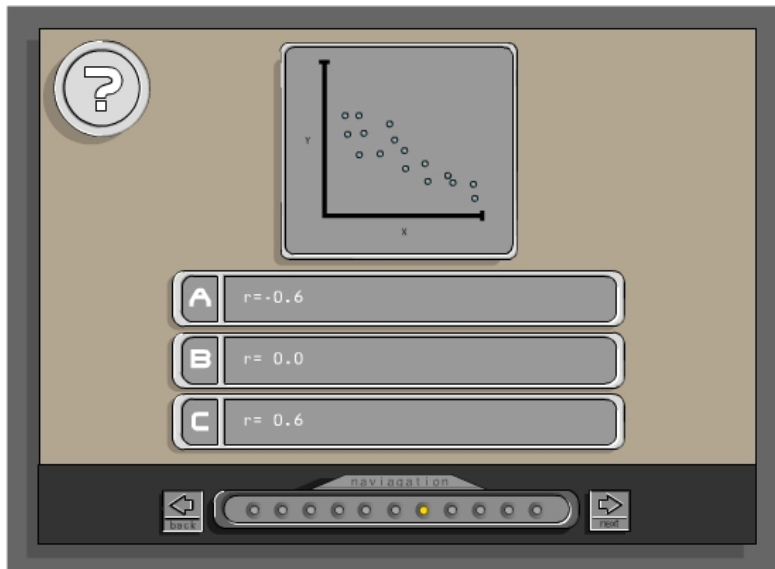
<u>Figure 6</u>. Screen shot example of the review questions at the end of module two.

Because module three contains more material than any other module, the content presentation is divided into five sections. Content topics include the realistic spread of data points, the best-fitting line, the regression line as a running mean, the components of the regression equation, and calculating the regression line. Each of these topics is explained by the agent and demonstrated by graphics and animations. Many of the animations in this module are similar to those in previously described modules; they consist of scatter plots with animated data points to illustrate particular points. Module three also contains animated equation explanations and worked examples (Clark & Mayer, 2003). These extra animations were designed to draw users' attention to important parts of the equation and to give them step-by-step demonstrations of how to perform specific tasks with the regression equation (Clark & Mayer, 2003).

The data-driven animation in module three is meant to help users understand how the regression functions as the "best-fitting line" among a set of data points. The animation, a screen shot of which is displayed in figure seven, presents users with a data set that forms a positive, roughly linear relationship. Users are informed that the red line on the data set represents an

incorrect regression line, and are then instructed to correct the regression line by changing its

slope and intercept. They are told that they will know when they have succeeded in creating the

best-fitting line because the amount of error (indicated in the "Average Error" box in figure

seven) will be at a minimum. Once users have completed the activity, the agent reviews the

major concepts that users should have recognized. The final screen summarizes the module's

content with animated bullets synchronized with the agent's narration in order to increase

retention and transfer.
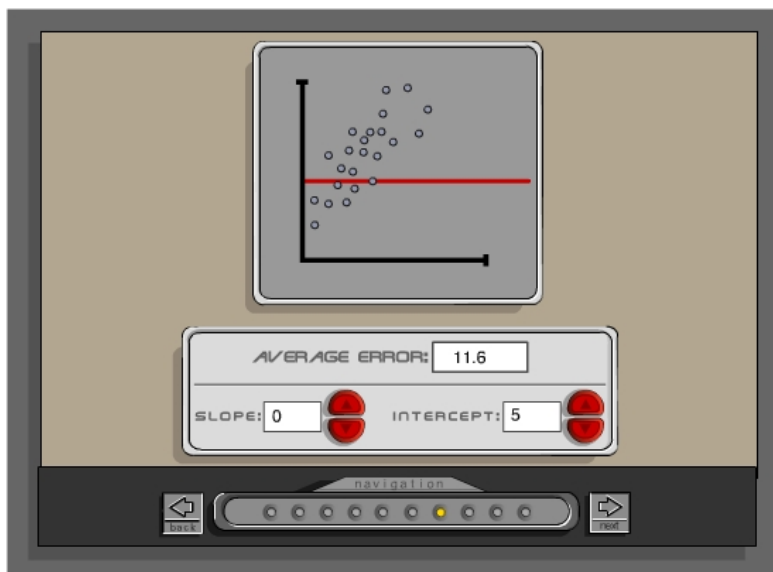
Figure 7. Screen shot of the best-fitting line data-driven animation in module four.

Module four: the assumptions of simple linear regression. This module introduces certain

statistical assumptions that must be taken into account when one is considering and calculating

the regression line. As users begin the module, they are introduced to the following objectives:

(a) understand the basic assumptions of simple linear regression, and (b) understand the

consequences of violating these assumptions.

The content of the module is presented in the next three sections, each of which present

one of the three basic statistical assumptions that must be considered when working with simple

linear regression: (a) the assumption of linearity, (b) the assumption of normal distribution, and (c) the assumption of homoscedasticity. The graphics and animations in this module are slightly different than the graphics that have been described previously. They are two-dimensional but were created to simulate three-dimensional environments. Figure eight contains a screen shot of one of these graphics. Two of the three dimensions are the standard axis of the scatter plot; the third "dimension" represents the distribution of one variable's scores for the sub-scores of the other variable (the yellow distribution curves in figure eight). These simulated three-dimensional graphics are similar to graphics used in many statistics texts (Howell, 1990).
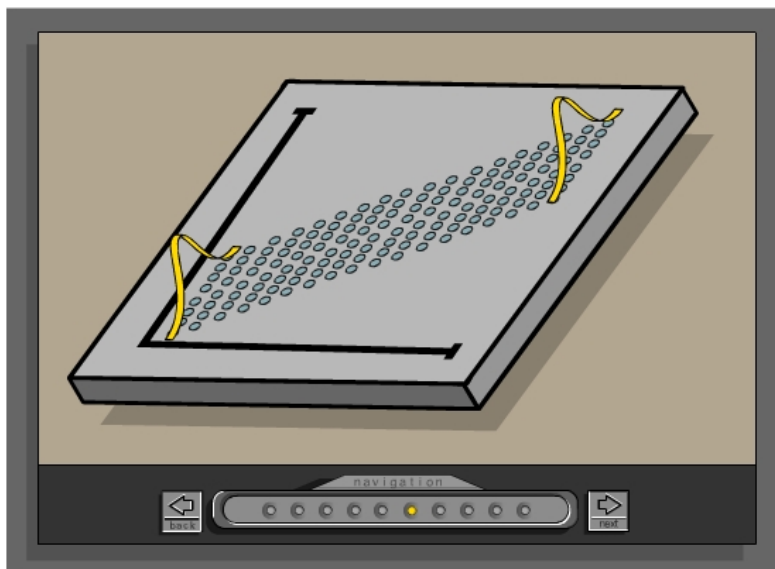
Figure 8. Screen shot of a three-dimensional graphic used to teach the statistical assumptions of simple linear regression.

There are no data-driven animations in this module of the SLRL. Because of the abstract nature of the content, the SLRL designer was unable to conceive an interaction that would communicate the principles of the statistical assumptions in a concise and instructionally powerful manner. The fixed-path animations described in the previous paragraph are intended to convey all of the essential points of these assumptions.

The questions in module four are similar to the questions in other modules. They provide a sample data set and ask users whether or not it would be appropriate to apply simple linear regression statistics to the data set. Users are given explanatory feedback (positive or negative) based on their responses.

The final screen summarizes the module's content with an animated, bulleted list synchronized with the agent's narration (similar to the conclusion screens in others modules). The conclusion includes an explanation of the consequences of violating the simple linear regression assumptions.

Module five: prediction error and the standard error of the estimate. Module five begins with the standard presentation of objective topics. In this module, users should learn about (a) the basic concept of residuals (error), (b) the purpose for calculating the standard error of the estimate, and (c) appropriate interpretation of the standard error of the estimate.

Three data-driven animations precede the actual instructional content of module five. These animations are similar to the data-driven animation in module one. Figure nine shows the first of these animations, which is a sample scatter plot with data points representing a perfect correlation. Users are then given a value for the "x" variable and are asked to predict a score for the "y" variable. Figure ten depicts the question screen that appears once they have made a prediction. On this screen, users are asked how certain they are that their prediction was accurate. The scatter plots on the next two screens depict moderate and zero correlations respectively. Each time users make a prediction, they are asked to indicate their level of certainty concerning the prediction they made. These animations were designed to help users understand (through visualization) the concept that as the correlation coefficient approaches zero, prediction error increases.
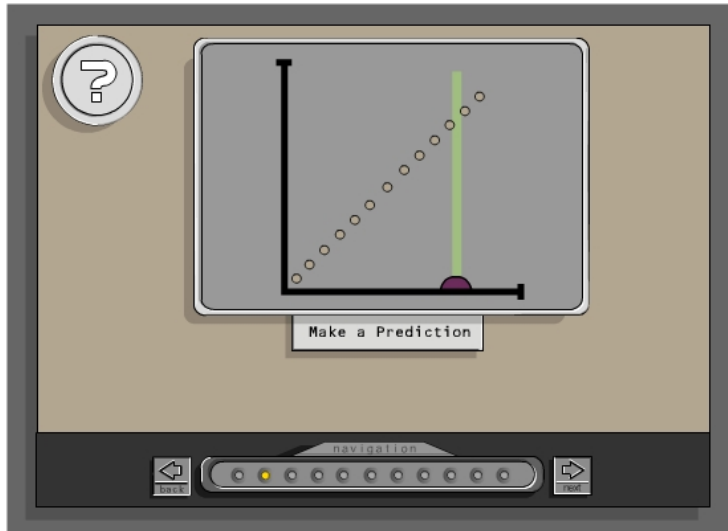
Figure 9. Screen shot example of the "make a prediction" animations in module five.
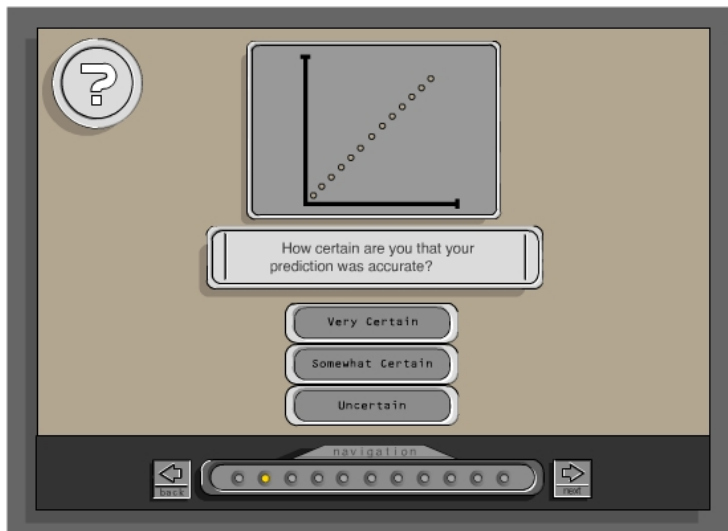
Figure 10. Screen shot of the question users answer after making a prediction.

The next three screens of module five contain narration and animations outlining the major concepts and principles associated with statistical error. These sections also contain an animated worked example that demonstrates the process of calculating the standard error of the estimate.

Figure 11 contains a screen shot of the data-driven animation in module five. In this animation, users see a sample scatter plot with a red regression line. They are instructed to "click

and drag" the slider (the red knob in figure 11) in order to modify the value of the correlation coefficient. As users drag the slider, they see two things change: (a) the amount of error (indicated in the box in figure 11 labeled "standard error of the estimate"), and (b) the spread of the data and the regression line. As the correlation coefficient approaches zero, the data points spread farther from the regression line; a blue line represents the distance between the regression line and each data point. The animation is in real time so that users can see the changing results as the slider is being dragged.

Module five concludes with a summary identical to the summaries the other modules: an animated, bulleted list synchronized with the agent's audio narration.
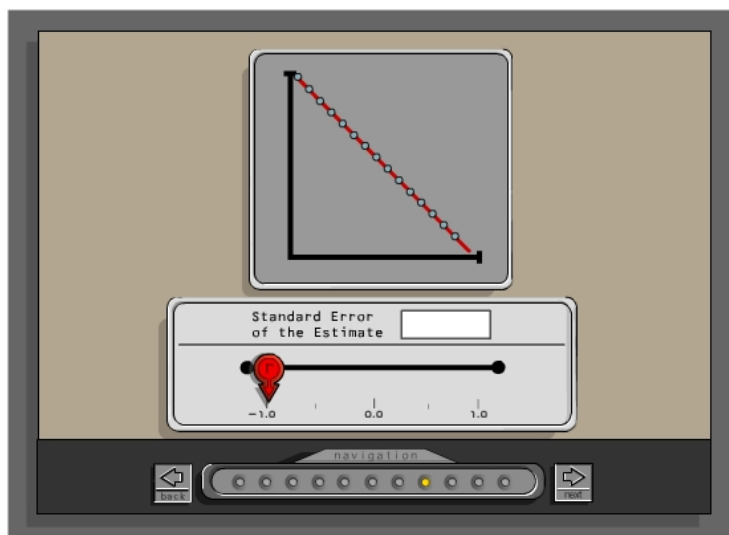
Figure 11. Screen shot of the standard error of the estimate slider in module five.

Lesson conclusion. The final lesson conclusion is similar to the summaries of each module; it simply summarizes the major points and concepts of the entire SLRL. The key concepts are presented in bulleted-list format as the agent reviews the pertinent material. The purpose of this section is to bring closure to the SLRL and to help increase the possibility of users recalling and transferring the learned information (Gagne et al., 1992). Users also encouraged to review the entire lesson or selected modules within the lesson.

Evaluation

According to Gagne et al. (1992), evaluating the effectiveness of instruction is one of the most important phases of any instructional design project. Carefully planned evaluative methods can be the means of improving an instructional product even before it reaches the latter stages of development. For this reason, the process of evaluating the SLRL began, prior to its design and development phases, with the identification of stakeholders and the selection of evaluative criteria.

As mentioned previously, four primary groups of stakeholders for the SLRL were identified at the beginning of the project: (a) professors of statistics, (b) IPT 550 students, (c) instructional designers and developers working on the project, and (d) professors and employers who expect IPT 550 students to know and apply the content from the course.

Each group of stakeholders was interviewed or surveyed so that the designer could develop a list of evaluation criteria and standards with which to judge the SLRL. Their responses to the interviews and surveys were analyzed and compiled into an initial list of questions, criteria, and standards. Each group reviewed the list and made minor revisions and suggestions. There were no contradictions or problems between stakeholder groups in regard to the questions, criteria, and standards. The following is a list of the primary evaluation questions agreed upon by stakeholders (a complete list of the evaluation questions, criteria, and standards can be found in Appendix D): (a) did the lesson stand-alone? (b) did the lesson function properly? (c) could users perform the necessary lesson tasks with the interface? (d) did the students enjoy the lesson? (e) did the graphics, animations, and interactive animations help students learn? (f) were the learners able to meet the specific objectives of the lesson? and (g) did the lesson demonstrate principles of sound instructional design?

In the context of the questions, criteria, and standards, the SLRL was subjected to three evaluative methods: (a) "connoisseur-based" formative evaluation (Flagg, 1990, p. 132), (b) "one-to-one" testing (Gagne et al., 1992, p. 336), and (c) small-group evaluation (Gagne et al., 1992). The specific processes and results of these methods are described below.

Connoisseur-Based Formative Evaluation

The purpose of a connoisseur-based, or expert, formative evaluation is to examine and appraise the instruction itself, not its effects (Flagg, 1990). Flagg (1990) writes, "the most common connoisseur-based studies in formative evaluation involve experts in the subject matter, media, design, and utilization" (p. 132). The evaluative activities implemented in this evaluation stage were grouped into three categories: (a) content evaluation activities, (b) graphical and aesthetic evaluation activities, and (c) instructional design evaluation activities

Content Formative Evaluation Activities

One of the most important evaluative efforts in the design and development of the SLRL was the connoisseur-based evaluation of the actual statistical content. The formative evaluation activities consisted of a series of interviews with the SME and project sponsor Dr. Stephen Yanchar. Interviews were conducted frequently (one or two times per week) over the course of the design and development of the project (eight weeks) in order to ensure that materials accurately conveyed the principles of simple linear regression. The interview questions were designed to elicit formative feedback on the following evaluation questions (listed previously): (a) did the graphics, animations, and interactive animations help students learn more than simple text or textbook activities? and (b) were the learners able to meet the specific objectives of the lesson?

There were five specific areas in which the SME proved invaluable in providing formative feedback and guiding the design and development of the project: (a) instructional objectives (GIOs and SLOs); (b) lesson outline and topic sequence; (c) graphics, animations, and data-driven animations; (d) the SLRL script; and (e) the simple linear regression test.

The frequency of interviews allowed the designer and the SME to go through several iterations of review and revision before finalizing specific materials. No formal evaluation instruments were required to conduct these expert reviews; their products were the comments, suggestions, feedback, and notes that resulted from the discussions during the interviews. These products were immediately reviewed and implemented in the SLRL objectives, graphics, animations, data-driven animations, script, and test (final versions of the objectives, test, and script can be found in Appendices A-C). The specific data collected during these interviews is too abundant to be listed here, but the process used to revise the SLRL script will illustrate the nature of the suggestions made during this portion of the connoisseur-based formative evaluation.

The SLRL designer wrote the initial draft of the script based on the instructional objectives and the outline created during the task analysis. The SME reviewed this draft and made suggestions for improvement. The suggestions focused mainly on specific wording and proper use of terminology. The SME also made important additions to the script, including examples of real world applications and worked examples. The designer modified the script based on the feedback, and the SME reviewed the script again. This process continued until the script was in satisfactory condition. The SME did not make recommendations concerning the instructional strategies and methods used in the lesson because he considered his sole responsibility to be the perfection of the instructional content.

Graphical and Aesthetic Formative Evaluation Activities

Most instructional design texts and models have little or nothing to say concerning the aesthetic appeal of instructional products (Misanchuk et al., 2000). Misanchuk et al. even state that the field of instructional design suffers from a "chronically low level of media design quality in products generated by practitioners in the field" (2000, Introduction, Why this book, p. 3). The SLRL instructional designer's consultations with a professional illustrator and a professional graphic designer were the formative evaluation activities that assessed the aesthetic quality of the lesson materials. These evaluations were conducted as interviews in which the illustrator and graphic designer gave feedback in regard to the (a) color palette, (b) screen layout, and (c) general look and feel of the lesson.

These interviews occurred only as needed (four times) during the design and development process. The products of these evaluations were the notes taken during the interviews. The notes were analyzed in order to ensure that the aesthetic feedback from visual experts did not conflict with principles of sound instructional design or interfere with the presentation of lesson content (Misanchuk et al., 2000). As soon as feedback was analyzed within this context, it was implemented immediately. For example, during one of the interviews, the graphic designer observed that the color palette was too dull. She suggested that a bright color be added to the palette so that the colors were more attractive and engaging. Based on this feedback, the instructional designer selected a bright shade of red and used it in the graphics and animations. This simple addition increased the aesthetic appeal and added instructional value by drawing learners' attention to important parts of the graphics and animations. Other specific feedback from the graphic designer and illustrator resulted in slight modifications to the look and feel of buttons, the placement and spacing of elements in the navigation tools, and the fonts used

throughout the lesson. A screen shot that exemplifies the color palette, screen layout, and general look and feel of the SLRL can be found in Appendix E.

Instructional Design Formative Evaluation Activities

Finally, the SLRL was formatively evaluated from an instructional design perspective. These evaluative activities focused on the instructional events and strategies employed in the SLRL. At two different stages of the design and development process, two instructional designers evaluated the lesson. The first of these reviews occurred after a low-fidelity version of the lesson was produced; the second came after the high-fidelity version of the lesson had been completed. In each case, the instructional designers reviewed the material as if they were learners. Designers were asked to use "think-aloud" protocol concerning instructional design issues (Rosson & Carroll, 2002, p. 243). The SLRL designer recorded the thoughts, opinions, and feedback of the reviewers as they spoke. These reviews focused on the evaluation question, "does the lesson demonstrate principles of sound instructional design?"

No formal evaluation instruments were used during these evaluations of the SLRL; their products were the comments, suggestions, feedback, and notes that resulted from the designer's reviews of the product prototypes. One important example of feedback from these design reviews occurred early in the development process. As the instructional designers reviewed the initial prototypes of the SLRL, they immediately recognized that the modules were too long. They suggested that the modules be divided into small sections through which users could actively navigate. This would allow users to study at their own pace and be more actively involved in the instruction. This suggestion resulted in the division of module content and subsequent modification to the navigation tools (as described in the "Description of the Product" section). In subsequent design reviews, instructional designers made suggestions about the

textual display of instructions, the review questions in each module, and the data-driven

animations. Notes from these design reviews were reviewed and analyzed based on their

feasibility and practicality within the project's timeline; actual changes were implemented in the

SLRL if they were possible and reasonable. For example, one instructional designer suggested

that the questions in each module be based on a series of sophisticated branching conditions so

that as users' understanding increased, they would encounter more difficult questions. The SLRL

designer recognized the value of such a strategy but was unable to implement it in the lesson due

to insufficient time and resources. Details regarding the specific instructional design of the lesson

can be found in the "Description of the Product" section.

One-to-One/Usability Formative Evaluation Activities

According to Gagne et al., "one-to-one" evaluative methods are observations of a

learner's performance when he or she is presented with a specific learning task (1992, p. 336).

This type of formative evaluation provides valuable feedback during the development phase of

the instructional design process (Gagne et al., 1992). Rubin (1994) suggests that designers

subject their product to a series of "usability tests" in order to "gather information from the target

users to improve the design of a[n instructional] product" (p. 16). The methodologies used in this

phase of the formative evaluation were adapted from the work of Rosson and Carroll (2002),

who define usability as "the quality of a system with respect to ease of learning, ease of use, and

user satisfaction" (p. 9). Rosson and Carroll advocate a user-centered, scenario-based approach

to usability. In this method, target users see prototypical models of the software product. The

evaluator then gives users a typical user task to complete. As users attempt to complete the task,

the evaluator makes observational notes about what the user does. Users are also asked to utilize

"think-aloud" protocol so that the evaluator can better understand their motives, intentions, and expectations with regards to the instructional materials (Rosson & Carroll, 2002, p. 243).

The one-to-one evaluations of the SLRL occurred during three phases of the product's development: (a) low-fidelity phase, (b) mid-fidelity phase, and (c) high-fidelity phase. During each of these phases, at least two members of the actual target audience (graduate students in the BYU IPT department) participated according to the Rosson and Carroll method described above. A summary of each phase of usability testing is described below.

Low-Fidelity Prototype Phase

The first of the one-to-one evaluations was conducted with the use of a low-fidelity prototype that included rough representations of the interface, instructions, content, and sequence of the SLRL (Rosson & Carroll, 2002). It consisted of simple yet demonstrative pencil drawings on paper that approximated the final product as closely as possible. Three members of the target audience reviewed the low-fidelity prototype in a face-to-face interview setting using "think-aloud" protocol. Because of the nature of the low-fidelity prototype, it was not possible to address every evaluation question at this stage. Rather, this prototype focused on the following evaluation questions: (a) did the lesson stand-alone? (b) could users perform the necessary lesson tasks with the interface? and (c) did the students enjoy the lesson?

Extensive observational notes were collected during the low-fidelity prototype phase of development. Notes from these interviews were reviewed, analyzed, and used to inform and guide the development of the mid-fidelity prototype described below. There were several significant consistencies in user feedback that resulted in modifications to the user interface and the display of information. These revisions included (a) the arrangement and functionality of the navigation bar, (b) the interface and user tasks of the data-driven animations, (c) the functionality

of the main menu, and (d) the functionality of the lesson's agent. In two cases the opinions and suggestions of users were contradictory. The designer resolved the inconsistency with a decision based on intuition and experience and in both cases the problems did not recur in subsequent rounds of testing.

During the low-fidelity prototype phase, much was learned about the user interface. In many cases in this early stage, participants were required to imagine the final state and appearance of the lesson. Even so, during each interview participants openly expressed their sense of engagement with and enjoyment of the lesson content. In addition, users were able to describe expected functionality that went far beyond the functionality of the paper and pencil prototype. These ideas were usually preceded with the phrase, "It would be cool if . . . ." This not only provided the designer with an additional list of possible functionality, but also served as evidence that the instructional materials engaged the target users.

Mid-Fidelity Prototype Phase

The second phase of one-to-one learner evaluation utilized the mid-fidelity prototype of the SLRL. This prototype was a more precise representation of the SLRL content and interface. It included paper printouts of computer-generated graphics and interfaces that approximated the final product without interactivity and animation. The mid-fidelity prototype evaluation was conducted in identical fashion to the low-fidelity prototype: three members of the target audience reviewed the prototype and gave evaluative feedback that included "think-aloud" protocol (participants were not the same as those selected previously). Data collected during this activity focused primarily on the following evaluation questions: (a) did the lesson stand-alone? (b) Could users perform the necessary lesson tasks using the interface? and (c) did the students enjoy the lesson?

As was the case with the low-fidelity prototype evaluation activities, many field notes were collected during the evaluation of the mid-fidelity prototype. Data from the face-to-face interviews was reviewed, analyzed, and used to guide and inform the development of the high-fidelity version of the lesson. Three major revisions made to the SLRL fell into these categories: (a) the user interface (including the navigation bar), (b) the interactive data-driven animations, and (c) the functionality of the main menu. For example, during the mid-fidelity prototype evaluation, each user suggested that the circles in the navigation tool function as buttons. This functionality was noted and later implemented in the high-fidelity prototype and the final SLRL. Also noted during this phase was that users did not take the time to read instructional text on the screen. Based on this observation, instructions for specific user tasks were added to the audio script and read by the agent.

Learners also gave suggestions about specific data-driven animations. The most notable improvements were to the animations in modules one and five, where learners make predictions about example data sets (described in the "Description of the Product" section). In this case, users described exactly what they expected to see as they interacted with the animation and moved the cursor over certain areas of the screen. The designer was able to implement the users' descriptions exactly as they were specified.

A final example of user feedback obtained during the mid-fidelity prototype phase dealt with the main menu. When users returned to the main menu of lesson topics they were often confused and could not recall which modules they had already completed. Users suggested a sophisticated improvement that would record the modules they had visited with a check mark. This particular suggestion was not implemented due to time constraints. Instead, the designer simply placed a large number to the left of the module title. This solution, though not ideal,

solved the problem in subsequent versions. In addition to providing detailed feedback about lesson revisions and improvements, participants again reported that the content was enjoyable and accessible.

High-Fidelity Prototype Phase

The third phase of one-to-one learner testing was the evaluation of a high-fidelity prototype of the lesson prior to its small group implementation. This version of the lesson consisted of a fully functional, animated, and interactive module of the SLRL. The specific module evaluated during the high-fidelity phase was module two: the review of correlation. Two previously unselected members of the target audience participated in the evaluation. Using "think-aloud" protocol, each user reviewed the high-fidelity version in a face-to-face interview setting. The high-fidelity prototype evaluation addressed the following evaluation questions: (a) did the lesson stand-alone? (b) did the lesson function properly? (c) could users perform the necessary lesson tasks with the interface? and (d) did the students enjoy the lesson?

The data from each of these interviews was collected, reviewed, and used to refine and guide the final development stages of the entire SLRL. Users' feedback focused mostly on the issue of technical functionality and revealed several programming bugs (malfunctioning buttons, missing audio, etc.) and technical errors (typographical mistakes, editing errors, etc.). Despite these bugs and errors, users enjoyed participating in this prototype and had especially positive reactions to the data-driven animations and other interactive activities.

There were relatively few comments that addressed the interface. Users did suggest the addition of some type of visual cue to inform them that the narration for each section was complete. Based on this feedback the designer programmed the "next" button to flash in yellow when the narration was complete. The lack of more significant feedback regarding the user

interface was most likely because many of the previous problems had been corrected during the iterative evaluation process.

In summary, the process of developing and evaluating prototypes in progressively higher fidelity served an important purpose in the development of the SLRL. Major problems and issues with the interface and the usability were identified and corrected early in the development process. These corrections resulted in a product that was easier to use, and thus, more satisfying to the end-users (Rosson & Carroll, 2002).

Small-Group Implementation Evaluation

The small-group implementation evaluation of the SLRL served as an evaluation of the materials. Gagne et al. (1992) suggest that "summative evaluation is usually undertaken when development of an instructional entity is in some sense completed, rather than ongoing. Its purpose is to permit conclusions about how well the instruction has worked" (p. 293). The small-group evaluation of the SLRL occurred in the IPT 550 course during the fall semester of 2003 as students reached the simple linear regression unit of the course. The product implemented for evaluation was the entire SLRL (five modules with introduction and conclusion).

Description of Subjects

Fifteen members of the target audience participated in the small-group evaluation of the SLRL. Each participant completed a questionnaire (described in the "Instruments" section) that provided basic information about the group. All 15 participants were students accepted into masters or doctoral programs in the IPT department and enrolled in the IPT 550 Empirical Inquiry and Statistics course. Fourteen of the participants took the course for credit and one student audited the course. Eleven participants had taken statistics courses as undergraduates; four of the participants had no prior experience with statistics. Two participants reported that

learning statistics is a "very challenging" learning task, three participants deemed it "very easy," the other 10 considered statistics a "moderately challenging" subject of study. All 15 participants reported that they were very familiar with WWW browsers and computers in general.

Instruments

Evaluative instruments provide information about users' perceptions of the instructional product and measures of their performance (Gagne et al., 1992). They also serve as a means of collecting feedback about the product from end-users. Two instruments were developed for the purpose of evaluating the SLRL: a performance assessment and a post-lesson questionnaire.

The performance assessment was developed parallel to, but separate from, the instructional materials. Its purpose was to provide answers to the following evaluation questions: (a) did the graphics, animations, and interactive animations help students learn? and (b) were the learners able to meet the specific objectives of the lesson? The designer consulted with assessment expert and SME Dr. Richard Sudweeks regarding the alignment of the performance assessment with the instructional objectives. The assessment consisted of 12 questions and the test was scored out of a possible 24 points. Four of the questions were dual response items that focused on proper application of simple linear regression in research situations. Another four questions were multiple-choice items that focused on the relationship between correlation and regression. The final four questions were short essay items that focused on the following topics: (a) the nature of the regression line, (b) the assumptions of simple linear regression, (c) the accuracy of prediction, and (d) proper interpretation of the standard error of the estimate. A copy of the performance assessment can be found in Appendix B. The performance assessment served as the pretest and the posttest for the small-group implementation evaluation.

The post-lesson questionnaire was also developed parallel to the instructional materials. The development process included a pilot and revision of the questionnaire during the high-fidelity prototype of the SLRL (described previously). The purpose of the questionnaire was to provide answers to the following stakeholder questions: (a) was the lesson able to stand-alone? (b) did the lesson function properly? (c) could users perform the necessary lesson tasks with the interface? and (d) did the students enjoy the lesson? The questionnaire consisted of 16 questions intended to establish each participant's perception of statistics, background in the subject, and overall user satisfaction with the instructional materials. (Students also responded to questions about their perceptions of the SLRL; for example, most desirable and least desirable aspects of the product). A copy of the post-lesson questionnaire can be found in Appendix F.

Procedures

The small-group implementation took place during a two-hour block of the regular IPT 550 meeting time. Students were informed that participation was not compulsory and would not affect their grades in the course. All 15 students voluntarily remained during the entire two-hour block. Prior to accessing the SLRL on the WWW, participants completed the performance assessment. This initial assessment served as pretest of students' knowledge of simple linear regression. After completing the pretest, each student was provided with a laptop and headphones in order to facilitate access to the SLRL. Students remained in the same classroom, but they did not collaborate at any time. While students worked through the SLRL, the designer made observational notes. The designer did not provide help or instruction while students completed the lesson. After finishing the SLRL, participants again completed the content assessment (posttest). They were then given the post-lesson questionnaire, during which they were permitted to access the SLRL as they answered specific questions about the instructional

materials. Participants completed the pretest, lesson, posttest, and questionnaire in an average of one hour and 45 minutes. Within one week of the small-group implementation, the designer conducted follow-up interviews with three participants. The purpose of these interviews was to further explore the participants' experiences during the implementation of the SLRL. The results from the evaluation are summarized below.

Qualitative Results and Discussion

Qualitative data was collected from post-lesson questionnaires, follow-up interviews, and observational notes taken during the small-group implementation. Interview and questionnaire items were specifically designed to provide answers to the following evaluation questions: (a) did the lesson stand-alone? (b) did the lesson function properly? (c) could users perform the necessary lesson tasks using the interface? and (d) did the participants enjoy the lesson? Below is a summary of participants' interview and questionnaire responses categorized according to the appropriate evaluation question. A summary table of questionnaire and interview responses can be found in Appendix G.

Did the lesson stand-alone? Criteria and standards associated with this evaluation question (contained in Appendix D) indicated that the instructional materials should be "stand-alone." In other words, stakeholders desired to have instructional materials that required no additional support from subject matter or technical experts. Observations made during the small-group implementation indicated that students did not require additional support as they completed the lesson. The lesson's ability to stand-alone is attributable to the number of prototype and script revisions in the project's design and development. Although the lesson allowed students to study without help or interaction, several students indicated that they would

have liked to have someone available to answer content questions or with whom to converse about the concepts.

Did the lesson function properly? Stakeholders required that the lesson be free from technical bugs and errors. Participants in the small-group implementation indicated that the lesson had three significant categories of technical errors: (a) audio errors, (b) spelling mistakes, and (c) malfunctioning buttons.

The audio errors occurred because of mistakes in audio editing that resulted in several phrases being repeated during the narration. Audio errors and spelling mistakes were most likely the results of insufficient time for quality control before the implementation. Students reported that these errors were bothersome but that they did not inhibit their learning in any way. Students also reported a malfunctioning button in module five. This bug resulted in a loop that sent users to a previous screen instead of advancing them to the module conclusion. Participants worked around the problem by returning the main menu. Because the bug occurred at the end of the module, students did not miss any of the instructional content. Students reported that this bug, while inconvenient, did not affect their learning.

Stakeholders also wanted the lesson to load within a reasonable amount of time (less than 3 seconds). Precise load times were not measured, but observations at the implementation suggested that the SLRL did not suffer from excessive load times.

Could users perform the necessary tasks with the interface? The SLRL was implemented as an entirely stand-alone unit of instruction; implementation participants were given the opportunity to provide feedback about their ability to perform necessary tasks using the user interface. This evaluation question referred primarily to the navigation controls and the data-driven animations. A majority of the participants responded that they would have liked to be able

to access the main menu from any point in the lesson. Participants recommended that a "menu" button be placed on the navigation bar to accomplish this. One participant reported that he was confused by the interface for the data-driven animation in module five; no other participants reported similar confusion. The lack of significant problems with the user interface is attributed to the rounds of prototyping and revising during the development process.

Did participants enjoy the lesson? Stakeholders specified that the SLRL should be designed in such a way that it is engaging and enjoyable to students. Every effort was made in design and development to accomplish this, and for the most part students reported that the lesson was enjoyable. When asked to indicate the most enjoyable part of the lesson, participants responded overwhelmingly in favor of the data-driven animations. They also reported that the data-driven animations engaged them in the lesson content and motivated them to continue learning statistics. All of the participants indicated that they would like to use more lessons of a similar nature.

Participants also answered questions about their least favorite aspects and the most discouraging portions of the lesson. Many of the participants reported that the worked examples and statistical equations (located in modules three and five) were the most discouraging parts of the lesson. One student said of these worked examples, "They represent what I hate most about statistics." This trend in the responses indicates that participants enjoyed those parts of the lesson that were the least like traditional textbooks and did not enjoy the parts that resembled traditional textbook methods.

Quantitative Results and Discussion

Quantitative data was collected from the content assessment, which served as the pretest and posttest in the small-group implementation. These data answered the following evaluation question: were the learners able to meet the specific objectives of the lesson?

The mean score for the pretest was 10.8 out of 24 possible points; the standard deviation was 4.72. The mean score for the posttest was 21.4 with a standard deviation of 1.96. The lower variance in posttest scores is most likely due to a ceiling effect. The average gain score from pretest to posttest was 10.6. A one-way analysis of variance revealed that pretest and posttest means were significantly different (α= .05), $F(1,14) = 119.41$, $p = .000$.

As mentioned previously, data regarding participants' previous experiences with statistics were also collected. An analysis of covariance was performed to analyze the difference between pretest and posttest means with this factor statistically controlled. The difference between means was still significant (α= .05), $F(1,12) = 16.82$, $p = .001$. Previous experience, however, was not significant, $F(1,12) = 2.189$, $p = .165$. Based on these data it appears that the instructional design made a difference even when previous experience was statistically controlled. The array of pretest and posttest scores can be found in Appendix H.

In addition to being statistically significant, the data collected seem to suggest *practically* that learning occurred. As mentioned previously, participants took the pretest, completed the SLRL, and took the posttest all within a two-hour period. They did not leave the room where the implementation took place and did not collaborate at anytime. The influences on learning were therefore somewhat controlled. On the other hand, the successive administration and identical nature of the tests may have given participants the advantage of anticipating the posttest questions while completing the SLRL.

In summary, the qualitative data collected during the small-group implementation indicates that participants believed the quality of the instructional materials to be high. In addition, the quantitative improvements of their test scores after participating in the SLRL indicate that learning occurred. Although the collected data do provide positive results concerning the instructional quality of the SLRL, more research needs to be conducted to determine its true value. Future studies could compare the SLRL to stand-alone, web-delivered tutorials *without* graphics, animations, and data-driven animations. Experiments that compare the SLRL to other methods of teaching statistics could also be conducted (e.g., traditional lectures, textbooks, etc.).

Critique

Product Strengths

The strengths of the SLRL originate from the basic assumption of the SME and the designer that graphics, animations, and data-driven animations are useful and effective instructional tools for teaching basic statistical concepts. The product's greatest strengths were the data-driven animations contained within each module; these animations provided students with visual representations of important statistical concepts. The ability to manipulate information in the graphic and then watch the subsequent effects proved to be a powerful learning tool and a method for engaging the students in the content.

The fixed-path animations synchronized with the agent's audio narration were another strength of the SLRL. The audio narration afforded users the opportunity to focus their attention on key animations in the lesson, and the combination of visual and audio stimulation served communicated the content in an instructionally effective and sensory-engaging manner.

Product Weaknesses

A significant weakness of the SLRL is that, in certain aspects, it is not easily modified. Ironically, this weakness is a result of one of the lesson's strengths: the synchronized animations and audio narration. The synchronized sequences were meticulously crafted so that an animation is triggered at the very instant that the agent mentions a key point or concept. Therefore, if a statistics instructor, for example, chose to modify the narration of a particular section, a significant amount of time and resources would be required to make all of the necessary changes. This weakness makes desired characteristics such as continuous improvement very difficult, if not impossible. At this point in time, there appears to be no good solution to this problem. The tradeoff, therefore, seems to be the instructional effectiveness of carefully designed animations and the cost to produce and improve them.

Another weakness of the SLRL was the presence of technical errors throughout the lesson. The most significant of these errors were poorly edited audio files and spelling errors in the lesson's text. Participants in the small-group implementation easily identified these errors, and while they most likely did not significantly affect learning outcomes, the errors were a distraction. The audio errors are a significant problem given the weakness described in the previous paragraph. The spelling errors, on the other hand, are easily corrected. Probable explanations for these problems can be found in the Process Weaknesses section.

Process Strengths

Upon the project's initiation, the designer and the SME agreed upon a set schedule for delivering, reviewing, and revising prototypes of the SLRL and its script. As mentioned in the "One-to-One Evaluation" section, this rigorous schedule included a series of prototypes that were tested by members of the target audience. The improvements made during these rounds of

prototyping and usability testing contributed considerably to the lesson's success during the small-group implementation.

The detailed schedule of specific deadlines and deliverables was another important strength of the design and development process. The deadlines provided the designer and the SME with checkpoints that ultimately allowed the product to be delivered in a timely manner. The successful adherence to the original schedule was due, at least in part, to the manner in which the SME supported the project and made himself available to review prototypes and versions of the script as necessary. With the schedule in place and the SME so willing to provide necessary support, the deadlines set during the design and development processes were met.

Process Weaknesses

The design of the SLRL was the undertaking of a single instructional designer with the support of a SME. The instructional designer was solely responsible for the development of the SLRL. In other words, the designer was responsible for audio recording, audio editing, graphic design, graphics production, type editing, animations production, programming, and quality control. This was primarily due to the lack of supporting resources for the project. The designer was therefore unable to elicit outside help and lacked the time necessary to provide sufficient quality control of the product before the small-group implementation. In fact, the designer was already aware of most of the errors and mistakes identified by participants during the implementation; he simply did not have the time or resources to make the necessary corrections. In an ideal situation, the designer would have been managing and coordinating the efforts of a team of graphic designers, animators, and programmers to create the lesson. In the case of the SLRL, a team effort would have greatly improved the quality of the instructional materials.

Schedule

Table 1 contains a summary of the estimated and actual number of hours required to complete each task, and the projected and actual delivery dates for each phase of the project. Time estimates for each phase of the project were made during the project's design phase. A timeline in the form of a Gantt chart was also created to outline major deadlines. Because the SLRL was implemented during the regular flow of a semester course, there was little flexibility in the schedule; the dates on which specific phases of the project were completed varied only slightly from the projected dates established at the project's inception. However, the actual hours spent on the project far exceeded the estimated hours. This was due to the fact that the designer grossly underestimated the amount of time some phases of the project would require (e.g., animations, audio recordings, and audio editing).

Table 1

Estimated and Actual Schedule for the SLRL

Task	Projected delivery	Actual delivery	Estimated hours	Actual hours
Instructional Objectives	15-Sep-03	15-Sep-03	5	9
Lesson Outline	15-Sep-03	15-Sep-03	5	7
Low-fidelity Prototype	26-Sep-03	24-Sep-03	10	15
Mid-fidelity Prototype	29-Sep-03	30-Sep-03	10	15
High-fidelity Prototype	1-Oct-03	3-Oct-03	20	27
Script	3-Oct-03	6-Oct-03	10	14
Measurement Instruments	17-Oct-03	22-Oct-03	15	23
Complete SLRL	17-Oct-03	22-Oct-03	120	180
Totals			195	290

Budget

The primary costs of the project were materials and the time of personnel. The materials used to develop the project were minimal. Necessary materials included the educational version of Flash MX software, the educational version of Adobe Illustrator, computer printouts, and copies. The equipment that facilitated the implementation included the computers and other software packages already owned by the University. BYU provided the necessary facilities, including power, server space, and Internet access.

The most expensive cost was the labor associated with designing and developing the SLRL; however, the designer received no wages or salary for the time that he dedicated to the project. Personnel costs were calculated as if the designer were receiving the typical wages of a graduate student employed by the IPT department.

Another significant cost of the project was the consulting time of BYU faculty members who acted as subject matter experts and design consultants: Dr. Paul Merrill, Dr. Stephen Yanchar, Dr. Richard Sudweeks, and Dr. Stephanie Allen. The wages and salaries of the personnel involved in the project are confidential. Approximate salaries were determined based on calculations obtained from the BYU Office of Research and Creative Activities (ORCA). The material costs, on the other hand, are accurate.

Table 2 contains an itemized budget of estimated and actual costs in terms of personnel hours, materials costs, and equipment costs. Major discrepancies in the estimated and actual budget are due to the designer's inaccurate estimation of the amount of personnel hours required to complete certain phases of the project.

Table 2

Estimated and Actual Costs for the SLRL

Resource	Rate	Estimated quantity	Actual quantity	Estimated cost	Actual cost
Subject matter experts	$25/hr	40	40	$1,000	$1,000
Design consultants	$25/hr	30	10	$750	$250
SLRL designer/developer	$15/hr	195	290	$2,925	$4,350
Macromedia Flash MX	$100	1	1	$100	$100
Adobe Illustrator	$100	1	1	$100	$100
Computer printouts	$.05 ea.	100	50	$5	$2.50
Photocopies	$.05 ea.	300	500	$15	$25
Totals				$4,895	$5,827.50

References

Aberson, C. L., Romero, V. L., Berger, D. E., Healy, M. R. (2000). An interactive tutorial for teaching statistical power. *Journal of Statistics Education,* 10, 3. Retrieved September 2, 2003, from http://www.amstat.org/publications/jse/v10n3/aberson.html.

Alesandrini, K. L. (1984). Pictures and adult learning. *Instructional Science,* 13, 63-77.

Anglin, G., Towers, R., & Levie, H. (1996). Visual message design and learning: The role of static and dynamic illustrations. In D. H. Jonassen (Ed.). *Handbook of research for educational communications and technology* (pp. 755-794). New York: Simon & Schuster Macmillan.

Clark, C. C., & Mayer, R. E. (2003). *E-learning and the science of instruction.* San Francisco, CA: Jossey-Bass/Pfeiffer.

Dick, W., Carey, L., & Carey, J. O. (1996). *The systematic design of instruction* (5th ed.). Glenview, IL: Scott, Foresman.

Dallal, G.E. (2000). *Is statistics hard?* Retreived September 5, 2003, from http://www.tufts.edu/~gdallal/hard.htm.

Emery, D. (1993). Developing effective instructional graphics. *Journal of Interactive Instruction Development*, 6(2), 20-23.

Flagg, Barbara N. (1990). *Formative evaluation for educational technologies.* Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Friedman, H. H., Friedman, L.W., & Amoo, T. (2002). Using humor in the introductory statistics course. *Journal of Statistics Education*, 10, 3. Retrieved September 2, 2003 from www.amstat.org/publications/jse/v10n3/friedman.html

Furlong, N. E., Lovelace, E. A., & Lovelace, K. L. (2000). *Research methods and statistics: An*

integrated approach. San Diego: Harcourt College Publishers.

Gagne, R.M., Briggs, L.J., & Wager, W.W. (1992). *Principles of instructional design* (4th ed.).
New York: Holt, Rinehart and Winston.

Giesbrecht, N. (1996). *Strategies for developing and delivering effective introductory-level
statistics and methodology courses*, ERIC Document Reproduction Service, No. 393-668,
Alberta, BC.

Gronlund, N.E. (1991) *How to write and use instructional objectives* (4th ed.). New York:
Macmillan.

Hong, E., & O'Neil, H. F. Jr. (1992). Instructional strategies to help learners build relevant metal
models in inferential statistics. *Journal of Educational Psychology*, 84(2), 150-159.

Howell, D. C. (1999). *Fundamentals statistics for the behavioral sciences* (4th ed.). California:
Duxbury Press.

JSE Mission Statement (n.d.). Retreived September 4, 2003, from
http://www.amstat.org/publications/jse/jse_mission.html.

Ip, E. (2001). Visualizing multiple regression. *Journal of Statistics Education*, 9, 1. Retrieved
September 2, 2003, from http://www.amstat.org/publications/jse/v9n1/ip.html.

Kennedy, P.E. (2002). More on venn diagrams for regression. *Journal of Statistics Education*,
10, 1. Retrieved September 2, 2003, from
http://www.amstat.org/publications/jse/v10n1/kennedy.html.

Levie, W. H. (1987). Research on pictures: A guide to the literature. In D. M. Willows and H. A.
Houghton (Eds.), *The psychology of illustration*, volume 1(pp. 1-50). New York:
Springer-Verlag.

Mathieson, K., Doane, D., & Tracy, R. L. (1995). A program for visualizing comparisons between two normal distributions. *Journal of Statistics Education*, Volume 3, number 1. Retrieved September 2, 2003, from http://www.amstat.org/publications/jse/v3n1/mathieson.html

Mayer, R. E. (1989). Models for understanding. *Review of Educational Research*, 59, 43-64.

Merrill, P. F., & Bunderson, C. V. (1981). Preliminary guidelines for employing graphics in instruction. *Journal of Instructional Development*, 4(4), 2-9.

Misanchuk, E., Schweir, R., & Boling, E. (2000). *Visual Design for Instructional Multimedia.* Retrieved September 2, 2003, from http://www.indiana.edu/%7Evdim/Intro/TOC1.HTM

Moore, D. S. (1997). *Statistics: concepts and controversies* (4th ed.), New York: W. H. Freeman and Co., Inc.

Norman, D. A. (1990). *The design of everyday things*. New York: Currency Doubleday.

Osguthorpe, R. T. (1985). Conducting literature searches for instructional development projects. *Journal of Instructional Development*, 8(4), 20-24

Peeck, J. (1987). The role of illustrations in processing and remembering illustrated text. In D. M. Willows and H. A. Houghton (Eds.), *The psychology of illustration*, volume 1 (pp. 115-151). New York: Springer-Verlag.

Rieber, L. P. (1990). Animation in computer-based instruction. *Educational Technology Research and Development*, 38(1), 77-86.

Rosson, M. B. & Carroll, J. M. (2002). *Usability Engineering: Scenario-Based Development of Human-Computer Interaction*. Morgon Kauffman Publishers, San Francisco.

Ruben, A., & Bruce, B., Statistics Workshop (software), (1991), BBN, Inc., Cambridge, MA.

Spool, J. M., Scanlon, T., Schroeder, W., Snyder, C., & DeAngelo, T. (1997). *Web site usability: A designer's guide*. North Andover, MA: User Interface Engineering.

Velleman, P. F., & Moore, D. S. (1996). Multimedia for teaching statistics: Promise and pitfalls. *The American Statistician,* 50 (3), 217-225.

Venn, J. (1880). On the diagrammatic and mechanical representation of propositions and reasonings*, The London, Edinburgh, and Dublin Philosophy Magazine and Journal of Science,* 5*,* 1-18.

West, R. W., & Ogden, T. R. (1998). Interactive demonstrations for statistics education on the world wide web. *Journal of Statistics Education*, 6, 3. Retrieved on September 2, 2003, from www.amstat.org/publications/jse/v6n3/west.html

Xiaoquan, X., & Jones-Marshall, G. (1995). Computer animation for EFL learning environments. *Selected readings from the annual conference of the International Visual Literacy Association*. Chicago, Illinois, October 18-22. (ERIC Document Reproduction Service No. ED 391 518).

Yanchar, S. (2004). Empirical inquiry and statistics course syllabus. Retrieved February 16, 2004, from http://www.byu.edu/ipt/courses.html.

Appendix A

SLRL General Instructional Objectives (GIO) and Specific Learning Outcomes (SLO)

GIO 1: Understand the basic principles of simple linear regression

SLO 1.1: Describe how simple linear regression relates to correlation

SLO 1.2: Describe the purpose of simple linear regression

SLO 1.3: Contrast the purposes of simple regression and correlation

GIO 2: Understand the basic assumptions of the method of least squares

SLO 2.1: Explain the principle of linearity

SLO 2.2: Explain the principle of homoscedasticity

SLO 2.3: Explain the assumption of normal distribution among each separate value of x

SLO 2.4: Describe the consequences of violation of the assumptions

GIO 3: Understand the regression equation

SLO 3.1: Select the necessary components for calculating regression

SLO 3.2 Explain how the regression equation relates to the equation of a straight line

GIO 4: Apply the regression equation correctly

SLO 4.1: Interpret the meaning of the calculated Y prime for different individuals

SLO 4.2: Identify research situations in which use of regression analysis would be

appropriate

SLO 4.3: Identify research situations in which use of regression analysis would *not* be

appropriate

GIO 5: Understand the standard error of the estimate

SLO 5.1: Describe the purpose for using the standard error of the estimate

SLO 5.2: Understand the concept of residuals

SLO 5.3: Explain how the standard error of the estimate and standard deviation of the dependent variable are similar

SLO 5.4: Describe the definitional formula for the standard error of the estimate

GIO 6: Apply the formula for the standard error of the estimate correctly

SLO 6.1: Interpret the meaning of the calculated standard error of the estimate

GIO 7: Understand the coefficient of determination and the coefficient of non-determination

SLO 7.1: Explain the purpose for calculating the coefficient of determination

SLO 7.2: Describe variance accounted for

SLO 7.3: Describe unexplained variance

SLO 7.4: Describe the definitional formula of the coefficient of determination

SLO 7.5: Compare and contrast the coefficient of determination and the coefficient of non-determination

Appendix B

Simple Linear Regression Test

Questions 1-4:

For each research topic, decide if it would be appropriate to use simple linear regression to analyze the data. Circle "Yes" if simple linear regression would be appropriate and "No" if simple linear regression would not be appropriate.

1. Yes   No   You want to know the proportion of variability in ACT scores that is related to GPA.

2. Yes   No   You want to analyze the direction and relationship between test anxiety and procrastination.

3. Yes   No   You want to predict scores on a performance test based on scores on a written test.

4. Yes   No   You want to find out how a computer-based lesson compares to a textbook in teaching statistics.

Questions 5 and 6: Circle the best answer

5. How are regression and correlation related?

    a. The definitional formula for calculating each statistic is a mathematical model for a straight line.

    b. Both explain the magnitude and the direction of the relationship between two variables.

    c. In order for regression to be most useful, there must be a positive correlation.

    d. The accuracy of predictions made by using regression depends on the strength of the correlation.

6. How do regression and correlation differ?

    a. Correlation analyzes the relationship between two variables; simple linear regression analyzes the relationships between three or more variables.

    b. In calculating correlation, no distinction is made between the independent and dependent variables.

    c. Simple linear regression is more prone to error than correlation.

    d. Correlation describes the relationship between two variables; regression simply predicts the value of one variable based on another.

Answer questions 7-9 using the following information:

An instructional designer wants to understand how scores on a post-training assessment are correlated with two other variables. Based on the data collected he draws the following scatter plots:
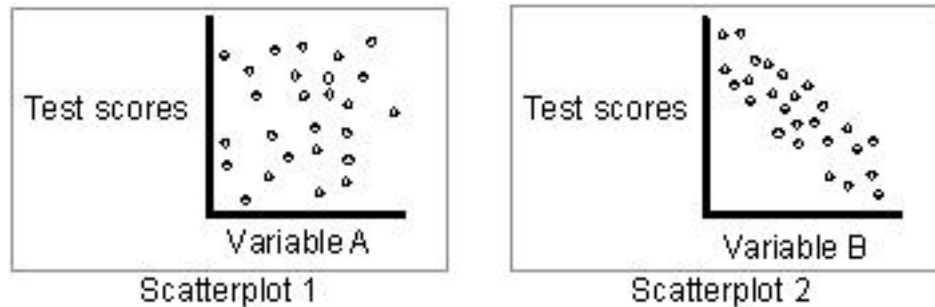
7. Which correlation coefficient most accurately describes Scatter Plot 1?

    a. –0.8

    b. –0.4

    c. 0.0

    d. 0.4

    e. 0.8

8. Which correlation coefficient most accurately describes Scatter Plot 2?

    a. –0.8

    b. –0.4

    c. 0.0

    d. 0.4

    e. 0.8

9. The instructional designer would like to use one variable (either A or B) to estimate the assessment score for individuals in his target audience. Predict how accurate the designer's estimations would be if he were to use each variable. Which variable or set of variables would you recommend he use? Explain the logic behind your recommendation.

10. An educational researcher has asked you to help him conduct a study using simple linear regression. The researcher collects data and records it on this scatter plot. What recommendations would you make about how to proceed with the statistical analysis? Explain the logic behind your recommendation.
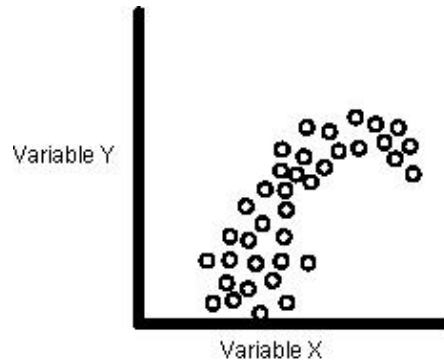
11. Explain the purpose of the regression line. Include in your answer an explanation of why it is called the best fitting line. Explain how the regression equation calculates the best fitting line.

12. You are at an academic conference on education. During one of the presentations you attend, the presenter reports that he has found a variable that can predict ACT scores, which range from 18 to 36. He calculated the standard error of the estimate to be 5. Explain what this statement means.

Appendix C

SLRL Agent Script

Welcome Screen

Welcome to the Simple linear regression lesson. During the next few minutes I will introduce to you the statistical concepts that make up the principle of simple linear regression. Press the menu button to go to the menu of lesson topics.

Menu Screen

If this is your first time studying this lesson, I recommend you proceed through the topics in the order listed. If you are reviewing the lesson, feel free to go to any section. Simply click on the section that you would like to study. When you complete a section you will automatically return to this menu.

Introduction: the purpose of Simple Linear Regression

The purpose of simple linear regression is to use one variable to make predictions and estimations about another variable. For example: A graduate school admissions committee is trying to make decisions about which applicants they will admit for the upcoming semester. In order to decide which students to accept, the school needs to predict how well the applicants will do in graduate school. Specifically, the admissions committee wants to predict the applicants' GPA. What kind of data do they have on hand to make that kind of prediction? For one thing, they know the GPA's of previous students, which range from 2.5 to 4.0. The mean for previous students is 3.4. So, what kind of prediction can the admissions committee make? With no other information, the best prediction for any individual applying for admissions is the mean from the data they have, available. By predicting the mean, the committee will be closer, on average, than

if they were to guess arbitrarily. However, this does not provide them much help them in making admissions decisions.

In order to make a more accurate prediction, the admissions office can use other information they know about the students. For example, the committee has record of the student's scores on the GRE. So, they dig out the GRE records from previous years and find a range of scores. They then combine the two scales by putting the GPA values on the Y axis and the GRE scores on the X axis. At this point the committee can see that students with low scores on the GRE also earned low GPA's in graduate school, and students with high GRE scores earned high GPA's. With these data from previous years, the admissions committee can make better predictions about how well applicants will do. If the committee knows that a student scored 1800 on the GRE, for example, they can predict his graduate GPA will be 3.3.

Why don't you give it a try, predict the graduate GPA for a student who got a 2000 on the GRE entrance exam. Click on the Y axis to select the GPA value that you would predict for that student.

Using what you have just learned about the purpose of simple linear regression, decide which of the following research questions should be answered using simple linear regression:

To conclude this section let's quickly review: The basic purpose of simple linear regression is to use one variable (x) to help us make predictions about another other variable (y). During the rest of the lesson we will go into greater detail about how to use and apply simple linear regression.

<center>Review Correlation</center>

Before we get any deeper into simple linear regression, let's quickly review correlation.

The purpose of correlation is to analyze how two variables are related to each other. For example, an educational researcher might be interested in how the amount of study time correlates with test scores or how stress levels correlate with performance.

The relationship between the variables can be expressed in the form of a statistic called the correlation coefficient, or "r". The correlation coefficient is a number between -1 and 1 that tells us two important things about the relationship between the two variables: first, the direction of the relationship, and second, the magnitude (or strength) of the relationship. Let's briefly review each point.

Point number one: Direction. The direction of the correlation between two variables is either positive or negative. A positive correlation means that low scores in one variable are associated with low scores in the other variable and, similarly, high scores in one, are associated with high scores in the other. For example, weight and height are positively correlated. In most cases, the taller someone is, the more they weigh. When two variables are positively correlated the correlation coefficient is between zero and positive one.

A negative correlation means that high scores in one variable are associated with *low* scores in the other variable, and vice versa. *(Pause)* The correlation between a mothers smoking habits and the birth weight of her child is an example of a negative correlation. As the number of cigarettes smoked *increases,* the birth weight of the child *decreases.*

Point number two: Magnitude. The magnitude is the strength of the correlation between the two variables. The strongest possible correlation is one where the two variables change uniformly all the time. For example, every time we add one cup of water to a bucket, the weight of the bucket increases by eight ounces. It will happen every time without fail! The correlation coefficient for this example would be positive one. This is a perfect *positive* correlation because

both variables are increasing uniformly. We can also have a perfect negative correlation where every time one variable increases, the other decreases. In this case the correlation coefficient would be negative one. The closer the coefficient is to positive one or negative one, the stronger the correlation between the two variables will be.

The opposite of a perfect correlation is zero correlation. Zero correlation means that there is no relationship between the two variables. For instance, there is zero correlation between the color of a person's hair and the length of their hair. It would be absurd to think that people with the same length of hair consistently have the same color of hair. The closer the correlation coefficient is to zero, the weaker the relationship between the two variables.

Take a moment to experiment with the correlation coefficient. Use the slider to change the value of the correlation coefficient and observe how the data points change. As you observe the change in the strength and the direction of the data, think about how this relates to what you already know about simple linear regression. When you are done, click the next button.

From the correlation slider activity, you should have recognized that the closer the coefficient is to zero the more uniform the changes in the variables are. In fact, perfect correlations form an imaginary straight line. The closer the coefficient is to zero, the less uniform the changes. Using what you have just learned about the spread of data and the correlation coefficient take a look at this scatter plot. What would you estimate the correlation coefficient to be for these data?

Before we conclude this section, let's briefly summarize correlation. Correlation is an analysis of the relationship between two variables. The correlation coefficient tells us two things about the two variables that we are analyzing: the direction of the relationship and the magnitude, or strength of the relationship. By the way, some students get confused about the

difference between correlation and regression. Remember that correlation simply measures the relationships. Regression, on the other hand, uses the correlation of two variables to make predictions about one of the variables. We'll talk more about that later.

<div align="center">The Regression Line</div>

Let's talk a little more about regression. If two variables are perfectly correlated, their correlation coefficient is either exactly positive one or exactly negative one, then prediction is somewhat simple. Notice that the data here forms a perfectly straight line. To make a prediction, we can use a student's score on X to predict what their Y value will be using the line as an indicator. With the line, we can even predict someone's score on Y for a previously un-encountered score on X. This line is called the regression line. Until now, we have been discussing examples that have two variables with a perfect linear relationship. In perfect linear relationships it is easy to see where the regression line is. However, actual data almost never assume a perfect linear relationship. Click on the button to see a more accurate example of what actual data look like. Now, with data that look like this, where do we create a regression line for prediction? Here? Here? Maybe here? In these cases the best regression line is a line that comes closest to every single data point. This line is called the "best fitting line." The best fitting line, is a running mean of the two variables. If we have one only one variable then we have a single average, or mean. If we have two variables we can have a running mean. In order to know exactly where the best fitting line lies on the data set, we use a regression equation.

So let's talk about using the regression equation to make predictions. There are two important points that we will consider as we discuss the regression equation: Point one: The regression equation is actually an equation for a straight line, and Point two: The regression equation isn't an equation for just any line, it is an equation for the best fitting line.

Many of today's software programs can calculate the regression line for you in a matter of seconds, so you may never use the actual equation. Even so, it is important to understand the concepts behind the equation so you can appropriately analyze and interpret data.

Let's start with point one: The regression equation is actually an equation for a straight line. You may remember from math class that a straight line is defined by two pieces of information. First, we need to know where the line crosses the Y axis, which is referred to as the *intercept.* Most statistics texts label the intercept with a lower case a. The second piece of information we need is the *slope* of the line. The slope is the amount of difference in Y associated with a one-unit difference in X (in math you would call this the rise over the run). A slope of one means that every time X increases by one unit, Y also increases by one unit. If the slope is five, then every time x increases by one, then Y increases by five. And so on. Statistics texts label the slope with a lower case b.

Now take a look at a definitional formula for the regression equation. What do you see? You should recognize that the equation contains a, the intercept, and b, the slope. If you recall from your math classes, this is also the formula for a straight line. Y equals bX plus a.
In the case of regression, we use the equation of the straight line to predict a value for Y. To make a prediction, all we need to do is find an individual's score on X, plug it into the equation, and solve for Y. This assumes that we already know the line's slope and the intercept. We will discuss that next.

The second point that we need to discuss is that the regression equation is the equation for the best fitting line, or the line that comes the closest to all of the data points. If we expand the definitional formula you will be able to see what information we need in order to calculate the best fitting line. In order to calculate the slope of the line we need the correlation coefficient for

the two variables, and the standard deviations of both variables. To calculate the intercept we need the correlation, deviations, and also the mean of both variables. It is probably not important to memorize this definitional formula to a tee, but you should understand the following statement: the regression equation calculates the best fitting line by using the correlation coefficient of the two variables, the mean of both variables, and the standard deviation of both variables. Using these values, the slope and the intercept for the regression line take into account the location of every point in the data set and come as close as possible to each of them. When solving for Y using the regression equation we actually call the predicted score Y prime. Y prime is different than Y because our prediction may not be accurate. Our regression line may predict that a students GPA will be 3.3 when it really turns out to be 3.5. When we make predictions using data that is less than perfect, we will inevitably make mistakes. These mistakes are called prediction error. So, the best fitting line is actually the line that minimizes the differences between y and y prime, or prediction error.

Let's take a break. Here is a sample data set analyzing the amount of time spent studying vocabulary and the number of vocabulary words memorized. You can see that the data form a roughly linear relationship. The red line is the regression line for this data set. You can tell that this is by no means the best fitting line, your job is to manipulate the slope and the intercept of the regression line until it is the best fitting line. You will know when the line fits best because the amount of error is at its lowest possible value.

The purpose of this activity is to demonstrate that the regression line fits best because it minimizes the amount of mistakes, or error, in our predictions.

Before we move on let's review some of the important points that we have discussed about the regression line. The regression line is a best fitting line that comes as close as possible to

every point in the data. The regression line is a running mean of the two variables. We can use the regression equation to calculate the best fitting line.

<div align="center">Statistical assumptions of simple linear regression</div>

There are three important assumptions that must be taken into account when performing simple linear regression. If any of these are violated the ability of X to serve as a predictor for Y comes into question. They are: the assumption of linearity, the assumption of normal distribution, and the assumption of homoscedacticity

The assumption of linearity

The defining feature of a linear relationship is that whenever the value of X changes by a given amount, the amount of change in Y is consistent. Two variables can be related in to each other in nonlinear ways. The most common alternative to linearity is curvilinear.

In curvilinear relationships the amount of change in Y is not consistent across all values of X. For example, an exponential relationship shows that there is a relationship between the two variables, but it does not fit our definition of linearity.

Another common type of curvilinear relationship is best described with an example explanation. Think for a moment about the relationship between the level of student anxiety and test performance. The relationship is best describes by stating that as anxiety increases from low to moderate, performance increases, but as anxiety increases from moderate to high test performance decreases.

Simple linear regression is not appropriate where a curvilinear exists between two variables. The ability to make predictions using simple linear regression is based on the linearity of the data set so that the regression line that comes closest to all of the points of data is actually a straight line. Remember that the regression equation gives us a straight line and will therefore

give us inaccurate predictions. Carefully analyzing the scatterplot of your data will help you determine whether or not you have a linear relationship.

The assumptions of normal distribution and homoscedasticity

The other important assumptions of simple linear regression address the entire population from which the sample comes. Let's look at it this way, if a researcher is analyzing GPA and GRE scores for graduate students, he knows that there are hundreds of thousands of individuals in this population. The researcher's study is only a small sample of the entire population. Even though the researcher can only see his sample, he must make certain assumptions about the entire population. Now, if we look at the entire population we can see that each GRE score has a subpopulation of GPA scores. The assumptions of normal distribution and homoscedasticity deal with the subpopulations of the X scores. A three-dimensional view best illustrates these assumptions.

Simple linear regression assumes that each subpopulation of the predictor variable assumes a normal distribution. The method of least squares uses the subpopulation mean as the best estimate because it contains the least amount of prediction error. Previously we said that the regression line is a running mean. So, if we view the entire population, the regression line passes through the mean of each subpopulation of X values. If the scores of the subpopulation are skewed or bimodal, the mean is no longer the center of the distribution, nor is it the most frequently occurring score. In these cases the regression line will not be the function that comes closest to the actual scores of Y. Therefore you can't rely on its ability to predict.

The assumption of homoscedasticity is related to the assumption of normal distribution. Homo means *same* and *scedasticity* means "tendency to scatter." In other words, simple linear regression assumes that the standard deviation of the Y scores is similar for every subpopulation

of X. The assumption of homoscedasticity states that there should be just as much variability among individuals who score 2 on X as there is among individuals who score 20 on X. If this assumption were violated, the regression equation would make more accurate predictions for some values of X and less accurate predictions for other values of X.

Good. Before we conclude this section let me ask you a couple of questions. Let's say that you want to use regression to make predictions about how instructional design will affect test scores in an elementary school classroom. When you plot the data it looks like this. Should you use simple linear regression in this case? Select an option below.

Okay. Let's say that you plot the data and it looks like this. Should you use simple linear regression in this case? Select an option below.

One more. This time your data looks like this. Should you use simple linear regression in this case? Select an option below.

Before we conclude this section let's review the assumptions of simple linear regression. When using simple linear regression we assume that the correlation between the data forms a linear relationship. We also assume that the distributions of x subpopulations are normal and that they have the same tendency to scatter. If these assumptions are violated we will be unable to make accurate and valid predictions using simple linear regression.

The accuracy of our prediction (error)

To begin our discussion about prediction error let me ask you a few questions. Take a look at this data set. Predict the Y score for an individual who scored where the purple mark is on the x axis. Click on the Y axis to select the value that you would predict for that student.

Now let's change the dataset a bit. Predict the Y score for an individual who scored where the purple mark is on the x axis. Click on the Y-axis to select the value that you would predict for that student.

Let's do it one more time. Predict the Y score for an individual who scored where the purple mark is on the x axis. Click on the Y-axis to select the value that you would predict for that student.

Why were you less certain about the last example than you were certain about the previous examples? Well, the answer has something to do with error. You correctly assumed that the more variance in scores the more mistakes we would make when making predictions.

Previously in the lesson we introduced the idea of error. We said that when we use the method of least squares we will inevitably make mistakes in our predictions. The prediction error of one individual is simply how far the actual score deviated from the predicted score, or Y prime. The error, or residual, will not be the same for every score, so in order to evaluate the *overall* accuracy of our regression line we need to calculate the average difference between all of the actual y scores and the predicted values of Y prime. The statistic we calculate to average the error in our predictions is called the *standard error of the estimate.* Once again it is important to note that computer software can quickly calculate the standard error of the estimate for a dataset but it is important to understand the fundamentals so that you can appropriately interpret the statistic.

To understand the standard error of the estimate, look at this example. In this study, the researcher is using the IQ of a parent to predict the IQ of his or her child. The researcher first collects all the IQ scores for the parents (X) and the children (Y). Next, the researcher calculates a slope and intercept and builds a regression equation. He calculates Y prime for every value of

X. Then, the researcher finds the difference between each y and y prime value. Then, he squares all the differences, and adds them up. This sum is the total amount of error in the data set. To get the standard error of the estimate, he first divides the sum of the error by one minus the number of participants in the study, then he calculates the square root of that number. In this case, the standard error of the estimate is 9.68. To the researcher this means that when he uses the parent's IQ to predict a child's IQ, he will off by an average of 9.68 IQ points. In many cases, his prediction will be within 9.68 points, but in other cases he will be farther off the mark. The standard error of the estimate is the average of the amount of prediction error in the dataset. Error is inevitable, that is why we call the regression line the best fitting line, not the perfect fitting line. But when we use the best fitting line, we are making the most of the predictability of the dataset. The process of calculating the best fitting line is sometimes called "the method of least squares." Here is the reason: When we calculate the amount of error in our dataset, we square the difference between Y and Y prime, the regression line is the line that minimizes these squared differences, and hence the method of *least* squares.

Now, there is an important principle to discuss about error. The general rule of regression analysis is that as the correlation between X and Y becomes stronger, the predictions will be more and more accurate. If the two variables form a perfect correlation, we will make no mistakes in our prediction and the error will be zero. On the other hand, if there is zero correlation between the two variables, the best prediction we can make for Y is the mean, which, as you can imagine is full of error.

Why don't you try it? Slide the knob across possible values of the correlation coefficient. As you slide it, keep an eye on the scatter plot and the amount of error in the data.

In summary, the stronger the correlation between two variables the more accurate your predictions will be. In contrast, the weaker the correlation, the more error you will have in your predictions.

Take a look at these scatter plots. Pick which one you think would yield the least amount of error for making predictions. Click on your selection.

Which one do you think would yield the most amount of error? Click on your selection. This process is called the method of least squares because it minimizes the squared differences of Y and Y prime.

To conclude this section let's review what we have discussed. Prediction error is the difference between the predicted score and the actual score. To find out how accurate our predictions are, we can calculate a statistic call the standard error of the estimate. We ended this section by discussing that if there is a strong correlation between the two variables, there will be less error when we make predictions. If the there is a weak correlation, there will be more error in our predictions

Coefficient of determination and coefficient of non-determination

A statistic of interest in both correlation studies and simple regression analysis is the coefficient of determination. To understand this statistic let's discuss an example. Let's say that we are analyzing the cumulative GPA for University students. This circle represents all of the variation among GPA scores. There are an infinite number of variables that could account for variance in GPA scores, for example, the students major, the amount of sleep the student got, how many credits the student took etc. etc. The fact is that we don't know why all of this variance occurs in GPA scores. We may have some hunches but we aren't sure.

We use correlation to analyze the relationship between GPA scores and other variables that could be related to it, such as those mentioned above. Here is another circle that represents the variance in another variable, say for instance, the amount of time students' study per week. We want to

To better understand this statistic, let's imagine that all of the variance in Y scores can be represented by a circle. Now, the purpose of correlation is to find out how much variance in Y is systematically related to variance in another variable. So, let's say that we have another circle that represents all of the variance in variable X. If there is a correlation between the two variables we can conceptually break the total variability in Y into two parts:

The variability in Y that is associated with changes in X, or variance accounted for,

and the variability in Y that is not associated with changes X, or variance not accounted for.

If you add up these two components, they equal to the total variability in the Y scores.

This is the idea behind the statistic called the coefficient of determination. The coefficient of determination tells us the proportion of Y variability that is explained by variation in X. In other words, variance accounted for over total variance. This correctly implies that the coefficient of determination is expressed in percentage form. For example, a researcher analyzing the correlation between the IQ of a parent and the IQ of a child might report that fourteen percent of the differences in children's IQ vary systematically with the differences in their parent's IQ. The remaining 86 percent of the differences among children are not related to the variation in parental IQ. This unexplained variance is known as the coefficient of non-determination. So, the coefficient of determination and the coefficient of non-determination are inverse statistics; as one increases the other decreases, and vice versa.

Take a moment and slide the knob across different values of the coefficient of determination. Not e that the higher the coefficient of determination, the greater the overlap between the two variables, or the more variance accounted for.

Good. As you moved the slider you should have recognized that the coefficient of determination describes the amount of variance in a variable that is accounted for by another variable. It is very important to note that the coefficient of determination does not imply a causal relationship. It is simply a register of how the variables vary together.

Before we conclude, I'll make one more point about the coefficient of determination. The amount of variance accounted for, or the value of the coefficient of determination, is a function of the value of the correlation coefficient. In fact, the easiest way to calculate the coefficient of determination is to simply square the correlation coefficient. Squaring the correlation coefficient gives us a percent of variance accounted for. Therefore, the stronger the correlation between two variables, the greater amount of variance accounted for.

To better understand the relationship between the coefficient of determination and the correlation, check out this slider activity. This one is a little different than the one you just did. Here the knob on the slider represents the correlation coefficient. As you move the slider you will notice the circles overlap will match the numerical value of the coefficient of determination in the box. Give it a try.

You should have recognized that the coefficient of determination is a function of the correlation coefficient.

Conclusion

During this lesson we have expanded on the correlation between two variables to cover situations in which we want to predict one variable from our knowledge of the other variable. This is simple linear regression. Using the regression equation we can calculate a regression line that comes closest to all the data points in our dataset. We also discussed the standard error of the estimate as a measure of the accuracy of our prediction. We learned that the regression line is sometimes called the "best fitting line" because it minimizes the amount of error in our predictions. We concluded the lesson by discussing the coefficient of determination as an index of the variance in one variable accounted for by another variable. I hope that you have enjoyed this lesson. Feel free to return and review specific parts of the lesson or the entire lesson.

Appendix D

Evaluation Questions, Criteria, and Standards

Stakeholder question 1: Did the lesson stand-alone?

Criteria: The SLRL should be standalone. It should not require additional help from

technical specialists or subject matter experts in accessing or comprehending lesson

material. The SLRL should be delivered on time and be used in the normal instructional

sequence of the 550 course.

Standard 1A: When participating in the SLRL, users receive no additional content help.

Standard 1B: The SLRL will be delivered no later than October 20, 2003.

Stakeholder question 2: Did the lesson function properly?

Criteria: The SLRL should be free from technical bugs that distract the user from lesson

content or impede learning all together.

Standard 2A: The SLRL will be 100 percent free from technical bugs and errors.

Standard 2B: Users should wait no longer than three seconds while activities and course

content loads.

Stakeholder question 3: Could users perform the necessary lesson tasks with the interface?

Criteria: The interface should be intuitive and user friendly so that users are able to

perform the required interactive tasks with ease and navigate within the lesson without

outside help.

Standard 3A: Users will need no outside guidance in understanding navigation and

content tasks.

Stakeholder question 4: Did the graphics, animations and interactive animations help students learn?

Criteria: The nature of the SLRL should be such that it takes the content above the level presented by textbooks. Empirical evidence should suggest that the lesson helps students understand the content more than a textbook or classroom.

Standard 4A: Using a traditional pre-post, control-experiment design, differences in mean scores will be statistically significant at an alpha level of .05.

Standard 4B: Using a traditional pre-post, control-experiment design, differences in mean scores will be practically significant.

Stakeholder question 5: Did the students enjoy the lesson?

Criteria: The nature of the SLRL and its presentation of the statistical content should be engaging and enjoyable to students.

Stakeholder question 6: Were the learners able to meet the specific objectives of the lesson?

Criteria: Users should be able to demonstrate their understanding of the SLRL content by fulfilling learning outcomes specified in the general instructional objectives.

Standard 6A: At the conclusion of the lesson, users will be able to achieve 85% or higher on a criterion referenced test on simple linear regression.

Stakeholder question 7: Does the lesson demonstrate principles of sound instructional design?

Criteria: The flow of content and instructional strategies used in the SLRL should be based in the research of the instructional design field and be "best practice" applications.

Standard 7A: Expert instructional designers and the body of research in the field should approve of the instructional methods and strategies applied in the SLRL.

Appendix E

SLRL Example Screen Layout and Color Palette

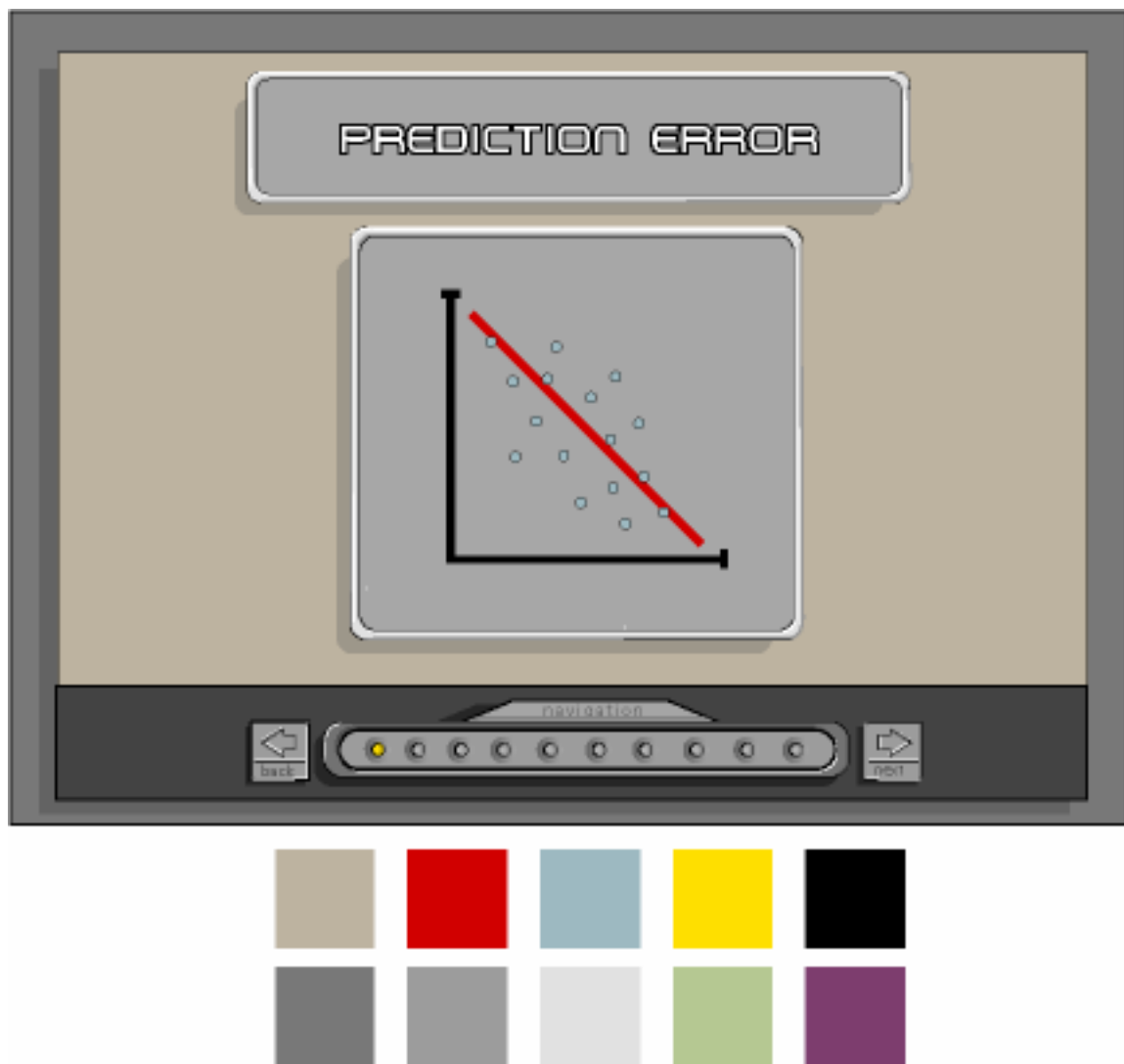Figure E1: SLRL layout and color palette.

Appendix F

Post-lesson Questionnaire

1. Have you taken statistics classes prior to IPT 550? *(Circle one)*
  Yes No

  If Yes, when and where did you take statistics courses?

2. How challenging is it for you to learn Statistics?
Circle a number.

| very | 1 | 2 | 3 | 4 | 5 | 6 | 7 | very |
| easy | | | | | | | | challenging |

For each one of the following questions try to remember the section that is referred to (you may go to that section and review it). Mark "yes" or "no" in order to answer the questions about each specific section. Write in the space named "Comments" any comments that you have about the section.

3. Introduction to Regression:
  Yes No Did this section function properly?
  Yes No Was the instruction clear?
  Yes No Were the graphics/animations helpful?
  Yes No Were the interactive activities helpful?
  Yes No Did you enjoy this section?

  Comments:

4. Review of Correlation:
  Yes No Did this section function properly?
  Yes No Was the instruction clear?
  Yes No Were the graphics/animations helpful?
  Yes No Were the interactive activities helpful?
  Yes No Did you enjoy this section?

  Comments:

5. The Regression Line:
       Yes No   Did this section function properly?
       Yes No   Was the instruction clear?
       Yes No   Were the graphics/animations helpful?
       Yes No   Were the interactive activities helpful?
       Yes No   Did you enjoy this section?

       Comments:

6. Assumptions of Simple Linear Regression:
       Yes No   Did this section function properly?
       Yes No   Was the instruction clear?
       Yes No   Were the graphics/animations helpful?
       Yes No   Were the interactive activities helpful?
       Yes No   Did you enjoy this section?

       Comments:

7. The Accuracy of the Prediction (Error)
       Yes No   Did this section function properly?
       Yes No   Was the instruction clear?
       Yes No   Were the graphics/animations helpful?
       Yes No   Were the interactive activities helpful?
       Yes No   Did you enjoy this section?

       Comments:

8. Conclusion:
       Yes No   Did this section function properly?
       Yes No   Was the instruction clear?
       Yes No   Were the graphics/animations helpful?
       Yes No   Were the interactive activities helpful?
       Yes No   Did you enjoy this section?

       Comments:

Answer the following questions by writing the title that corresponds to the part of the lesson and explain why you chose it.

9. Which part(s) of the lesson was your favorite(s)? Why?

10. Which part(s) of the lesson was your least favorite(s)? Why?

11. Which part(s) of the lesson did you learn the most from? Why?

12. Which part(s) of the lesson did you learn the least from? Why?

13. Which part(s) of the lesson motivated you most to continue learning statistics? Why?

14. Which part(s) of the lesson discouraged you from learning statistics? Why?

15. Would you like more statistics lesson with a similar approach? Why or why not?

16. Other comments or suggestions?

Appendix G

Summary Table of Post-lesson Questionnaire Responses

Table G1

Summary of Responses to the Post-Lesson Questionnaire

Question	Responses	Number
1 Have you taken statistics before?	Yes	12
	No	3
	1	0
	2	3
	3	3
	4	3
	5	3
	6	1
2 How challenging is it for you to learn statistics?	7	2
	No comments	6
	Impressed, Good Job, good instruction	6
	Too fast	1
3 Introduction to regression	Good end of section exercises	2
	No comments	7
	Good data-driven animation	2
	Error in audio	2
4 Review of correlation	Great job	4
	No comments	6
	Error in audio	2
	Good data-driven animations/ visuals	5
5 The regression line	Very good	2
	No comments	5
	Error in audio	2
	Very good	3
	Great 3d images	3
	Spelling mistake	1
6 Assumptions of Simple linear regression	3d images confusing	1
	No comments	6
	Pace to slow	1
	Technical bug on review questions	3
	Great data-driven animation	3
	Clear understanding	1
	Audio error	1
7 The accuracy of prediction	Very good	2
8 Conclusion	No comments	11
	Very good	1
	Too fast	1

	Clear understanding	1
	Not necessary	1
	Regression line module	2
	Graphics and animations	3
	Data-driven animations	9
9 Favorite part of lesson	Assumptions of simple linear regression	1
	No comments	2
	Assumptions of simple linear regression	4
	Worked example	3
	None	3
	Summary of each section	1
	Simple questions	1
10 Least favorite part of the lesson	Accuracy of prediction	1
	No comments	2
	Regression line module	4
	Assumptions of simple linear regression	4
	Graphics and animations	2
	Accuracy of prediction	1
11 Part learned most from	All of the modules	2
	No comments	5
	Latter sections	1
	Prediction error	3
	None	2
	Correlation review (already knew)	3
12 Part learned least from	Assumptions of simple linear regression	1
	No comments	3
	Review of correlation	1
	Graphics and animations	5
	All of it	3
	Examples	1
	Regression line module	1
13 Part that motivated you to continue learning stats	None of it	1
	No comments	6
	Discouraged before lesson	1
	Worked example/equations	4
14 Part that discouraged you from learning stats	None	3
15 More lessons with similar approach	Yes	15
	No comments	4
	Good job	9
16 Other comments	Menu button throughout	2

Appendix H

Pretest Scores, Posttest Scores, and Experience

Table H1

Pretest Scores, Posttest Scores, and Statistics

ID	Pretest	Posttest	Experience
98	4	19	0
1	6	20	0
2	6	19	1
44	15	20	1
70	12	22	1
35	12	21	1
79	12	23	1
11	15	23	1
20	7	18	1
31	13	24	1
13	8	21	0
94	23	24	1
30	10	21	0
76	10	22	1
17	9	24	1