Decision Based Learning Course Design & Implementation for Introductory Statistics

Austin Heath

Follow this and additional works at: https://scholarsarchive.byu.edu/studentpub_uht

BYU ScholarsArchive Citation
Heath, Austin, "Decision Based Learning Course Design & Implementation for Introductory Statistics" (2021). Undergraduate Honors Theses. 204.
https://scholarsarchive.byu.edu/studentpub_uht/204

This Honors Thesis is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in Undergraduate Honors Theses by an authorized administrator of BYU ScholarsArchive. For more information, please contact ellen_amatangelo@byu.edu.
DECISION BASED LEARNING COURSE DESIGN & IMPLEMENTATION FOR INTRODUCTORY STATISTICS

by

Austin Heath

Submitted to Brigham Young University in partial fulfillment of graduation requirements for University Honors

Statistics Department
Brigham Young University
July 2021

Advisor: Lynne Nielsen
Honors Coordinator: Del Scott
ABSTRACT

DECISION BASED LEARNING COURSE DESIGN & IMPLEMENTATION FOR INTRODUCTORY STATISTICS

Austin Heath
Statistics Department
Bachelor of Science

Submitted to Brigham Young University in partial fulfillment of graduation requirements for University Honors

ABSTRACT

Researchers in multiple industries (biomedicine, engineering, etc.) cite the selection of an appropriate statistical test as a common problem. Experts draw on a framework of conceptual and procedural knowledge to navigate when to use statistical methods. Students also struggle determining the correct statistical method to use for a given research question. This is because they lack the opportunity to practice recognizing a host of features in each research question that provide clues for experts as to which method is most appropriate. “Decision Based Learning” (DBL) is a teaching method designed to help teachers and students address this struggle. In this study we create and implement a decision model for choosing which of the 14 statistical methods, taught in an introductory education statistics course, is most appropriate for the research questions. We compared the performance on method selection test questions of 1021 students using DBL and a comparison group of 930 students during the Winter semester of 2021 at Brigham Young University. The treatment and comparison groups were composed of randomly assigned sections of students. The performance on examinations lacks sufficient evidence to show that DBL significantly improved method selection capabilities. The recommendations in this study provide potential improvements in the presentation and delivery of decision-based learning for introductory education statistics.


**TABLE OF CONTENTS**

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td>i</td>
</tr>
<tr>
<td>Abstract</td>
<td>iii</td>
</tr>
<tr>
<td>Table of Contents</td>
<td>v</td>
</tr>
<tr>
<td>List of Figures</td>
<td>vii</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>ix</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Literature Review</td>
<td>4</td>
</tr>
<tr>
<td>Research Questions</td>
<td>5</td>
</tr>
<tr>
<td>Methods</td>
<td>6</td>
</tr>
<tr>
<td>Limitations</td>
<td>22</td>
</tr>
<tr>
<td>Results &amp; Discussion</td>
<td>23</td>
</tr>
<tr>
<td>Implications for Research &amp; Practice</td>
<td>32</td>
</tr>
<tr>
<td>Conclusion</td>
<td>34</td>
</tr>
<tr>
<td>Cited</td>
<td>35</td>
</tr>
<tr>
<td>Appendix</td>
<td>38</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Box 1: Statistical Procedure Selection Example Question 6
Figure 1: Decision Model for Statistical Method Selection 8
Figure 2: Feature Identification Example 10
Figure 3: Learning Module Definitions Section Example 12
Figure 4: Learning Module Example Problem Section Example 12
Figure 5: Learning Module Summary Section Example 13
Figure 6: Learning Module Practice Section Example 13
Figure 7: Learning Module Review List 14
Figure 8: Synthesis point learning module (part 1) 18
Figure 9: Synthesis point learning module (part 2) 18
Table 1: Feedback Plan for Quizzes & Exams 19
Table 2: Summary of Treatment Group vs Comparison Group Performance on Exam 3 23
Boxplot 1: Treatment Group vs Comparison Group Performance on Exam 24
Graph 1: Treatment Group vs Comparison Group Performance on 13 Individual Questions 24
Table 3: 13 Procedure Questions on Exam 3 25
Table 4: Summary of Treatment Group vs. Comparison Group Performance on Final Exam 27
Boxplot 2: Treatment Group vs Comparison Group Performance on Final Exam 27
Graph 2: Treatment Group vs Comparison Group Performance on 13 Individual Questions 28
Table 5: 13 different Procedure Questions on Final Exam 28
Barchart 1: Student Perceptions - DBL Usability 30
Barchart 2: Student Perceptions - DBL vs. Other Supplementary Materials 31
Barchart 3: Student Perceptions - Negative, Neutral, and Positive 31
ACKNOWLEDGEMENTS

I am grateful to Dr. Ken Plummer, creator of Decision Based Learning, for changing the way I pursue teaching and learning. For the better part of two years he mentored and guided the development of the decision model and implementation. I’m grateful to Dr. Lynne Nielsen for providing access to the introductory statistics course teaching assistants and teaching materials that contributed to the design of content and the decision model. I’m grateful for the dozens of hours volunteered by Dr. Bryan Bradley, Associate Director of the BYU Center for Teaching & Learning, to guide the development of the decision model. I’m grateful to my classmates Shane Huang, Mckay Gerratt, Sam Tyler, and Jacob Andros for the dozens of hours they each put into the decision model, problem bank, and learning module development.

I’m grateful to the Honors Program for pushing me to contribute faculty mentored research. Thanks to Hunter Hallows for introducing me to the Honors Program. My experiences with this thesis have been a pivotal part of my personal and professional development.
INTRODUCTION

The interpretation and conclusion of a research study depend on the use of appropriate statistical procedures (Mishra, Pandey, Singh, Keshri, And Sabaretnam, 2019). Inappropriately matching a statistical method to the analysis of a dataset risks “the loss of the scientific value of such research results and the loss of informativity” (Khusainova, Shilova, Curteva, 2016). While statistical computation has become increasingly automated with the development of computing, statistical method selection remains difficult, such that it “is a common phenomenon in the published articles in biomedical research” (Mishra, Pandey, Singh, Keshri, And Sabaretnam, 2019). Additionally, research from the American Society for Engineering Education Annual Conference and Exposition & Scales, Petlick, the selection of a statistical test as a “frequent” problem in the research process.

Teaching “statistical thinking” at the introductory education level is complicated by enrollment “from a variety of curricula” causing the examples and content [to not be] specific to individual needs (American Society for Engineering Education Annual Conference & Exposition & Scales, Petlick, 2004). Introductory statistics has “developed a reputation for being difficult, mechanical, and boring” as it “primarily [teaches] the technical aspects of using the statistical formulas” (American Society for Engineering Education Annual Conference & Exposition & Scales, Petlick, 2004). Put another way, “procedural steps too often claim students’ attention that an effective teacher could otherwise direct toward concepts” (GAISE College Report ASA Revision Committee,
2016) and functional skills. University students enrolled in introductory education statistics, therefore, struggle to conceptualize the various statistical tools and procedures.

**Procedure Selection and Conditional Knowledge**

To shift focus from the narrow procedural steps towards the broader conceptual construction, students must (1) identify relevant problem features, (2) associate domain-specific concepts with the features, (3) select an equation or procedure appropriate for the concept.

**Feature Identification**

Students look for clues revealing the aims (Jaykaran, 2010) and objectives (Mishra, Pandey, Singh, Keshri, And Sabaretnam, 2019) of the study. Students will take note of hypotheses or assumptions that guided the implementation of the research study. Identifying the groups of interest (Khusainova, Shilova, Curteva, 2016) improves the ability of the student to articulate the nature of comparison or investigation in the research study (Jaykaran, 2010). This investigative effort is meant to be foundational to the spirit of the statistics training (Wild, 1994).

**Domain Specific Concepts**

After identifying features within a research problem, students need to connect the features to domain specific concepts. Students struggle identifying an appropriate statistical method because “there is more to it than meets the eye; the form of the data, sample size, sample distribution, test power, and test robustness are all part of the
equation of test selection” (American Society for Engineering Education Annual Conference & Exposition & Scales, Petlick, 2004). Students connect groups of interest to domain specific concepts of data types (means, proportions, etc.) and data distributions (Mishra, Pandey, Singh, Keshri, And Sabaretnam, 2019). Students must connect statistical objectives of statistical design such as between group design, within subject design, pre/post, etc (American Society for Engineering Education Annual Conference & Exposition & Scales, Petlick, 2004).

**Method Selection**

Educators and researchers such as Khusainova, Shilova, Curteva, and Jaykaran have developed algorithms for statistical method selection. While these algorithms can help guide the feature identification and domain specific concept exploration steps, the authors disclose that there are still many open questions surrounding the goal of choosing the most efficient statistical method (Khusainova, Shilova, Curteva, 2016). Incrementally choosing the appropriate statistical method, therefore, requires the gradual acquittal of experience through “familiarization with the individual elements of statistical methods, the formation of a systemic knowledge about statistical methods, and conscious application of the methods in different situations” (Khusainova, Shilova, Curteva, 2016).

**Role of DBL**

In this study we introduce a decision-based learning approach for introductory education statistics students to improve students’ conditional knowledge, thus facilitating their success at selecting statistical methods. Decision-based learning (DBL) is a relatively
new and emerging teaching pedagogy that systematically targets conditional knowledge as a first-order learning activity (Plummer, Swan, & Lish, 2017). DBL students learn concepts and procedures relevant to solving meaningful problems. Initial empirical studies have provided some evidence of the potential of DBL in terms of student positive impressions and improved performance (Sansom, Suh, & Plummer, 2019).

DBL models teach students to identify the features of a problem, link the features to important concepts, and ultimately select an appropriate problem-solving procedure. As students use the model, they internalize key concepts. We will illustrate this process in the section about Teaching the DBL Model. Our DBL model organizes concepts needed to select an appropriate statistical method for introductory education statistics.

**LITERATURE REVIEW**

Educational psychologists have noted that strategic or conditional knowledge is the knowledge that distinguishes novices from experts (Sugiharto, Corebima, Susilo, & Ibrohim, 2018). It is the knowledge that controls other knowledge types like conceptual (i.e., the why) and procedural knowledge (i.e., the how to) (Turns & Van Meter, 2011). However, conditional knowledge (i.e., the when) is developed tacitly in the workplace rather than in academic settings where the focus is solely on conceptual and procedural knowledge (Walsh 2007; Biggs 2011). Bransford, Cocking, and Brown (1999) in their meta-summary of expertise note that conditional knowledge is not taught systematically in education in introductory. This is supported by Gobet (2005) as well as Zhu et al. (1996) and more recently by Raymond (2019) in their studies. DBL is a pedagogy
originally conceptualized as a teaching method designed to explicitly target conditional knowledge (Plummer, 2017).

**RESEARCH QUESTIONS**

We explore the implications of using a DBL model in an undergraduate introductory statistics course by comparing the exam performance (exam 3 and final exam) of students who utilized the DBL model and students that did not use the model. Due to COVID-19 extenuating circumstances, all students enrolled in introductory statistics during this semester took the course online. We randomly assigned half of the course sections between the DBL and the other half in non-DBL groups. Student perceptions of the DBL model were collected at the end of the semester.

We address the following two research questions:

1. Does the use of a DBL model improve student exam performance in introductory statistics?
2. How do students perceive the DBL model?
METHODS

Model Development

We created a decision-based learning model for statistical method selection. To create the model, we selected several problems from homework assignments and in-class examples that assess a student’s ability to identify the experimental design. Box 1 depicts a typical problem, which we will use to demonstrate the way the model was developed.

Box 1: Statistical Procedure Selection Example Question

1. A certain brand of fishing line claims to have an average breaking strength of 30 pounds. A group of fishermen become angry because this brand of line seems to break so easily and test 25 randomly selected lines of this brand. The mean breaking strength is 27.994 pounds with a standard deviation of 0.846 pounds. A plot of the data follows. Do these data provide sufficient evidence for the fishermen to conclude that the average breaking strength is less than claimed?

A course teaching assistant was shown the problem and was invited to think aloud about all the decisions that would be made to determine a strategy to solve the problem. Talking through the teaching assistant’s decision-making process allowed for the design of a decision model that emphasizes conditional knowledge, structuring conceptual and procedural knowledge around a series of decisions.

Upon seeing the problem in Box 1, the teaching assistant knew tacitly to use a one sample t hypothesis test for solving this problem. Upon further questioning, the teaching assistant articulated that this was due to the researcher’s effort to counter (alternate
hypothesis) the brand’s claim (null hypothesis). Thus, the first decision in the DBL model became identifying hypothesis tests vs confidence intervals. As this process of looking at typical problems and thinking aloud about the decisions necessary to solve the problem continued, an initial model was constructed.

To test the versatility of the model to handle various problem types, I interviewed students from past semesters of the introductory statistics course. The students were presented with several method selection problems and asked to think aloud while solving each problem, first on their own, and second with the model as a guide. This provided valuable information about parts of the model that were easier or more challenging for students to use and led to slight modifications of the model. Three of the students from past semesters, all of which are statistics majors, were so interested in the potential of the decision model, they volunteered to join the project and contribute their experiences and perspective on the creation of learning materials and problem scenarios.

The team of statistics majors, a former course instructor, and an instructional designer from the university’s center for teaching and learning met more than a dozen times to discuss the construction of the decision model. The presentation of procedure selection as shown in the course’s lecture materials was used as reference material. The team grappled with the varying procedural selection approaches they individually internalized and the differing presentations of the process found within the course lecture materials and the course equation sheet. Through this discourse, I documented the breaking points in the model and ensured that our next discussion centered around resolving our past
concerns. We created multiple versions of the decision model but failed to clearly structure a path to the one sample statistical procedures. The current course instructor was integral to providing perspective to ensure our model would be clearly aligned with the course curriculum and student experience.

The DBL model depicted in Figure 1 eventually emerged. We acknowledge that this model represents the decision tree created by a team of instructional designers and teaching assistants, with input from students guiding its development. If another statistics instructor were to develop a DBL model, it might be slightly different. For example, this model is limited to problems related to method selection, excluding nonparametric methods. A different instructor might choose to focus on or emphasize different topics in their decision tree.

Figure 1: Decision Model for Statistical Method Selection
Setting and Participants

The participants in the study were students in a first semester introductory statistics course for all majors, enrolled at Brigham Young University. Generally, the most common majors of students enrolled in introductory statistics are from the college of life sciences and nursing, followed by business school students, followed by the college of physical and mathematical sciences, followed by undeclared majors. For this study, we randomly assigned eight sections of the course to a treatment group and eight sections to a comparison group. The performance on questions relating to statistical method selection, on exam 3 and the final exam, of students in the Winter 2021 (treatment) cohort (N = 1021) were compared to the students in the Winter 2021 (comparison) cohort (N = 930). We have reason to believe the cohorts are comparable because all students participated in the same self-study, online learning environment with the only difference being the DBL modules for the treatment group.

Problem Bank Development

Using procedure selection problem examples from credit quizzes and practice exams, I documented the key features (seen below) that students would likely encounter in examination.
Working with a team of three volunteer statistics majors and an instructional designer from the university’s center for teaching and learning, I delegated the creation of problem examples across all statistical procedures. Using the problem template (above), each team member created at least five unique problem bank questions for each statistical method to fit within the context of five different problem scenarios. Each of the problems were reviewed by two of the other team members and an instructor of the course to ensure alignment with the class curriculum.

Problem Scenarios:

- Corn Farming (one sample t hypothesis testing, two sample t hypothesis testing, etc.)
- Anxiety Medication (one sample t hypothesis testing, two sample t hypothesis testing, etc.)
- Seat Belt Safety Testing (one sample t hypothesis testing, two sample t hypothesis testing, etc.)
- Exercise Heart Rate Study (one sample t hypothesis testing, two sample t hypothesis testing, etc.)
- Tree Pesticides (one sample t hypothesis testing, two sample t hypothesis testing, etc.)

The problem scenarios were intended to show students the problem features in the context of a story that influence procedure selection. For example, conceptualize the differences between two sample t hypothesis testing and one sample t hypothesis testing by seeing the differences (mean of two species of corn versus mean of one species of corn) in the context of the corn farmer. Students worked through examples of all the statistical procedures on practice exam 3 through the lens of the problem scenario of the anxiety medication. Similarly, the practice final presented an example of each statistical procedure through the lens of corn farming.

Learning Module Development

The team of volunteer statistic majors and an instructional designer from the university’s center for teaching worked through the first decision point (hypothesis test vs. confidence intervals) together on multiple problem examples. I documented the clues (problem features) used by the team members to move through the first decision point. Using course lecture materials, I created a learning module template (see below) with four sections: definitions, examples, summary, and practice.
Figure 3: Learning Module Definitions Section Example

Figure 4: Learning Module Example Problem Section Example
Figure 5: Learning Module Summary Section Example

In summary again

**Hypothesis testing** uses data from a sample to **assess a claim or comparison about a parameter.**

**Confidence interval estimation** uses data from a sample to **estimate a population parameter.**

**Key Terms**
- Equal to, less than, greater than, difference
- Level of significance ($\alpha$)
- Perform or carry out

**Remember**
- We are testing a **claim** (fail to reject or reject)
- 2-sample/matched pairs/ANOVA: Testing the equality of means or proportions (comparison)

**Key Terms**
- Interval, range, margin of error,
- Level of confidence, estimate
- Construct or calculate

**Remember**
- We are estimating a likely range of values for the parameter
- There is no prior claim of the population parameter (mean or proportion)

Figure 6: Learning Module Practice Section Example

Practice #1 of 3

Netflix would like to determine the true proportion of their current subscribers who would pay extra for a premium membership including access to more movies and TV shows. They sample 50 customers and found that 45 would pay extra.

Which procedure type should be used?
- Hypothesis test
- Confidence interval
Each team member designed two to three of the learning modules. Each of the 16 learning modules were reviewed by two of the other team members and an instructor of the course to ensure alignment with the class curriculum.

Figure 7: Learning Module Review List

<table>
<thead>
<tr>
<th>Name</th>
<th>Link</th>
</tr>
</thead>
<tbody>
<tr>
<td>How would a statistician approach this research question?</td>
<td>link</td>
</tr>
<tr>
<td>What type of procedure is most appropriate for this question?</td>
<td>link</td>
</tr>
<tr>
<td>Is this study examining a distribution of one variable or a relationship between two variables? Hypothesis Test</td>
<td>link</td>
</tr>
<tr>
<td>Is this study examining a distribution of one variable or a relationship between two variables? Confidence Interval</td>
<td>link</td>
</tr>
<tr>
<td>Is the response variable categorical or quantitative in this problem? 2 var Hypothesis Test</td>
<td>link</td>
</tr>
<tr>
<td>Is the response variable categorical or quantitative in this problem? 2 var Confidence Interval</td>
<td>link</td>
</tr>
<tr>
<td>Is the variable categorical or quantitative in this problem? 1 variable Hypothesis Test</td>
<td>link</td>
</tr>
<tr>
<td>Is the variable categorical or quantitative in this problem? 1 variable</td>
<td>link</td>
</tr>
</tbody>
</table>
Presentation of DBL

Students first interacted with DBL after learning the single sample statistical methods introduced in the course. The course instructor determined that DBL would be integrated in six practice quizzes, six credit quizzes, practice exam #3, and the practice final exam:

Quiz 27: Role-Type classification; Exploratory Data Analysis for Categorical to Quantitative Relationships

Methods Learned:
- One sample t-confidence interval for means (Lesson 18)
- One sample t-test for means (Lesson 21)
- One sample z-confidence interval for proportions (Lesson 25)
- One sample z-test for proportions (Lesson 26)

Quiz 29: Two-sample t procedures

Methods Learned:
- Matched pairs t-test and confidence interval (Lesson 28)
- Two sample t-test for means and confidence interval (Lesson 29)
- All previous methods

Quiz 31: Exploratory Data Analysis for Categorical-to-Categorical relationships; Two-way tables and conditional distributions

Methods Learned:
- ANOVA (Lesson 30)
- All previous methods
Quiz 34: Exploratory Data Analysis for Quantitative-to-Quantitative relationships:

Scatterplots and Correlation analysis

Methods Learned:
- Two sample z-test for proportions and confidence interval (Lesson 32)
- Chi-square test for independence (Lesson 33)
- All previous methods

Quiz 36: Cautions in correlation and regression analyses

Methods Learned:
- All previous methods

Quiz 38: Inference for regression predictions – Confidence Interval and Prediction Interval

Methods Learned:
- Linear Regression (Lesson 37 & 38)
- All previous methods

Practice exam 3 - seven questions based on an anxiety assessment scenario

- One sample t-confidence interval for means (Lesson 18)
- One sample t-test for means (Lesson 21)
- One sample z-confidence interval for proportions (Lesson 25)
- One sample z-test for proportions (Lesson 26)
- Matched pairs t-test (Lesson 28)
- Two sample t-test for means (Lesson 29)
- ANOVA (Lesson 30)
- Two sample z-test for proportions (Lesson 32)
- Chi-square test for independence (Lesson 33)

Practice final - eight questions based on a corn yield scenario
- Linear Regression (Lesson 37 & 38)
- All previous methods on practice exam 3

The problem bank, created by the DBL team, was used for a total of 12 questions on statistical model selection in the practice quizzes, 14 in credit quizzes, and an additional 15 in practice exams. The students in the treatment group used the decision model to work through the same questions as the comparison group. All credit quizzes and practice quizzes were delivered using an online learning management system (Canvas). I added the problems in their entirety to each of the online assignments for the control group. For the DBL group, a link to the DBL software was provided which required that they work through the decision model in order to see the problem scenario for the quiz. After working through the decision model, they would return to the learning management system to input their answer. An online administrator of the learning management system ensured that the DBL/control versions of quizzes and practice exams were made available to the correct course sections.

The DBL group received just-enough, just-in-time instruction through the decision tree, explaining the conceptual foundation for each decision. Students could self-select this instruction at any decision point by accessing the learning module. Working through each decision point, students collect the features and understanding that are
foundational to conditional knowledge: the how-to-decide-what-to-do-and when-to-do-it knowledge (McCormick, 1997). The last decision point in the model, see Figure 3, invites students to synthesize the conditions and concepts that are associated with a specific statistical procedure.

Figure 8: Synthesis point learning module (part 1)

Example: A research clinic is interested in determining the impact of a new anti-obessive compulsive medication on patients with a certain type of OCD. It is known that their patients with this type of OCD without any medication have an average OCD score of 52.5 (scale of 0-50). A sample of 45 participants with OCD participate in the study and their mean OCD score is 46.5.

Figure 9: Synthesis point learning module (part 2)
The model requires students to not only choose the correct procedure, but also choose the correct path to get there. Students cannot move through the model unless they understand key concepts and ideas. Each decision in the model must link to the concepts that inform that decision. Working through the example scenario in box 1, a student must incrementally choose the concepts that lead to the use of a one sample t hypothesis test.

Working through the decision model, students were given feedback on their decisions.

**Correct Path (Correct End Point & Correct Choices):**

Hypothesis Test -> examining the distribution of one variable -> quantitative variable -> one sample t test for means

**Possible Incorrect Path (Correct End Point With Incorrect Choice):**

Hypothesis Test -> examining the distribution of one variable -> categorical variable -> one sample t test for means

Table 1: Feedback Plan for Quizzes & Exams

<table>
<thead>
<tr>
<th>Quiz or Exam</th>
<th>Immediate Software Feedback</th>
<th>Annotated answers with highlighted procedures</th>
<th>Delayed Software Feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>Practice Quiz #27</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Credit Quiz #27</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Practice Quiz #29, 31, 34, 36, 38</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Credit Quiz #29, 31, 34, 36, 38</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Practice Exam #3</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Final Practice Exam</td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>
**Immediate Software Feedback**

After every incorrect choice, the software immediately warns the student that they have left the correct path needed to choose the appropriate statistical method.

**Annotated Answers**

After making a decision in the model, students are shown the statistical methods that are no longer appropriate based on their decisions up to that point. Additionally, an expert explanation for each step in the decision model can be studied to learn the features and concepts that guide an expert’s schema.

**Delayed Software Feedback**

The software warns the student they have left the correct path once they have reached the end of the model and selected a procedure. The students can see at which decision point they strayed from the correct path but are not shown the correct sequence of choices. They must proceed and determine the correct answers with help from the learning modules.

**Student Performance Data and Analysis**

The third midterm exam contained 13 questions related to statistical procedure selection. Each student answered different versions of the exam which displayed 3 to 4 questions from the bank of 13 questions related to procedure selection. The final exam contained 11 questions related to statistical procedure selection. Each student answered different versions of the exam which displayed 3 to 4 questions from the bank of 11 questions related to procedure selection. For each question, we calculated the score proportion of students that answered the question correctly for both the treatment and comparison
groups. We used these scores to compare the performance of the two groups with one another.

To investigate the effect of using the DBL model on exam performance, we used a two-sample z test for proportions to compare the percentage of correct answers between the treatment and comparison groups.

**Student Perceptions Data and Analysis**

The survey was administered as part of the semester course review in the form of an online survey on the university’s learning management system. In this survey, students were asked to rate the influence of technical issues on their DBL experience and the level that course resources (online textbook, r shiny applet, and DBL) enhanced their learning. 76% of the 1021 students in the treatment group answered the survey.
LIMITATIONS

The design team did not have access to past exam questions covering statistical method selection. The statistical method selection questions written for this study -- for use on the credit quizzes, practice quizzes, and practice exams – were reviewed and approved by the introductory statistics course manager. The format and design of the questions, however, may not fully align with the statistical method selection questions on exam 3 and the final exam.

Students were given a limited number of repetitions to learn the 14 statistical methods. While practice quizzes gave students 12 problems examples, complete with annotated answers and immediate feedback, and the practice exams an additional 15 problem examples, both were optional to complete. The graded credit quizzes gave students exposure to 14 problem examples spread out over multiple weeks of instruction.

Additionally, the analysis of exam results and student perceptions of DBL were collected with permission from the statistics course coordinator. I was not provided access to the raw data and therefore was unable to perform additional analysis into the demographics of the students within the course sections that were assigned to the treatment group and control group. The two sample z test for proportions included in the results section are considered default analyses. Future analyses should consider experiment design that allows for blocking (age, ACT score, etc.) beyond randomly dividing sections between the treatment and control.
RESULTS AND DISCUSSION

There was insufficient evidence to conclude a difference on Exam 3 and the Final Exam between the treatment group and comparison groups on the “choose the correct procedure” questions. These findings are supported by the results of the two-sample z test for proportions for exam three [p = .1402] and the final exam [p = .3979].

Table 2: Summary of Treatment Group vs Comparison Group Performance on Exam 3

<table>
<thead>
<tr>
<th>Group</th>
<th>Mean</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comparison</td>
<td>2.5</td>
<td>0.833</td>
</tr>
<tr>
<td>Treatment</td>
<td>2.46</td>
<td>0.819</td>
</tr>
</tbody>
</table>

Mean = average # of points scored on all procedure questions. Questions are out of 3 points, so students either scored a “3” or a “0” on each question.

Proportion = average percentage of students who correctly answered the procedure questions
Boxplot 1: Treatment Group vs Comparison Group Performance on Exam 3

Graph 1: Treatment Group vs Comparison Group Performance on 13 Individual Questions
Table 3: 13 Procedure Questions on Exam 3

<table>
<thead>
<tr>
<th>Question</th>
<th>DBL?</th>
<th>Number Correct</th>
<th>Total</th>
<th>Prop</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>No</td>
<td>165</td>
<td>197</td>
<td>0.838</td>
<td>2.5127</td>
</tr>
<tr>
<td>1</td>
<td>Yes</td>
<td>206</td>
<td>247</td>
<td>0.834</td>
<td>2.502</td>
</tr>
<tr>
<td>2</td>
<td>No</td>
<td>213</td>
<td>269</td>
<td>0.792</td>
<td>2.3755</td>
</tr>
<tr>
<td>2</td>
<td>Yes</td>
<td>248</td>
<td>312</td>
<td>0.795</td>
<td>2.3846</td>
</tr>
<tr>
<td>3</td>
<td>No</td>
<td>262</td>
<td>284</td>
<td>0.923</td>
<td>2.7676</td>
</tr>
<tr>
<td>3</td>
<td>Yes</td>
<td>288</td>
<td>324</td>
<td>0.889</td>
<td>2.6667</td>
</tr>
<tr>
<td>4</td>
<td>No</td>
<td>236</td>
<td>295</td>
<td>0.8</td>
<td>2.4</td>
</tr>
<tr>
<td>4</td>
<td>Yes</td>
<td>245</td>
<td>322</td>
<td>0.761</td>
<td>2.2826</td>
</tr>
<tr>
<td>5</td>
<td>No</td>
<td>161</td>
<td>212</td>
<td>0.759</td>
<td>2.2783</td>
</tr>
<tr>
<td>5</td>
<td>Yes</td>
<td>167</td>
<td>229</td>
<td>0.729</td>
<td>2.1878</td>
</tr>
<tr>
<td>6</td>
<td>No</td>
<td>212</td>
<td>263</td>
<td>0.806</td>
<td>2.4183</td>
</tr>
<tr>
<td>6</td>
<td>Yes</td>
<td>257</td>
<td>315</td>
<td>0.816</td>
<td>2.4476</td>
</tr>
<tr>
<td>7</td>
<td>No</td>
<td>233</td>
<td>263</td>
<td>0.886</td>
<td>2.6578</td>
</tr>
<tr>
<td>7</td>
<td>Yes</td>
<td>273</td>
<td>305</td>
<td>0.895</td>
<td>2.6852</td>
</tr>
<tr>
<td>8</td>
<td>No</td>
<td>151</td>
<td>203</td>
<td>0.744</td>
<td>2.2315</td>
</tr>
<tr>
<td>8</td>
<td>Yes</td>
<td>163</td>
<td>232</td>
<td>0.703</td>
<td>2.1078</td>
</tr>
<tr>
<td>9</td>
<td>No</td>
<td>259</td>
<td>303</td>
<td>0.855</td>
<td>2.5644</td>
</tr>
<tr>
<td>9</td>
<td>Yes</td>
<td>266</td>
<td>307</td>
<td>0.866</td>
<td>2.5993</td>
</tr>
<tr>
<td>10</td>
<td>No</td>
<td>219</td>
<td>265</td>
<td>0.826</td>
<td>2.4792</td>
</tr>
<tr>
<td>10</td>
<td>Yes</td>
<td>268</td>
<td>325</td>
<td>0.825</td>
<td>2.4738</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>11</td>
<td>No</td>
<td>227</td>
<td>275</td>
<td>0.825</td>
<td>2.4764</td>
</tr>
<tr>
<td>11</td>
<td>Yes</td>
<td>228</td>
<td>287</td>
<td>0.794</td>
<td>2.3833</td>
</tr>
<tr>
<td>12</td>
<td>No</td>
<td>191</td>
<td>223</td>
<td>0.857</td>
<td>2.5695</td>
</tr>
<tr>
<td>12</td>
<td>Yes</td>
<td>183</td>
<td>226</td>
<td>0.81</td>
<td>2.4292</td>
</tr>
<tr>
<td>13</td>
<td>No</td>
<td>266</td>
<td>288</td>
<td>0.924</td>
<td>2.7708</td>
</tr>
<tr>
<td>13</td>
<td>Yes</td>
<td>284</td>
<td>305</td>
<td>0.931</td>
<td>2.7934</td>
</tr>
</tbody>
</table>

*Note: Every student got about 3-4 of these questions

*Mean*: Mean amount of points students earned on the given question – Variable is binary (either 3 for correct or 0 for incorrect)

*Total*: Total number of students in either DBL or non-DBL sections that got this question.

*Number Correct*: Number of students who had this question that got it correct.
Table 4: Summary of Treatment Group vs. Comparison Group Performance on Final Exam

<table>
<thead>
<tr>
<th>Group</th>
<th>Mean</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comparison</td>
<td>2.48</td>
<td>0.819</td>
</tr>
<tr>
<td>Treatment</td>
<td>2.46</td>
<td>0.811</td>
</tr>
</tbody>
</table>

Mean = average # of points scored on all procedure questions. Questions are out of 3 points, so students either scored a “3” or a “0” on each question.

Proportion = average percentage of students who correctly answered the procedure questions.

Boxplot 2: Treatment Group vs Comparison Group Performance on Final Exam
Graph 2: Treatment Group vs Comparison Group Performance on 13 Individual Questions

Table 5: 13 different Procedure Questions on Final Exam

<table>
<thead>
<tr>
<th>Question</th>
<th>DBL?</th>
<th>Number Correct</th>
<th>Total</th>
<th>Prop</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>No</td>
<td>246</td>
<td>288</td>
<td>0.854</td>
<td>2.5625</td>
</tr>
<tr>
<td>1</td>
<td>Yes</td>
<td>229</td>
<td>308</td>
<td>0.744</td>
<td>2.2305</td>
</tr>
<tr>
<td>2</td>
<td>No</td>
<td>146</td>
<td>187</td>
<td>0.781</td>
<td>2.3422</td>
</tr>
<tr>
<td>2</td>
<td>Yes</td>
<td>206</td>
<td>252</td>
<td>0.817</td>
<td>2.4524</td>
</tr>
<tr>
<td>3</td>
<td>No</td>
<td>353</td>
<td>424</td>
<td>0.833</td>
<td>2.4976</td>
</tr>
<tr>
<td>3</td>
<td>Yes</td>
<td>393</td>
<td>448</td>
<td>0.877</td>
<td>2.6317</td>
</tr>
<tr>
<td>4</td>
<td>No</td>
<td>235</td>
<td>396</td>
<td>0.593</td>
<td>1.7803</td>
</tr>
<tr>
<td>4</td>
<td>Yes</td>
<td>276</td>
<td>464</td>
<td>0.595</td>
<td>1.7845</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>222</td>
<td>252</td>
<td>264</td>
<td>0.841</td>
<td>2.5227</td>
</tr>
<tr>
<td>6</td>
<td>181</td>
<td>206</td>
<td>202</td>
<td>0.896</td>
<td>2.6881</td>
</tr>
<tr>
<td>7</td>
<td>366</td>
<td>426</td>
<td>406</td>
<td>0.901</td>
<td>2.7044</td>
</tr>
<tr>
<td>8</td>
<td>343</td>
<td>370</td>
<td>434</td>
<td>0.79</td>
<td>2.371</td>
</tr>
<tr>
<td>9</td>
<td>191</td>
<td>212</td>
<td>202</td>
<td>0.901</td>
<td>2.7028</td>
</tr>
<tr>
<td>10</td>
<td>252</td>
<td>271</td>
<td>278</td>
<td>0.906</td>
<td>2.7194</td>
</tr>
<tr>
<td>11</td>
<td>185</td>
<td>183</td>
<td>229</td>
<td>0.808</td>
<td>2.4236</td>
</tr>
</tbody>
</table>

**Mean:** Mean amount of points students earned on the given question – Variable is binary (either 3 for correct or 0 for incorrect)

**Total:** Total number of students in either DBL or non-DBL sections that got this question.

**NumCorrect:** Number of students who had this question that got it correct.
**Student Perceptions**

Survey responses demonstrated students had an adequate understanding of how to use the DBL model, with 93% reporting they did not have technical issues. 11 of the 776 students that responded to the end of semester survey reported that they strongly agreed with having technical issues while using DBL.

**Barchart 1: Student Perceptions - DBL Usability**

In an end of semester survey, students were asked to rate the level that course supplementary resources (online textbook, exploratory data analysis tool, and DBL) enhanced their learning. 76% of the 1021 students in the treatment group answered the survey.
Student perceptions in the charts above are organized by negative (strongly negative and negative), neutral (slightly negative, neither negative nor positive, slightly positive, and no answer), and positive (positive and strongly positive). The general positive sentiment towards DBL’s role in student learning is a promising sign for the continuing role in the introductory stats course moving forward in future semesters.
IMPLICATIONS FOR RESEARCH AND PRACTICE

The results presented here represent a first attempt to measure the impact of a decision-based learning model for method selection on student performance in introductory statistics. We have insufficient evidence to conclude that using the decision model -- within practice quizzes, credit quizzes, and practice exams -- improves student performance on exam questions targeting method selection. The outcome points to several considerations for future instructional practice and statistics education research. Despite the decision model receiving the highest rating (57% agreed or strongly agreed) for learning enhancement as compared to the other class resources listed on the end of semester survey, we recognize that some students didn’t find the model helpful. We should consider a variety of strategies for integrating decision-based models with conceptual and procedural instruction, introducing the model in smaller pieces. As an alternative to expert model development, instructors might also consider allowing students or groups of students to create their own decision models, perhaps with guidance from the instructor or teaching assistants.

It may also be fruitful to have students evaluate a larger number of method selection problems. Students in the treatment group were required to answer 14 method selection problems over the course of the semester as they learned 14 statistical methods. We recommend instructors use a hybrid approach of teaching the decision model in class discussions and within the online portal. Previous efforts to have students practice procedure selection questions inside and outside the model have been promising.
(Plummer et al., 2017). Instructors can also ensure the structure and tone of the problems align with the problems students will be assessed with in examination.

Due to the positive student impressions towards the presentation of the decision model, researchers may consider investigating the impact of DBL on student affective development or retention on the material in a longitudinal study. Other areas of interest include the impact of DBL on student and teacher preparation time.
CONCLUSION

We have presented the results of an exploratory study to investigate the impact of explicit instruction about conditional knowledge on the exam performance of students in introductory statistics. Statistics instructors in an “attempt to cover the technical side of statistics, a clear pattern of how and why individual tests and models should be selected are lost” (American Society for Engineering Education Annual Conference & Exposition & Scales, Petlick, 2004). Statistics instruction often lacks the depth or organization to present conditional knowledge. Decision-based learning (DBL) helps students identify the features of a problem, link the features to important concepts, and ultimately select an appropriate problem-solving procedure. While we did not find sufficient evidence to conclude that DBL significantly improves introductory statistics students ability to select appropriate statistical methods, the positive perceptions of students give hope to future implementation of the DBL pedagogy. The positive effects of DBL on various problem solving skills in other subjects (Plummer, 2017; Sansom, 2019) show an opportunity to improve statistical problem solving worth working towards.
References:


Publication Date:February 8, 2019

https://doi.org/10.1021/acs.jchemed.8b00754


Performance of Treatment vs. Control Group on Credit Quizzes

<table>
<thead>
<tr>
<th>Quiz</th>
<th>Question Number</th>
<th>Prop Correct non-DBL</th>
<th>Prop Correct DBL</th>
</tr>
</thead>
<tbody>
<tr>
<td>27</td>
<td>1</td>
<td>0.531784841</td>
<td>0.846153846</td>
</tr>
<tr>
<td>27</td>
<td>2</td>
<td>0.674816626</td>
<td>0.891089109</td>
</tr>
<tr>
<td>29</td>
<td>7</td>
<td>0.722636816</td>
<td>0.896089385</td>
</tr>
<tr>
<td>31</td>
<td>9</td>
<td>0.901719902</td>
<td>0.939293598</td>
</tr>
<tr>
<td>31</td>
<td>10</td>
<td>0.359163592</td>
<td>0.817275748</td>
</tr>
<tr>
<td>34</td>
<td>6</td>
<td>0.622691293</td>
<td>0.916861827</td>
</tr>
<tr>
<td>34</td>
<td>7</td>
<td>0.561426684</td>
<td>0.901176471</td>
</tr>
<tr>
<td>36</td>
<td>8</td>
<td>0.916775033</td>
<td>0.955916473</td>
</tr>
<tr>
<td>36</td>
<td>9</td>
<td>0.721716515</td>
<td>0.906032483</td>
</tr>
<tr>
<td>36</td>
<td>10</td>
<td>0.898569571</td>
<td>0.954703833</td>
</tr>
<tr>
<td>AVERAGE:</td>
<td></td>
<td>0.691130087</td>
<td>0.902459277</td>
</tr>
</tbody>
</table>
A two-sample z-test for difference in proportions yielded these results:

sample estimates:
   prop 1   prop 2
0.9018384  0.6898213

**p-value < 2.2e-16**

95 percent confidence interval:
0.1999479  0.2240863