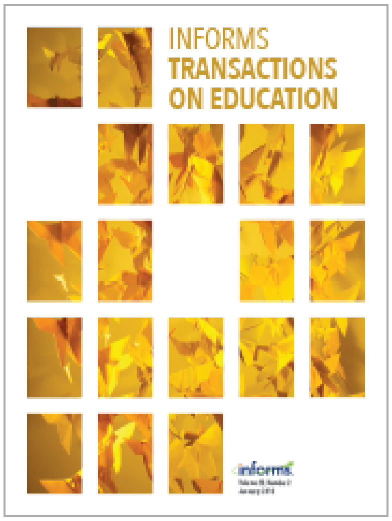


INFORMS Transactions on Education

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

STRATA: A Spreadsheet Tool for Multidimensional Analysis of Operations Research/Management Science Assessment Test Data

Susan W. Palocsay, Scott P. Stevens, Luis J. Novoa

INFORMS TRANSACTIONS ON EDUCATION
Vol. 21, No. 1, September 2020, pp. 18–33
ISSN 1532-0545 (online)
http://pubsonline.informs.org/journal/ited

STRATA: A Spreadsheet Tool for Multidimensional Analysis of Operations Research/Management Science Assessment Test Data

Susan W. Palocsay,[a] Scott P. Stevens,[a] Luis J. Novoa[a]

[a] Department of Computer Information Systems and Business Analytics, College of Business, James Madison University, Harrisonburg, Virginia 22801

Contact: palocssw@jmu.edu, https://orcid.org/0000-0002-2836-4100 (SWP); stevensp@jmu.edu, https://orcid.org/0000-0003-4691-4599 (SPS); novoadlj@jmu.edu (LJN)

Received: June 26, 2019
Revised: December 24, 2019
Accepted: February 28, 2020
Published Online in Articles in Advance: September 9, 2020

https://doi.org/10.1287/ited.2019.0229

Abstract. Accrediting organizations in higher education have shifted emphasis from indirect measures such as student course evaluations, focus-group interviews, and employer surveys to direct measurement of student performance using appropriate assessment instruments. In response, we developed a spreadsheet tool for reports and analysis of test assessments (STRATA) to document and evaluate quantitative data from test questions linked to learning objectives in an introductory undergraduate management science (MS) course. STRATA automatically generates a variety of tables and charts that allow us to examine our students' achievement at three different levels: individual question item, MS topic (subsets of items), and total score. It also tracks test results over time and produces summary reports. In this paper, we discuss STRATA's features, its usefulness in meeting accreditation requirements, and how it can provide faculty with valuable feedback in ways that encourage meaningful conversations about course content and pedagogy.

Supplemental Material: The supplemental materials are available at https://doi.org/10.1287/ited.2019.0229.

Keywords: accreditation issues • assessment • teaching management science • closing the loop • assurance of learning • learning outcomes

1. Introduction

Colleges, notably schools of business and engineering, are under considerable external pressure to monitor and improve their performance, so major accrediting organizations place strong emphasis on the assessment activities of the schools they evaluate. As a result, operations research (OR)/management science (MS) faculty are under increasing demands to participate in assessment activities for their discipline. These activities often include the use of course-embedded exams for direct measurement of students' knowledge and skills on specified learning objectives. Longitudinal data on exam results must then be collected, analyzed, and documented in reports. To meet requirements for closing the loop (Banta and Blaich 2010), these results must be used to guide instructional and curriculum changes for continuous improvement. In this paper, we present an interactive spreadsheet-based tool to support comprehensive analysis of data from assessment tests and to assist in report generation with granularity ranging from individual test questions to course topics to overall scores. The *spreadsheet tool for reports and analysis of test assessments* (STRATA) can be customized to summarize test results by course, instructor, and/or session in a variety of graphs and tables. It will also perform tests for statistical significance to facilitate detailed examination of students' performance at the question item level.

STRATA was developed after several years of analyzing test data manually in Excel. It was designed for accessibility, ease of use, and automation of in-depth comparative analyses of assessment data collected from multiple instructors over time. We also considered using standard statistical software such as SPSS or R. However, their general-purpose nature would require substantial coding to produce the output needed for assessment reporting and user expertise with the software. In contrast, STRATA can be run by anyone with basic Excel knowledge, and its instructions, available functions, tables, and charts are tailored to specific tasks associated with "assurance of learning." Overall, STRATA is a flexible tool with an intuitive user interface for dynamic visualization and analyses of assessment test data. Although our presentation in this article is focused on MS, STRATA can be applied to data from any objective test questions, including true/false and matching formats, as well as standard multiple choice.

This paper is organized as follows. First, we provide a brief overview of the general assessment process for context and background on the motivation for the development of STRATA. This is followed by a summary of relevant literature on the role of direct assessment testing within the process as it pertains to OR/MS courses. Then we describe the key components of STRATA and demonstrate its application in assessing an introductory undergraduate MS course. Using example learning objectives and test questions, we illustrate how STRATA has been used to evaluate student performance and guide improvements on course learning outcomes. Directions for initializing STRATA, assigning questions to topic groups, and importing data are provided as an appendix, in conjunction with an open-source version of the tool. We conclude with a brief discussion of how STRATA tables and charts are used at the aggregate level to form the basis of an assessment report for input to the accreditation process. Finally, we share some remarks on the evolution of our conceptions of assessment through development and implementation of STRATA.

2. Background

Demands for accountability in higher education date back to the 1970s, when government agencies and legislatures began to ask for evidence of student learning to justify budgetary allocations. These mandates initially led to requirements for accrediting organizations to add criteria on educational quality and effectiveness at the institutional level. Subsequently, the focus moved to the program level and resulted in the development of standards for outcome-based assessment by major accrediting bodies in both business (see references in Black and Duhon 2003) and engineering (see references in Olds et al. 2005).

The program-assessment process is generally portrayed as a cycle that starts with a strategic mission statement and determination of broad educational goals. Learning objectives associated with these goals are defined to articulate the knowledge and skills that students are expected to have upon completion of the program. After mapping objectives to the curriculum, appropriate assessments are selected for purposes of measuring student learning. Analysis and interpretation of assessment data then form the basis for planning and implementing future curricular and pedagogic modifications. This final step, often referred to as "closing the loop," feeds assessment results back into the program to continuously improve educational effectiveness.

While there are multiple approaches to assessment (Suskie 2009), both the Association to Advance Collegiate Schools of Business International (AACSB 2013) and the Accreditation Board for Engineering and Technology (ABET 2018) require assessment plans that include direct evidence of student learning. Consequently, most programs employ some form of individual student testing via either a standardized instrument (Ling et al. 2015; see also references in Douglas and Purzer 2015) and/or a locally designed exam (Hadjicharalambous 2014). There are a number of benefits associated with tests of course-related content that are constructed in-house, including tailoring of questions to specific learning objectives, ease of revision, and increased faculty participation in the assessment process (Smith et al. 2010). However, one of the drawbacks is the underlying cost of managing, analyzing, and reporting test results.

We addressed this disadvantage by developing STRATA to facilitate evaluation and documentation of assessment test outcomes for an introductory undergraduate MS course. Our faculty had previously used an iterative process to detail a list of learning objectives and write corresponding test items for the final exam. To meet the AACSB standard for assurance of learning, we designated a senior faculty member as assessment coordinator responsible for managerial tasks associated with collecting exam data, summarizing student scores, and annual report writing. We also implemented a feedback mechanism whereby our coordinator provides us with a detailed analysis of aggregated data at the test-item level and a confidential instructor's report on our students' relative performance. Although we found this feedback to be critical in reviewing our assessment results, it placed a huge administrative burden on the assessment coordinator. STRATA was born primarily in response to the need for automation of the tedious, time-consuming aspects of assessment data analysis and report production.

3. Literature Review

Course grades provide a summary measure of student performance; however, they do not provide sufficient detail on the degree to which students have met learning objectives for accreditation purposes (Banta et al. 2016). Since AACSB's assurance-of-learning standard went into effect in 2003, the Major Field Test for a Bachelor's Degree in Business (MFTB; Educational Testing Service 2016) has become popular for direct measurement and benchmarking of undergraduate competency in the business discipline, primarily as an exit exam. It incorporates questions from nine content areas: accounting, economics, management, quantitative analysis, information systems, finance, marketing, legal and social environments, and international issues. Ling et al. (2015) found evidence of a moderate correlation between MFTB scores and grade-point averages in a synthesis of 24 published studies. Other research has examined the relationship of MFTB performance to a variety of student characteristics, including gender, SAT scores, business major, culture, and motivation (see references in Ketcham et al. 2018).

The quantitative analysis module constitutes approximately 11% (13) of the 120-question MFTB and purports to cover a wide range of probability, statistics, and operations management (OM) topics that include traditional OR/MS techniques such as linear programming (LP) and project scheduling (Educational Testing Service 2016). Subscores on MFTB modules first became available in 2014, and we located two analyses of data from this component. A confirmatory factor analysis of MFBT items by Avery et al. (2014) raised concerns about quantitative analysis being the only one of the nine business content areas with a reported factor loading below the recommended 0.5 threshold. Settlage and Wollscheid (2015) looked at the relationship between ACT scores, major grade-point average, age, gender, major, and total number of courses taken in business subject areas to MFTB subscores. Their results showed that after controlling for demographic variables, the only statistically significant explanatory variable for the quantitative analysis subscore was the number of finance courses.

The need for valid assessment tests in science, technology, engineering, and mathematics (STEM) has also received a great deal of attention (Douglas and Purzer 2015). In this domain, the creation and implementation of concept inventories—standardized tests designed to evaluate students' understanding of core ideas—are prevalent (Evans et al. 2003). A distinction is made between conceptual and factual knowledge for formative assessment of strengths and weaknesses in reasoning about fundamental principles. Concept inventories use a multiple-choice format to facilitate identification of specific student misconceptions, measurement of the effects of pedagogical methods, and analysis of the instrument's psychometric properties (Goldman et al. 2010). The literature indicates that these types of standardized tests have been or are currently under development for a wide range of STEM subject areas (e.g., Allen 2006, Steif and Hansen 2007, Taylor et al. 2014). However, we did not find any concept inventories for industrial or systems engineering that encompass traditional OR/MS topics.

To earn and maintain full accreditation, both AACSB and ABET require detailed, systematic documentation of metrics to measure achievement levels on learning objectives over time. As a result, assessment plans commonly include some combination of course-embedded instruments such as assignments, projects, cases, labs, and exams. When they contribute toward the course grade, these types of classroom assessment tools are inherently motivating to students. By incorporating existing course materials, they are less burdensome to instructors, provide instant feedback, and can be easily modified. Our search for publications that involve the use of assessment methodologies in OR/MS courses uncovered the following works (in chronological order):

- Assessment of student performance demonstrating effectiveness of using a measurable-quantity translation method to formulate LP models (Stevens and Palocsay 2004)
- Pre- and posttesting of both actual and perceived knowledge for MBA students in six competency areas of business statistics (descriptive/graphical analysis, probability, statistical inference) and MS (LP, decision analysis, quality control) (Price and Randall 2008)
- Percentages of correct responses on final exam questions in online business statistics and MS MBA courses to measure application and interpretation of inferential procedures and mathematical formulation and Excel implementation of a network LP model (Sebastianelli and Tamimi 2011)
- Teaching notes describing an MBA assignment integrating regression analysis, simulation modeling, and decision trees for a location decision with sample assessment results (Klingenberg 2012)
- Cluster analysis of test results in business statistics to identify the most difficult learning objectives and those where instruction had a positive impact (Stevens and Palocsay 2012)
- Test instrument with empirical findings for a two-decision business scenario requiring LP model formulation, graphical solution, and written interpretation (Williams et al. 2016).

All these studies identify various performance issues in learning OR/MS-related quantitative subject matter. Our goal is to extend this body of literature in several directions. First, we deliver spreadsheet software designed to reduce the workload associated with capturing, analyzing, and reporting data on question items from any multiple-choice assessment test. Then we highlight STRATA's features and demonstrate its application within a formal assurance-of-learning process for an introductory MS course. We also share a sample of detailed findings from STRATA for MS test questions and discuss how they were used in our assessment process.

4. Design and Functionality of STRATA

We developed STRATA to be a user-friendly interactive software application in the Microsoft Excel environment. It is designed to perform four major functions in support of assessment test data analytics: data importation, question grouping, statistical analyses, and report generation. Execution of these functions is supported by an array of Excel features, including tables, charts, slicers, and macros written in

Visual Basic for Applications (VBA). However, knowledge of VBA programming is not required. STRATA's functionality is available to the user through a collection of five Excel worksheets: Data, Topics, GeneralSummary, Coordinator, and Drilldown. Users can easily navigate between and within these worksheets to investigate assessment test results via VBA buttons and/or worksheet tabs, as described later.

For demonstration purposes, we loaded STRATA with fictionalized data for approximately 3,000 students on an MS assessment test consisting of 63 multiple-choice questions that was administered over seven periods (i.e., semesters) by 10 different instructors. The questions are grouped into eight categories that correspond to the main topics in our introductory MS course. (STRATA reads comma-delimited files from course management systems such as Blackboard, as well as text and Excel files. If multiple versions of the assessment test were constructed by shuffling the order of question items, then we provide a separate utility tool to synchronize student answers to a single test form. Directions for formatting and importing test data and assigning questions to topics are provided in the appendix.)

STRATA opens to the GeneralSummary worksheet with an overview of test results, as shown in Figures 1 and 2. The top section of this report shows the overall success rate for each topic, computed by averaging the percentages of students with correct answers to questions in that category. For example, Figure 1 indicates that student performance was lowest in integer linear programming (ILP) and simulation, with average success rates of only 70.1% on question items in those categories, in comparison with 83.2% on forecasting questions. Figure 1 also provides descriptive statistics and histograms for test scores based on all 63 questions using all of the available data and for only the last two time periods. Although the means and medians are approximately the same here, the median is highlighted on each histogram because the mean is more sensitive to outliers. To summarize test results over time, this worksheet presents the chart in Figure 2, which graphs average test scores and the percentage of students receiving a passing score. The threshold for passing (here 60%) is set by the user when initializing STRATA.

From the GeneralSummary worksheet, users can move to either the Coordinator worksheet, where a variety of detailed reports can be generated, or the Drilldown worksheet for comprehensive test results. Both of these worksheets permit the user to interactively filter data based on instructor and/or academic

Figure 1. GeneralSummary Worksheet (Upper Section)

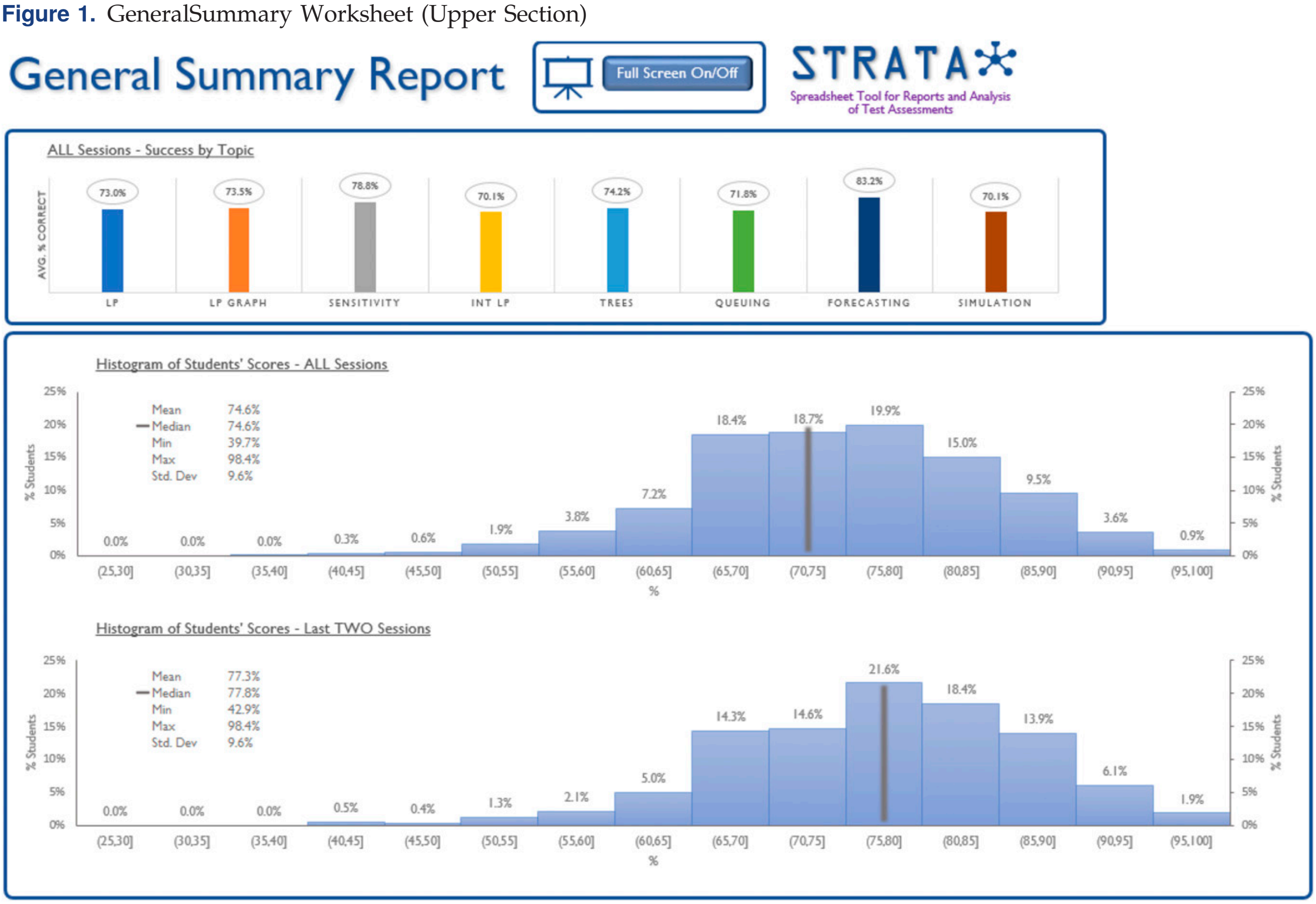
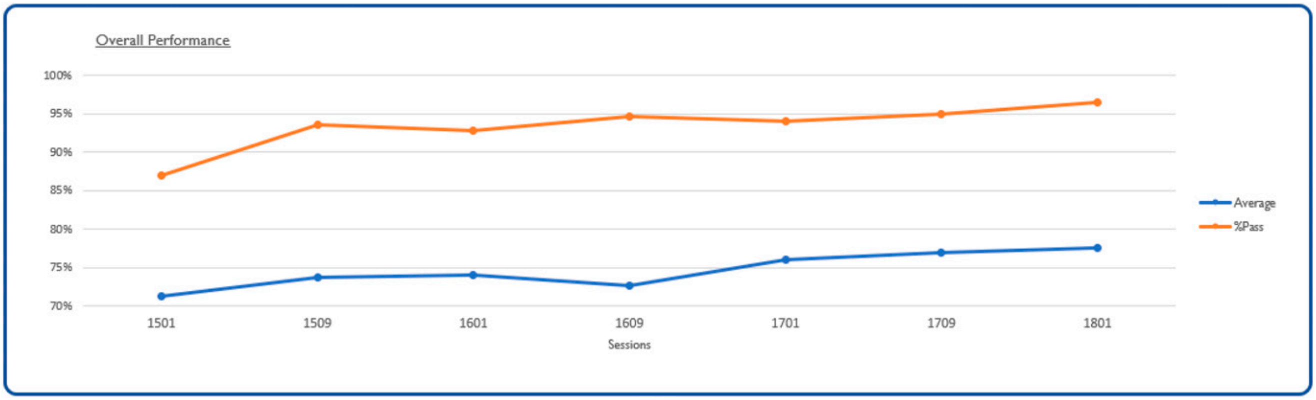


Figure 2. GeneralSummary Worksheet (Lower Section)



session. A bar chart at the top of the Drilldown worksheet shows the overall topic success rates for the sessions and teachers selected (Figure 3). It is followed by three descriptions of results for each test question in the lower portion of the worksheet (Figure 4). Relative frequencies for each answer are shown with correct answers highlighted and horizontal bar charts for visualization of distributions. For instance, it is easy to see that questions 1 and 6 had similar success rates of 63.9% and 66.1%, respectively. However, there was less variation in the answers chosen for question 1, with 21.0% of students incorrectly selecting A, whereas most of the errors on question 6 were split almost evenly between B (15.0%) and D (13.9%). The percentage of students who earned a passing score of 60% or higher on the test is also calculated here. Options are available for viewing questions in numerical order or by performance—from either best to worst or worst to best. Note that the assignment of question items to topic categories can be referenced as needed on the Topics worksheet (Figure 5).

Below the detailed relative frequencies, STRATA computes and displays the success rate and most popular incorrect answers for each question. To identify the most widespread misconceptions, warning signs are displayed for questions with a wrong-answer percentage that exceeds a user-specified threshold. In our fictional data, average success rates were similar on formulation of LP models (73.0%) and solving LPs graphically (73.5%). Yet most students (90.4%) answered question 2 on LP modeling correctly, whereas only 62.2% were successful on question 11 on LP graphing. And 21.9% of students made the same error on this question, indicating a misunderstanding that may be addressable by classroom examples or practice problems. This section of the Drilldown worksheet in Figure 4 helps to immediately flag 9 of the first 15 questions on the assessment test as having specific, identifiable errors above the threshold rate of 11%. To focus on recent test results, the user can turn on multiselection and limit analysis to the recent sessions. Figure 6 demonstrates this, limiting the report to sessions 1709 and 1801.

On the Coordinator worksheet, STRATA runs both descriptive and inferential statistical analyses to support examination of student outcomes on each test

Figure 3. Drilldown Worksheet for All Data (Upper Section)

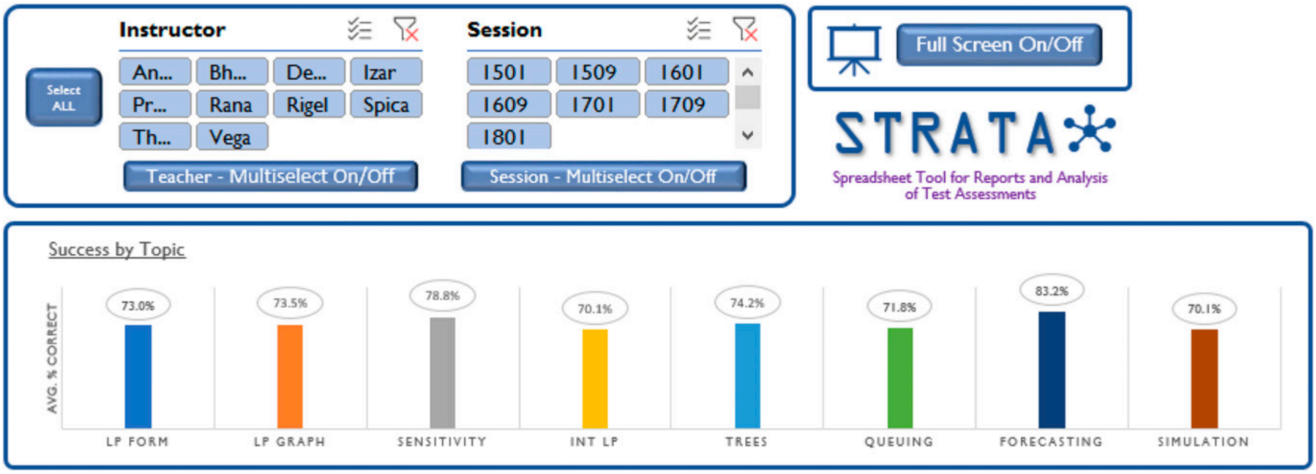
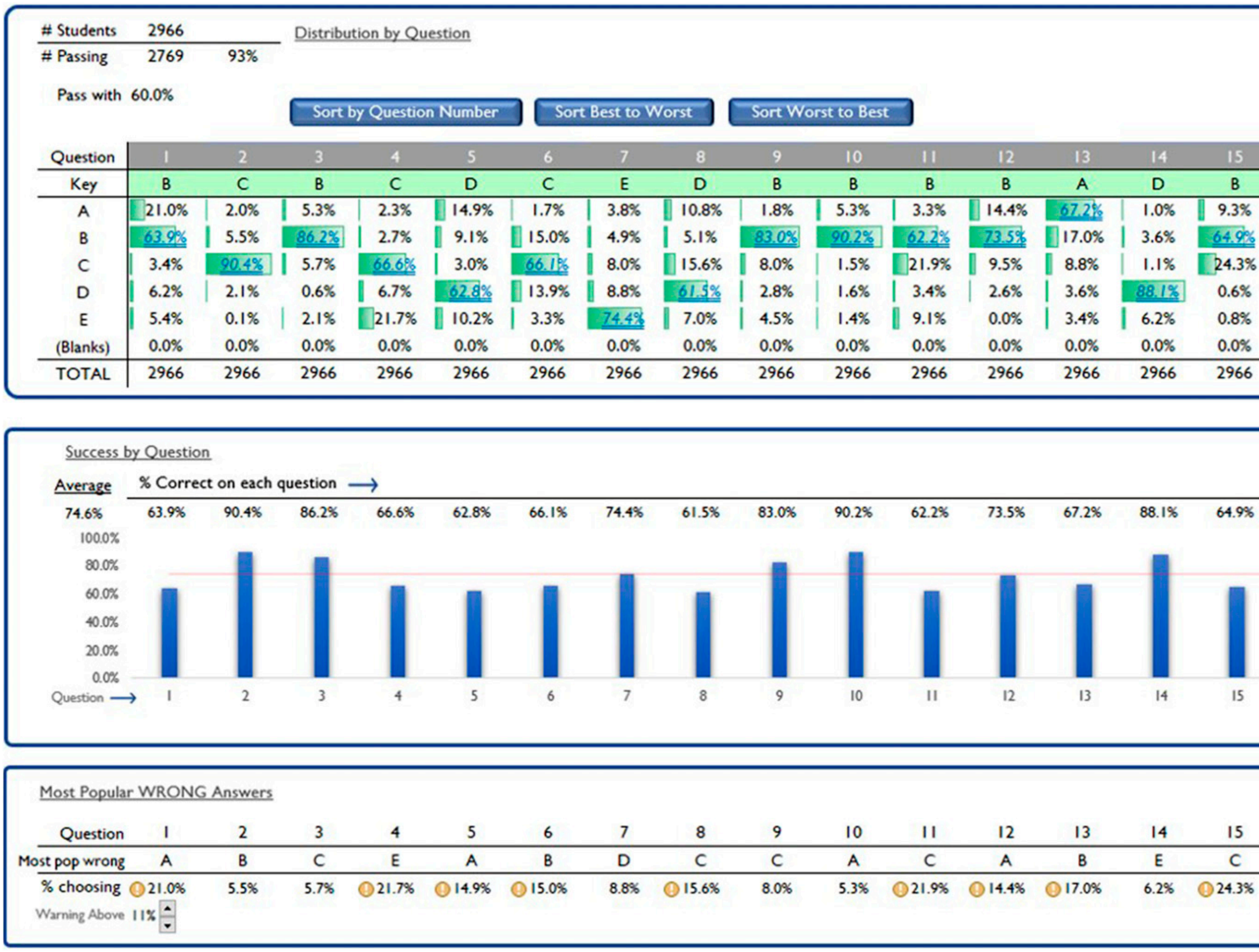


Figure 4. Drilldown Worksheet for All Data (Partial View of Lower Section)



# Students	2966	
# Passing	2769	93%
Pass with	60.0%	

Distribution by Question

Sort by Question Number | Sort Best to Worst | Sort Worst to Best

Question	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Key	B	C	B	C	D	C	E	D	B	B	B	B	A	D	B
A	21.0%	2.0%	5.3%	2.3%	14.9%	1.7%	3.8%	10.8%	1.8%	5.3%	3.3%	14.4%	67.2%	1.0%	9.3%
B	63.9%	5.5%	86.2%	2.7%	9.1%	15.0%	4.9%	5.1%	83.0%	90.2%	62.2%	73.5%	17.0%	3.6%	64.9%
C	3.4%	90.4%	5.7%	66.6%	3.0%	66.1%	8.0%	15.6%	8.0%	1.5%	21.9%	9.5%	8.8%	1.1%	24.3%
D	6.2%	2.1%	0.6%	6.7%	62.8%	13.9%	8.8%	61.5%	2.8%	1.6%	3.4%	2.6%	3.6%	88.1%	0.6%
E	5.4%	0.1%	2.1%	21.7%	10.2%	3.3%	74.4%	7.0%	4.5%	1.4%	9.1%	0.0%	3.4%	6.2%	0.8%
(Blanks)	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
TOTAL	2966	2966	2966	2966	2966	2966	2966	2966	2966	2966	2966	2966	2966	2966	2966

Success by Question

Average	% Correct on each question →														
74.6%	63.9%	90.4%	86.2%	66.6%	62.8%	66.1%	74.4%	61.5%	83.0%	90.2%	62.2%	73.5%	67.2%	88.1%	64.9%

Most Popular WRONG Answers

Question	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Most pop wrong	A	B	C	E	A	B	D	C	C	A	C	A	B	E	C
% choosing	21.0%	5.5%	5.7%	21.7%	14.9%	15.0%	8.8%	15.6%	8.0%	5.3%	21.9%	14.4%	17.0%	6.2%	24.3%

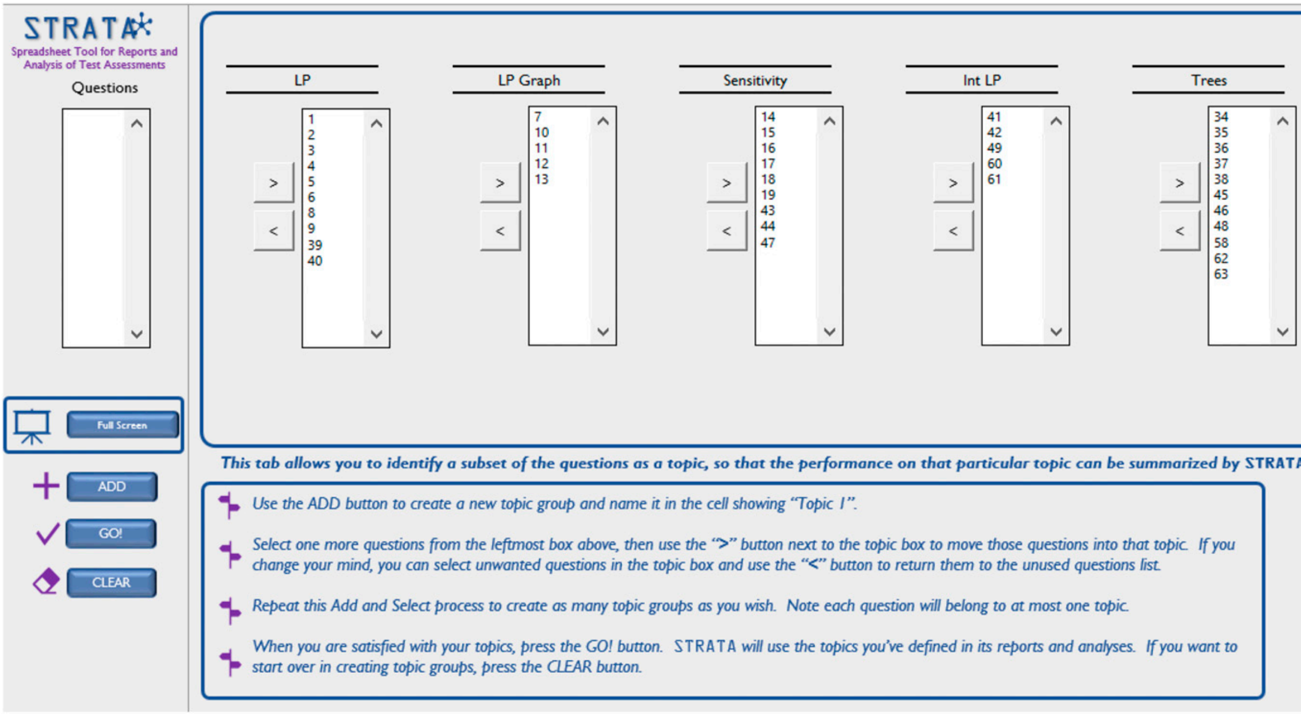
Warning Above 11%

question by instructor. In Figure 7, we see that five instructors taught a total of 848 students in the MS course during the last two semesters, with average test scores ranging from 74% to 81%. For statistical inference, we hypothesize that each instructor's students are a random sample drawn from a population whose success proportion is the same as the overall success proportion on that question for those semesters. Then green- and red-highlighted boxes visually depict statistically significant results (at the 0.05 level) that are above or below average, respectively, for each test item.

This allows the assessment coordinator to quickly identify questions on which everyone's students had about the same success rate. It also distinguishes those questions from test items where one or more instructors' students did particularly well or poorly in comparison with the general population. While coloring is based on a z-test of proportions for each question, there is no correction for repeated use of p-values, so these results should be viewed as more of a flag (green or red) than a formal statistical test. An instructor can track his or her students' performance on a particular question over semesters, and that question item becomes notable only when a color pattern persists over multiple semesters. If a single question is identified as being an instructor's biggest potential problem, though, then the results for that question in the next semester could be used in a single z-test.

Figure 7 demonstrates this for a user who chose to review the results of the two most recent sessions. It shows that question 3 was relatively easy, with 91% of students answering correctly and no significant differences among instructors. However, the overall success rate on question 8 was only 71%, with 86% of Thuban's students answering correctly in comparison with rates of 58% and 57% in Rana's and Bharani's classes, respectively. STRATA also displays the distribution of test scores for each teacher during these two semesters in side-by-side box plots. To put these results in proper context, the histogram of all test scores for those two sessions is repeated from the GeneralSummary worksheet (Figure 1) adjacent to the box plots.

Figure 5. Grouping of Questions into Topic Area



After exploring assessment results from the session perspective, the coordinator can then proceed to the next section of the Coordinator worksheet to generate a detailed summary report for each instructor. As an example, Figure 8 shows a partial screenshot for Spica, who has taught the MS course for four sessions. The average test score for Spica's students was 68% in her first semester with a 73% pass rate. In this session, her students had statistically significant lower performance (red boxes) on multiple test items on LP formulation, LP graphical solution, and queuing. However, compared with other instructors in that time period, Spica's students were stronger (green boxes) in LP sensitivity analysis. Over time, there is evidence of steady improvement in the weaker areas. By the fourth session, Spica has a test average of 77%, with 95% of her students achieving a score of at least 60% and solid or better performance on all the course topic areas. A button at the bottom of this worksheet will execute VBA code to write and save the assessment results for Spica to a new Excel workbook. This file, shown in Figure 9, can then be sent directly to the instructor. Because of the confidential nature of instructor-specific data, the Coordinator worksheet should be limited to use for continuous improvement planning. We discuss the sensitive aspects of this step in the assessment process and how we handle it in our conclusions.

Figure 6. Upper Section of Drilldown Worksheet for Two Most Recent Sessions

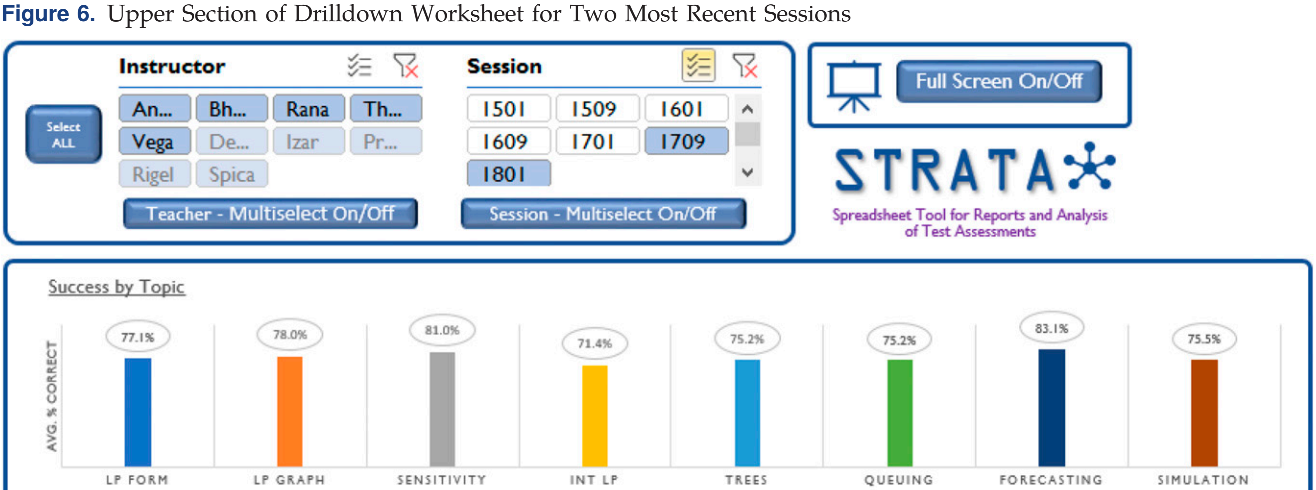


Figure 7. Coordinator Worksheet for the Last Two Sessions

Student Outcomes by Teacher (Session(s) fixed)

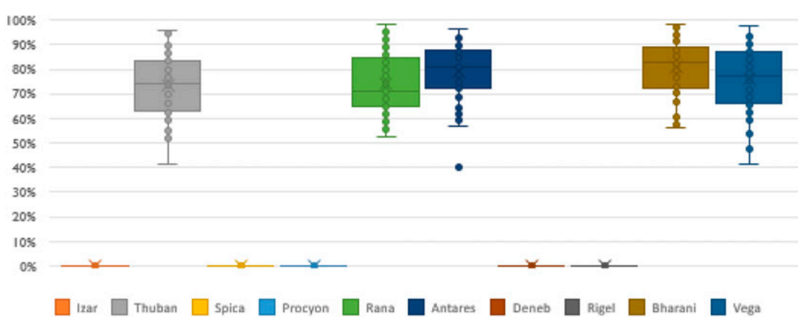
Session: 1501, 1509, 1601, 1609, 1701, 1709, 1801

Significance Level α = 0.05

Session - Multiselect

Instructor	#Students	Average	% Pass	1 B	2 C	3 B	4 C	5 D	6 C	7 E	8 D
Izar	---	---	---	---	---	---	---	---	---	---	---
Thuban	125	74%	94%	72%	94%	92%	74%	76%	60%	76%	86%
Spica	---	---	---	---	---	---	---	---	---	---	---
Procyon	---	---	---	---	---	---	---	---	---	---	---
Rana	65	74%	94%	71%	92%	89%	65%	68%	75%	78%	58%
Antares	186	79%	98%	74%	91%	90%	75%	77%	69%	72%	77%
Deneb	---	---	---	---	---	---	---	---	---	---	---
Rigel	---	---	---	---	---	---	---	---	---	---	---
Bharani	181	81%	97%	82%	97%	94%	69%	61%	70%	85%	57%
Vega	291	76%	95%	54%	94%	90%	74%	70%	63%	79%	73%
Overall	848	---	96%	68%	94%	91%	73%	70%	66%	78%	71%

Student Outcomes by Teacher



5. Example Assessment Results Using STRATA

The MS course is a core requirement at the sophomore level for admission to our college of business and covers traditional OR/MS topics using Excel spreadsheet modeling. Our approach to meeting the AACSB's assurance-of-learning standard in this course is based on a well-defined set of learning objectives with corresponding test items written by senior faculty members. Although outside the scope of this paper, we note that the process for developing these objectives and question items with carefully designed distractors requires substantial time and effort and is critical to obtaining actionable assessment results. We use constructed-response questions for evaluation throughout the semester; however, a multiple-choice format was selected for assessment testing to standardize scoring and provide more precise measurement of student knowledge (Simkin and Kuechler 2005). Both learning objectives and test items undergo regular review and revision by full-time faculty. In this section, we present some examples of how using STRATA to analyze test results has given us a better understanding of student misconceptions, as

Figure 8. Partial Teacher Results on the Coordinator Worksheet

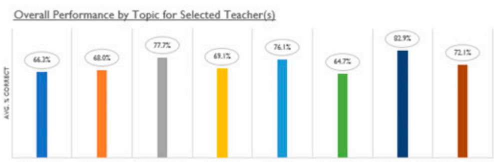
Student Performance over Time (Teacher fixed)

Instruc...: Ant..., Bha..., Den..., Izar, Pro..., Rana, Rigel, Spica, Thu..., Vega

Teacher - Multiselect

Session	#Students	Average	% Pass	1 B	2 C	3 B	4 C	5 D	6 C	7 E	8 D	9 B	10 B	11 B	12 B	13 A	14 D	15 B
1501	41	68%	73%	59%	59%	51%	39%	61%	44%	46%	34%	41%	51%	59%	46%	41%	98%	76%
1509	74	72%	96%	64%	85%	80%	59%	53%	65%	65%	55%	73%	84%	61%	54%	57%	91%	61%
1601	84	75%	92%	43%	94%	87%	58%	43%	65%	81%	68%	77%	88%	74%	69%	83%	87%	62%
1609	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
1701	83	77%	95%	57%	94%	93%	64%	61%	72%	73%	67%	88%	89%	73%	83%	82%	94%	71%
1709	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
1801	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
	Mean	73.6%	%Correct	**63.9%**	**90.4%**	**86.2%**	**66.6%**	**62.8%**	**66.1%**	**74.4%**	**61.5%**	**83.0%**	**90.2%**	**62.2%**	**73.5%**	**67.2%**	**88.1%**	**64.9%**

Overall Performance by Topic for Selected Teacher(s)



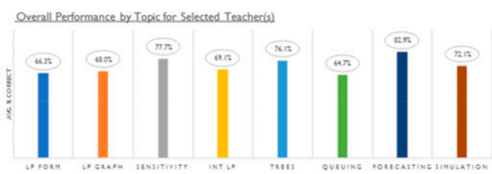
Session	LP Form	LP Graph	Sensitivity	Int LP	Trees	Queuing	Forecasting	Simulation
1501	53.4%	48.8%	81.3%	67.8%	73.6%	53.7%	81.1%	71.5%
1509	67.8%	64.1%	75.5%	70.5%	78.3%	58.6%	82.6%	68.2%
1601	69.0%	79.0%	73.7%	70.7%	77.1%	72.4%	84.1%	72.2%
1609	-	-	-	-	-	-	-	-
1701	74.9%	80.2%	80.2%	67.2%	75.6%	74.3%	83.7%	76.3%
1709	-	-	-	-	-	-	-	-
1801	-	-	-	-	-	-	-	-
Average	**66.3%**	**68.0%**	**77.7%**	**69.1%**	**76.1%**	**64.7%**	**82.9%**	**72.1%**

All Sections	LP Form	LP Graph	Sensitivity
1501	66.7%	67.5%	76.5%
1509	68.6%	68.7%	77.2%
1601	72.6%	70.9%	76.5%
1609	74.1%	74.5%	79.5%
1701	74.2%	76.3%	80.0%
1709	76.0%	77.6%	81.2%
1801	78.2%	78.4%	80.8%
Average	**72.9%**	**73.4%**	**78.8%**

Figure 9. Partial Individual Teacher Report

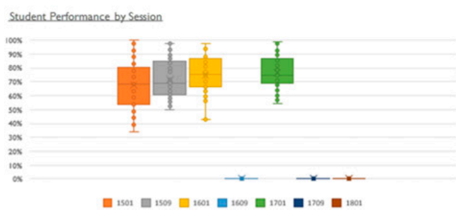
Individual Report

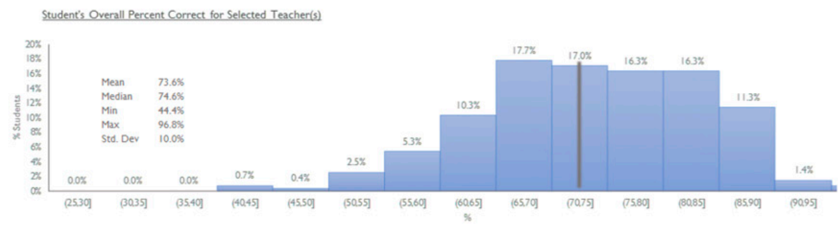
Teacher																				
Spica	Overall	2966	---	93%	64%	90%	86%	67%	63%	66%	74%	61%	83%	90%	62%	73%	67%	88%	65%
				Question	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	Session	#Students	Average	% Pass	B	C	B	C	D	C	E	D	B	B	B	B	A	D	B
	1501	41	68%	73%	59%	59%	51%	39%	61%	44%	46%	34%	41%	51%	59%	46%	41%	98%	76%
	1509	74	72%	96%	64%	85%	80%	59%	53%	65%	65%	55%	73%	84%	61%	54%	57%	91%	61%
	1601	84	75%	92%	43%	94%	87%	58%	43%	65%	81%	68%	77%	88%	74%	69%	83%	87%	62%
	1701	83	77%	95%	57%	94%	93%	64%	61%	72%	73%	67%	88%	89%	73%	83%	82%	94%	71%
		Mean	73.6%	%Correct	**63.9%**	**90.4%**	**86.2%**	**66.6%**	**62.8%**	**66.1%**	**74.4%**	**61.5%**	**83.0%**	**90.2%**	**62.2%**	**73.5%**	**67.2%**	**88.1%**	**64.9%**



Session	LP Form	LP Graph	Sensitivity	Int LP	Trees	Queuing	Forecasting	Simulation
1501	53.4%	48.8%	81.3%	67.8%	73.6%	53.7%	81.1%	71.5%
1509	67.8%	64.1%	75.5%	70.5%	78.3%	58.6%	82.6%	68.2%
1601	69.0%	79.0%	73.7%	70.7%	77.1%	72.4%	84.1%	72.2%
1609	-	-	-	-	-	-	-	-
1701	74.9%	80.2%	80.2%	67.2%	75.6%	74.3%	83.7%	76.3%
1709	-	-	-	-	-	-	-	-
1801	-	-	-	-	-	-	-	-
Average	**66.3%**	**68.0%**	**77.7%**	**69.1%**	**76.1%**	**64.7%**	**82.9%**	**72.1%**

All Sections	LP Form	LP Graph	Sensitivity
1501	66.7%	67.5%	76.5%
1509	68.6%	68.7%	77.2%
1601	72.6%	70.9%	76.5%
1609	74.1%	74.5%	79.5%
1701	74.2%	76.3%	80.0%
1709	76.0%	77.6%	81.2%
1801	78.2%	78.4%	80.8%
Average	**72.9%**	**73.4%**	**78.8%**





well as evidence of our teaching effectiveness and guidance for continuous improvement.

One of our MS learning objectives is for students to be able to graphically solve a linear program with two decision variables. We also have a set of objectives for essential algebra skills, including graphing, solving, and evaluating the numerical values of linear equations. Consider the LP graph and a subset of the questions about it with student outcomes shown in Figure 10: correct answers with success rates are in boldface green, and the most frequent incorrect answer is indicated in red with its percentage. The average success rate on these four questions was 50%. We can immediately see that performance on question 4 was strong and indicated that a majority (91%) of our students could correctly solve two simultaneous equations and evaluate the objective function at this point. However, question 2 only required plugging $x = 40$ $(y = 0)$ or $y = -10$ $(x = 0)$ into the objective function $C = x - 4y$, and yet it had a very low success rate of 27%. We were surprised to see most of our students using *both* intercepts in the objective without thinking that *either* of the two numbers should be used alone to specify its value. This error occurred even

Figure 10. LP Graphing Questions with Correct (Green) and Most Frequent Incorrect (Red) Rates

1. The feasible region for this problem is Region

   a) A **b) B (62%)** c) C d) D (35%) e) E

2. The original OFL shown has an x-intercept of 40 and a y intercept of –10. The value of the objective function C corresponding to this particular OFL is

   a) -80 b) -40 c) 0 **d) 40 (27%)** e) 80 (55%)

3. For this question only, pretend that the feasible region is Region B. Then the optimal point for the given objective function would be at Point

   a) #1 (31%) b) #2 c) #3 d) #4 (46%) e) #5

4. Algebraically find the coordinates of point #3. The value of the objective function, C, at this point is

   a) -87.5 (91%) b) -85 (9%) c) -78.625 d) -50 e) 0 when they could successfully compute the objective function value in question 4. This item continued to show poor discrimination over several sessions between students who otherwise demonstrated solid knowledge of the LP graphical solution procedure and those who did not. It was also duplicating question 4 in evaluating the learning objective on LP objective function value calculation. Following recommended guidelines for writing reliable test items (Kehoe 1994), we subsequently decided to delete question 2, making room for expansion of the assessment test in other areas.

Figure 10 also shows that more than half of students (62%) answered question 1 correctly. The most common error resulted from 35% of students failing to test a point to determine the feasible side of constraint 2 (that is, $2x \leq y$), a line through the origin. It suggested that our students needed to see more examples of constraints with positive slopes. The success rate on question 3 was also poor (31%), with a substantial group of students (46%) understanding that the optimal point occurs where the objective function line moves out of the feasible region but wrongly assuming that it always shifts downward for a minimization linear program. Our assessment coordinator provided an example of a one-product maximization linear program with a limited budget and a demand that depends on advertising spending for classroom use in illustrating a nonintuitive direction of optimization. The summary report from STRATA identified these problem areas, and individual instructor reports helped our faculty to monitor their students' outcomes. Two years later, the success rates for these two items had increased to 67% and 51%, respectively.

Another group of LP-related learning objectives focuses on sensitivity analysis. In this module of the MS course, we emphasize practical applications of what-if analysis to develop students' quantitative reasoning skills. Figure 11 provides a LP model with its corresponding Excel Sensitivity Report and a subset of the assessment test questions on changes

Figure 11. LP Sensitivity Analysis Questions with Correct (Green) and Most Frequent Incorrect (Red) Rates

5. If there were 707 hours of milling time available, rather than 907 hours, then

a) the total cost would increase by $1200.
b) the total cost would not change.
c) the total cost would decrease by $1200. (16%)
d) the total cost would drop by $2000.
e) **we could not predict the new total cost from the Excel report given. (68%)**

6. Suppose that the firm were required to produce at least 300 units, rather than at least 310 units. Then the optimal total cost for the company would be

a) **$19,458. (76%)**
b) $19,668.
c) $20,438. (14%)
d) impossible to say from Excel, since we are outside the appropriate RHS range.
e) impossible to say from Excel, since we are outside the appropriate OFC range.

7. The OFC range on product X has an allowable decrease of 12.375. It follows from this that, if the cost per unit for X dropped to $72, then

a) **the total cost for the company would drop by $80. (44%)**
b) the total cost for the company would remain unchanged. (37%)
c) the company would make more of product X.
d) the company would make less of product X.
e) none of these follow, because the OFC range applies to constraints, not variables.

8. Suppose that product W could be made for free. Then the sensitivity analysis shows that the amount of product W that the firm would produce in its optimal solution would

a) probably increase (shown by RHS ranging on constraint 6).
b) probably increase (shown by OFC ranging on the variable W). (26%)
c) definitely stay the same (shown by RHS ranging on constraint 6).
d) **definitely stay the same (shown by OFC ranging on the variable W). (45%)**
e) sensitivity analysis does not allow you to conclude any of the above.

in objective function coefficients (OFCs) and right-hand-side (RHS) values. Questions 5 and 6 both had relatively solid success rates of 68% and 76% compared with the previous five-year averages of 49% and 51%, respectively. An examination of the most common mistakes informed us that whereas there had been improvement, 16% of the students were still not checking the RHS range before applying the shadow price, and 14% continued to believe that a change within the RHS range of a binding constraint would not affect the optimal cost. We discussed teaching strategies to target these misunderstandings and saw an 11% increase in the success rate of question 5 and modest improvement (3%) on question 6 the following year.

We expected students to have difficulties with RHS range analysis but were surprised by the low success rates of 44% and 45% on questions 7 and 8, respectively. Although students recognized that these were OFC ranging problems, many of them did not fully understand what it means to say that the optimal solution (i.e., values of the decision variables) will remain the same while the optimal objective function value changes. After reviewing course materials, some of our instructors found that they had less class coverage of OFC range analysis because of an assumption that it was conceptually easier. Modifications made the next year led to a significant increase in success on question 7 (up to 61%), although question 8 continued to be problematic.

Our introductory MS course covers queuing models as a prerequisite to business process management in the OM course. Students have already completed a class in introductory statistics with exposure to the Poisson and exponential probability distributions prior to the MS course. Learning objectives for this topic include selecting the appropriate queuing model, determining mean arrival and service rates, converting between times and rates, and interpreting steady-state queuing statistics. Figure 12 describes a queuing scenario for car washing where the number of servers corresponds to the number of teams, with a sample of four questions from our assessment test. Most students (82%) found the correct statistic (W_q) in question 10, with only 5% choosing W instead of W_q. However, in question 9, 8% of the students minimized W_q instead of L_q (69%). There were also errors in the interpretations of system probabilities (question 11) and mean service rate versus mean service time (question 12). Faculty discussions of these results led us to increase the emphasis on applications of queuing statistics over computations. Within two years, average success on these four questions improved from 70% to 77%. The most persistent misconceptions in queuing have been about discrete state probabilities in conjunction with varying numbers of servers.

6. Concluding Remarks

Our MS course coordinator gathers assessment test data each semester for analysis and produces both summary and instructor reports using STRATA. Detailed success rates at the item level are then reviewed by full-time faculty to examine trends, identify strengths and weaknesses in student learning, and look for ways to address common errors. This targets the common stumbling block of closing the loop (Banta and Blaich 2010): applying what has been learned from the assessment process to continually improve the quality of the course. It is important to regularly revisit the content of courses and programs to ensure that they reflect the knowledge and skills that are needed for student success in the marketplace. However, closing the loop also entails improving the learning process itself, which is characterized

Figure 12. Queuing Questions with Correct (Highlighted) and Most Frequent Incorrect Rates

9. Russ argues that it's bad for business when a lot of cars are seen waiting for their washing to begin. He wants to choose a team size that will minimize the average number of cars that are waiting in this way. With his choice, the average number of cars waiting for their washing to begin would be

   a) 0.0111 cars (8%) **b) 0.1111 cars (64%)** c) 0.1333 cars d) 0.1482 cars e) 1.333 cars

10. If the students choose to use two teams of three students each, what is the average time that a car spends waiting for its washing to begin?

    a) 0.000556 min. b) 40 sec. **c) 2 min. (82%)** d) 8 minutes (5%) e) 8 minutes, 10 seconds

11. If the students choose to use two teams of three students each, what fraction of the time will one team be busy washing a car while the other team is waiting for a customer?

    a) About 11% b) About 17 % **c) About 33% (69%)** d) About 50% e) About 67% (15%)

12. If all six students work together on a single car, how long on average, does it take to wash it?

    a) 3 min. (11%) **b) 4 min. (64%)** c) 6 min. d) 10 min. e) 15 min.

by an interaction of faculty, students, and course materials. STRATA facilitates this improvement in several ways.

The academic quality of incoming students is variable and not directly under our control. STRATA assists us in flagging changes and leads to modification of our presentation of course material in light of them. When a concept is consistently troublesome for students, STRATA helps us to see if a new approach shows evidence of being a significant improvement over previous ones. It has aided us in identifying what is relatively easy or difficult for students in general. The information available to individual instructors has allowed them to quickly distinguish between concepts where almost all students do well or poorly and concepts where the performance of their own students is either significantly better or worse than that of their peers. By recognizing areas where focused attention and/or consultation with other faculty may be useful, instructors can substantially improve their teaching effectiveness.

At the end of each academic year, an updated longitudinal report on MS assessment results is sent to the administration. This report includes aggregated descriptive statistics and graphs from STRATA that do not contain instructor-specific information. These data are accompanied by a discussion of how we have applied findings from the previous year's assessment process to continually improve the value added to the core business curriculum by our MS course. As a result, it provides documentation in support of AACSB Standard 8, which requires use of "well-documented, systematic processes for determining and revising degree program learning goals; designing, delivering, and improving degree program curricula to achieve learning goals; and demonstrating that degree program learning goals have been met" (AACSB 2013, p. 32). This MS course report serves as input to the "portfolio of evidence demonstrating achievement of learning goals" (AACSB 2013, p. 33) at the program level for accreditation purposes.

In addition to showing progress in teaching MS topics over time, these reports have documented fall–spring seasonal patterns and negative impacts associated with larger class sizes as well as more teaching by adjunct faculty and instructors without doctoral degrees. Our department head has been able to effectively use these data in support of requests for more tenure-track faculty lines. We also note that use of STRATA is not limited to any specific subject area, and it is currently being applied to analysis of assessment test data for our introductory business statistics course.

The individual STRATA reports have been particularly helpful to our newest instructors who came on board with varying amounts of prior teaching experience. The immediate, objective feedback on specific learning objectives has given them reassurance about those aspects of their teaching that are on track and guidance for making adjustments to their methods and materials where needed. In our assessment program, instructor-level outcomes are kept confidential by the coordinator and only available to the teacher for continuous improvement purposes. This requires faculty ownership of the assessment process and a high degree of trust between faculty and the administration. We recognize that we are fortunate to be in an institution with a culture that places value on the use of assessment to inform pedagogical decision making. However, if privacy cannot be maintained, STRATA can still be used for summary and test-item analyses by loading data anonymously. Furthermore, with its flexible design, STRATA's capabilities can be applied to data from any assessment instrument with multiple-choice items, regardless of course content. We are now in the process of implementing STRATA for assessment of introductory business statistics.

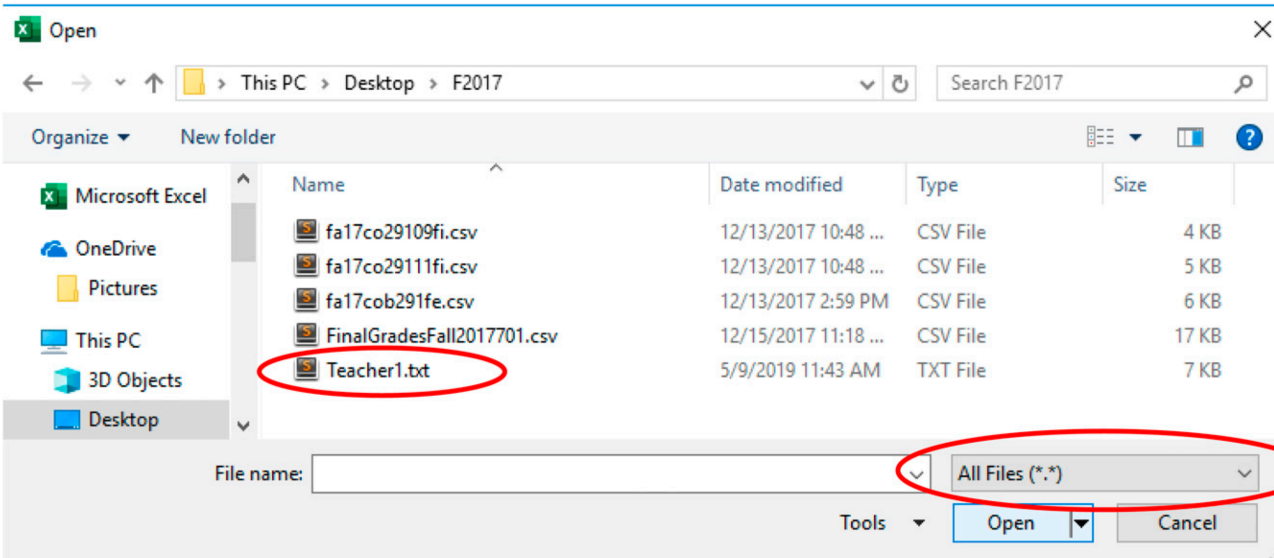
Numerous challenges are associated with the operation of a robust assurance-of-learning program—defining appropriate learning outcomes, developing a plan to measure those outcomes, implementing assessment methods, preparing reports on assessment data, and using assessment results to inform curricula changes and effect learning improvements in accordance with an institution's mission (AACSB 2019). Insufficient resources and a lack of faculty involvement are often cited as major barriers to establishing and maintaining a systematic assessment cycle with a successful feedback loop. STRATA was conceived not only to assist with key administrative tasks that must be performed but also to help make assessment information readily "actionable" (Kuh et al. 2015, p. 3). As a result, it is helping us to address these issues for our MS course.

Although we are currently blessed to be teaching in the analytics era (Liberatore and Luo 2010), history has shown us the importance of delivering relevant content with high instructional quality in our OR/MS courses. We supply the mathematical models that are at the center of advanced predictive and prescriptive analytics (Lustig et al. 2010). As corporations seek a higher level of analytical maturity, expertise in OR/MS modeling will become critical to improving business processes (Gartner 2018). We need to ensure that our students have the requisite knowledge and skills to build, implement, and interpret OR/MS-based solutions. Employment of STRATA is supporting a data-driven approach to increase student learning of analytics in our core undergraduate business curriculum.

Appendix. Getting Started with STRATA

STRATA is a powerful and easy-to-use tool for analyzing the results of multiple-choice assessments across sessions,

Figure A.1. Load Data File into Excel



teachers, and topics. Entering your data into STRATA is straightforward, but this guide should address any questions that might come up as you prepare to use STRATA.

1. **Before using STRATA, get your data into Excel.** Whether your assessment answers are recorded in a computer or on Scantron sheets, your assessment results will be coming back to you in one or more files. The one you will need for STRATA is the one that contains the raw data: the question-by-question answers given by each student. This file will probably also contain additional information. It may be a .csv file, a .txt file, an Excel file, or another file type that Excel is capable of importing. We'll get that data in shape before opening the STRATA program. If you do not already know how to import a file into Excel, we'll walk through it here.

Run Excel; then click on the **Open** icon. Click on **Browse** on the list of options that appears; then navigate to the directory that contains your file, as we did in Figure A.1. Click on the dropdown menu in the lower right corner of the popup and tell Excel that you want to see **All file** types. Now select your desired file, and click the **Open** button. If your file is a .csv or an Excel file, that's it—Excel should simply open the file.

In this tutorial, our file is a text (.txt) file, so there's a little more to do. When you click Open, you'll see a box like the one in Figure A.2. Click on **Next.**

You should now see something like the window shown in Figure A.3, where each piece of information is sensibly sorted into columns in the lower portion of the window. If it isn't, click on the various choices of the **Delimiters** checkboxes until the sorting is sensible; then click **Finish**.

Your sheet is now in Excel, looking something like the spreadsheet shown in Figure A.4.

STRATA will want column A to contain the ID of the teacher (if any) and column B to contain the ID of the session (if any). Column C should contain the answers to the first question

Figure A.2. Excel's Text Import Wizard (Step 1)

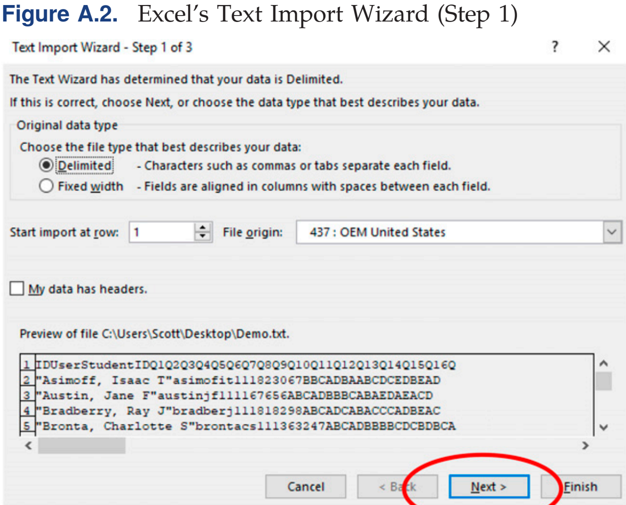


Figure A.3. Excel's Text Import Wizard (Step 2)

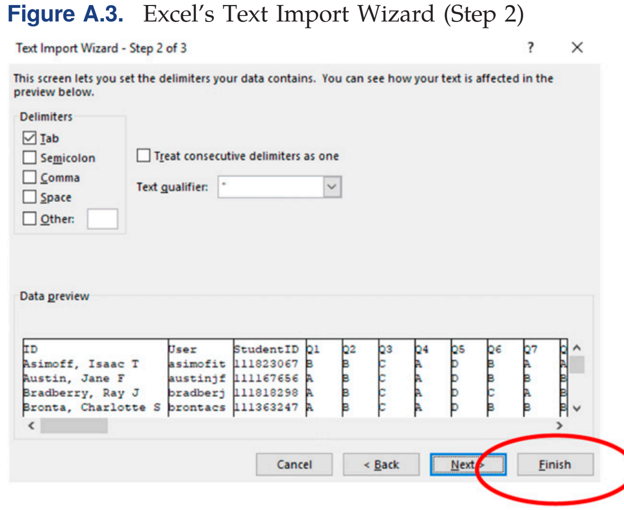
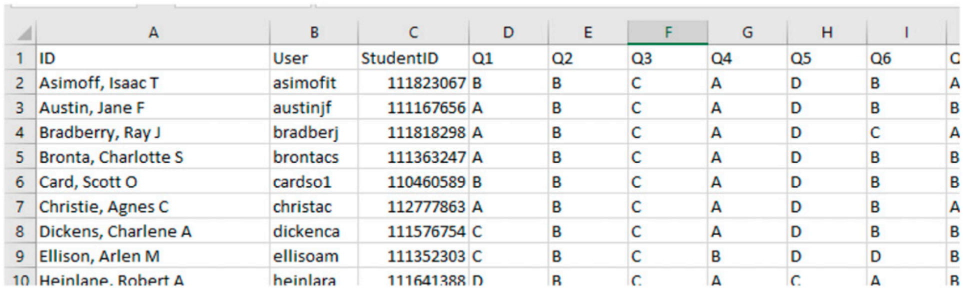
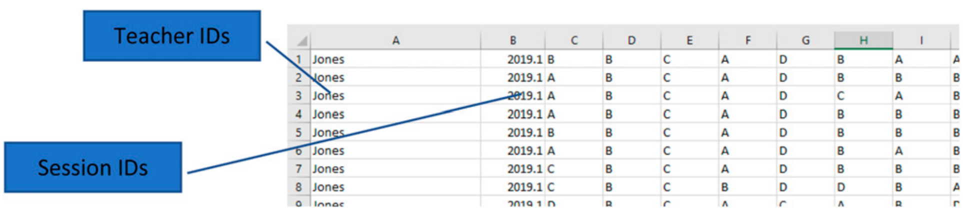


Figure A.4. Data File After Successful Importing into Excel



	A	B	C	D	E	F	G	H	I	
1	ID	User	StudentID	Q1	Q2	Q3	Q4	Q5	Q6	Q
2	Asimoff, Isaac T	asimofit	111823067	B	B	C	A	D	B	A
3	Austin, Jane F	austinjf	111167656	A	B	C	A	D	B	B
4	Bradberry, Ray J	bradberj	111818298	A	B	C	A	D	C	A
5	Bronta, Charlotte S	brontacs	111363247	A	B	C	A	D	B	B
6	Card, Scott O	cardso1	110460589	B	B	C	A	D	B	B
7	Christie, Agnes C	christac	112777863	A	B	C	A	D	B	A
8	Dickens, Charlene A	dickenca	111576754	C	B	C	A	D	B	B
9	Ellison, Arlen M	ellisoam	111352303	C	B	C	B	D	D	B
10	Heinlane, Robert A	heinlara	111641388	D	B	C	A	C	A	B

Figure A.5. First Two Columns are Used for Teacher and Session IDs

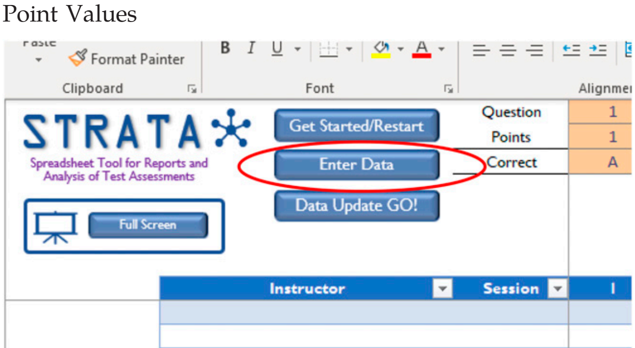


of the assessment. This means that in our example file, we're going to **delete the current column C and use columns A and B for Teacher ID and Session ID.** (The data shown include only students of Professor Jones in the first session of 2019, but the sheet can contain multiple teachers and sessions.) We'll also **delete row one, containing the headers**. The result will look like the spreadsheet shown in Figure A.5.

The session ID must be numeric. If you have multiple files from multiple teachers or sessions, you can combine them into a single Excel file at this time, or you'll be able to add each file like the preceding one into STRATA one at a time. In either case, the preparation is done, and it's time to open STRATA.[1]

When you open STRATA in Excel, it will prompt you to enter a name for your workbook. Doing this will ensure that you always have a pristine copy of STRATA around if you want to begin a new STRATA project later.

Figure A.6. Enter Student Data After Correct Answers and Point Values



2. **Tell STRATA about your exam.** When STRATA opens, follow the onscreen prompts to tell STRATA about your assessment instrument. Note that you may choose how many points to assign to each question; the default is one point per question. After entering your exam key, click on the **Enter Data** button shown in Figure A.6. STRATA is now ready to accept your students' info.

Copy the information from the file you created in step 1 by selecting it, choosing **Copy**, clicking on the blue cell beneath the word **Instructor** in STRATA, and pasting your information.

You can continue to paste additional files beneath any that you have already added—just begin in the first empty cell in the **Instructor** column. When you have finished, click on the **Data Update GO!** button. At any later date (e.g., after a later session) you can add additional results in the same way. Just be sure to click on **Data Update Go!** in Figure A.7 after adding additional results.

Figure A.7. Press the DATA UPDATE GO Button After Adding Student Results

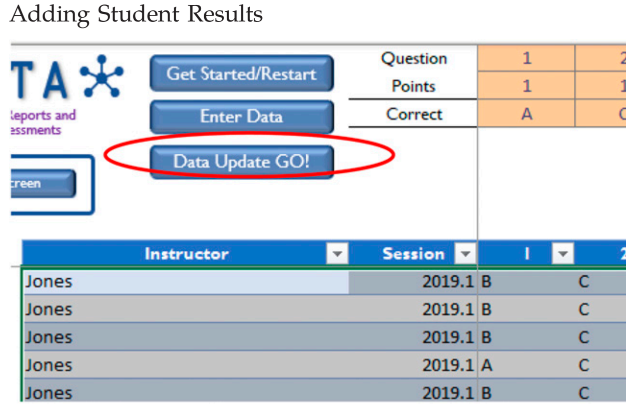
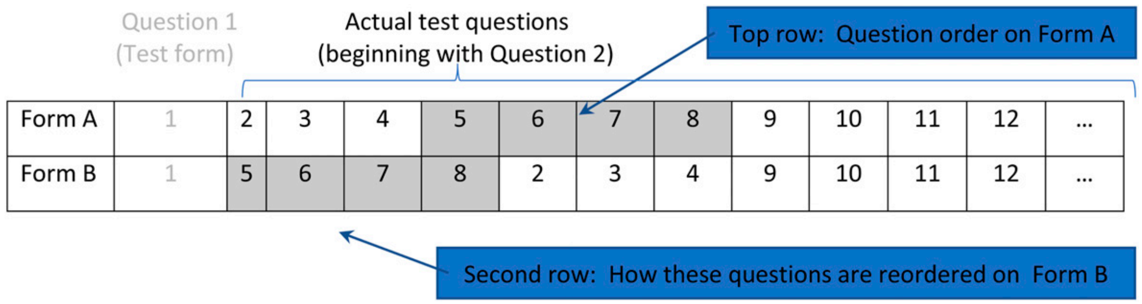


Figure A.8. Multiple Test Forms with Differently Ordered Answers



3. **Identify questions by topic.** STRATA now takes you to its **Topics** sheet. If you wish, you may now identify any question with the topic to which it belongs—simply follow the instructions on that sheet. You do not have to create topics, and not every question has to belong to a topic. When you are finished (or if you do not wish to use topics), click on the **Go!** button. The setup is complete.

A Supplementary Tool Allowing the Use of Multiple Test Forms: UNSCRAM

For additional security, multiple versions of an examination may be desired. The easiest way to accomplish this is simply to reorder the exam questions. STRATA cannot deal directly with multiple test forms, so we provide another spreadsheet tool, UNSCRAM, that provides that functionality. UNSCRAM rearranges each student's answers so that their order corresponds to the order of questions on the master test form. The resulting output can then be pasted directly into the STRATA data input area.

When UNSCRAM is opened, it asks for the number of questions on the examination and then provides an area in which the user records how these questions are reordered in the alternate test versions. UNSCRAM assumes that question 1 on the exam is used to record the student's test form and treats form A as the master test form. Up to four reorderings of questions may be used as different versions, corresponding to forms B, C, D, and E.

Suppose, for example, that form B was obtained from form A by moving a block of questions (questions 5–8 on form A) to the beginning of the form B test. We can imagine creating Form B with a simple cut-and-paste operation, as shown in Figure A.8.

This is exactly the information that UNSCRAM wants you to provide: If the questions on form A kept their original numbers, what would the numbering order on the new form be? On the actual test form B, of course, normal numbering will be applied, so students will see question 1, asking for the test form letter, and then question 2, which was question 5 on the original version, and so on.

Once key information is entered, click on the **Enter Data** button in UNSCRAM. Paste the data you created in step 1 into UNSCRAM. Now click on the **Convert to Master** button. UNSCRAM will reorder the answers so that they conform to the order on the master test form, form A, putting these results on the Data2 tab. At this point, paste this information into STRATA as described in step 2.

Endnote

[1] If the assessment was given using multiple test forms that scrambled the questions, then the student answers will need to be rearranged before using STRATA. We provide another tool, UNSCRAM, for doing this. Instructions for using UNSCRAM appear at the end of this tutorial.

References

AACSB (2013) 2013 eligibility procedures and accreditation standards for business accreditation. Accessed June 6, 2020, https://www.aacsb.edu/accreditation/standards/business.

AACSB (2019) Accreditation standard 8 (2013 business standards): Curricula management and assurance of learning—an interpretation. Accessed June 6, 2020, https://www.aacsb.edu/-/media/aacsb/publications/white-papers/aol_white_paper_standard_8.ashx?la=en&hash=9DF96D5B2A1E0B02E19ACF6984FFFE07D23D45CA.

ABET (2018) Accreditation policy and procedure manual (APPM), 2019–2020. Accessed June 6, 2020, https://www.abet.org/accreditation/accreditation-criteria/accreditation-policy-and-procedure-manual-appm-2019-2020/.

Allen K (2006) The statistics concept inventory: Development and analysis of a cognitive assessment instrument in statistics. Unpublished doctoral dissertation, University of Oklahoma, Norman.

Avery SL, McWhorter RR, Lirely R, Doty HH (2014) Seeking empirical validity in an assurance of learning system. *J. Ed. Bus.* 89(3):156–164.

Banta TW, Blaich C (2010) Closing the assessment loop. *Change* 43(1):22–27.

Banta TW, Ewell PT, Cogswell CA (2016) Tracing assessment practice as reflected in "Assessment Update." NILOA Occasional Paper 28, National Institute for Learning Outcomes Assessment, Champaign, IL.

Black HT, Duhon DL (2003) Evaluating and improving student achievement in business programs: The effective use of standardized assessment tests. *J. Ed. Bus.* 79(2):90–98.

Douglas KA, Purzer Ş (2015) Validity: Meaning and relevancy in assessment for engineering education research. *J. Engrg. Ed.* 104(2):108–118.

Educational Testing Service (2016) Major field tests for the bachelor's degree in business. Accessed June 6, 2020, https://www.ets.org/mft/about/content/bachelor_business.

Evans DL, Gray GL, Krause S, Martin J, Midkiff C, Notaros BM, Pavelich M, et al. (2003) Progress on concept inventory assessment tools. *Proc. 33rd ASEE/IEEE Frontiers Ed. Conf. 2003* (Institute of Electrical and Electronic Engineers, Boulder, CO), session T4G:1–8.

Gartner (2018) Combine predictive and prescriptive analytics to solve business problems. Accessed June 6, 2020, https://www.gartner.com/en/documents/3891993/combine-predictive-and-prescriptive-techniques-to-solve.

Goldman K, Gross P, Heeren C, Herman GL, Kaczmarczyk L, Loui MC, Zilles C (2010) Setting the scope of concept inventories for introductory computing subjects. *ACM Trans. Comput. Ed.* 10(2):1–29.

Hadjicharalambous C (2014) Commercial vs. internally developed standardized tests: Lessons from a small regional school. *J. Ed. Bus.* 89(1):42–48.

Kehoe J (1994) Basic item analysis for multiple-choice tests. *Practical Assessment Res. Evaluation* 4:Article 10.

Ketcham D, Nigro P, Roberto M (2018) Do persistence and passion matter: Evidence from the educational testing service major field test in business. *J. Ed. Bus.* 93(3):112–118.

Klingenberg B (2012) Teaching note: Learning outcome assessment using an integrative assignment on location decision making. *INFORMS Trans. Ed.* 12(3):140–146.

Kuh G, Ikenberry S, Jankowski N, Cain T, Ewell P, Hutchings P, Kinzie J (2015) Beyond compliance: Making assessment matter. *Change* 47(5):8–17.

Liberatore MJ, Luo W (2010) The analytics movement: Implications for operations research. *Interfaces* 40(4):313–324.

Ling G, Bochenek J, Burkander K (2015) Using the Major Field Test for a bachelor's degree in business as a learning outcomes assessment: Evidence from a review of 20 years of institution-based research. *J. Ed. Bus.* 90(8):435–442.

Lustig I, Dietrich B, Johnson C, Dziekan C (2010) The analytics journey. Accessed June 6, 2020, http://analytics-magazine.org/the-analytics-journey/.

Olds BM, Moskal BM, Miller RL (2005) Assessment in engineering education: Evolution, approaches and future collaborations. *J. Engrg. Ed.* 94(1):13–25.

Price BA, Randall CH (2008) Assessing learning outcomes in quantitative courses: Using embedded questions for direct assessment. *J. Ed. Bus.* 83(5):288–294.

Sebastianelli R, Tamimi N (2011) Business statistics and management science online: Teaching strategies and assessment of student learning. *J. Ed. Bus.* 86(6):317–325.

Settlage DM, Wollscheid JR (2015) Analyzing student performance in specific subject area indicators on the ETS major field test in business. *J. Ed. Bus.* 90(6):322–327.

Simkin MG, Kuechler WL (2005) Multiple-choice tests and student understanding: What is the connection? *Decision Sci. J. Innovative Ed.* 3(1):73–98.

Smith LJ, Clements C, Olson JE (2010) Local vs. standardized content assessment: Some management implications, or why bother? *J. Ed. Bus.* 85(5):249–257.

Steif PS, Hansen MA (2007) New practices for administering and analyzing the results of concept inventories. *J. Engrg. Ed.* 96(3):205–212.

Stevens SP, Palocsay SW (2004) A translation approach to teaching linear program formulation. *INFORMS Trans. Ed.* 3(3):38–54.

Stevens SP, Palocsay SW (2012) Identifying addressable impediments to student learning in an introductory statistics course. *INFORMS Trans. Ed.* 12(3):124–139.

Suskie L (2009) *Assessing Student Learning: A Common Sense Guide*, 2nd ed. (Wiley, Hoboken, NJ).

Taylor C, Zingaro D, Porter L, Webb KC, Lee CB, Clancy M (2014) Computer science concept inventories: past and future. *Comput. Sci. Ed.* 24(4):253–276.

Williams JAS, Rankin M, Gallamore K, Reid R (2016) Teaching note: Learning outcome assessment using an integrative assignment on location decision making. *INFORMS Trans. Ed.* 17(1):13–19.