Integrating Student Analysis of Error into the Design of Customized Online Modules for Teaching a Topic in Business Statistics

Jayson Kunzler, Brigham Young University — Idaho
D. Sammons, Idaho State University

Abstract: This paper describes an approach to student error analysis that was developed as part of a project to design, create, implement and evaluate customized instructional modules to teach the topic of normal distribution in an undergraduate business statistics class. In order to identify the required content for the modules, detailed steps within the ADDIE model were followed. These included using detailed objectives to conduct an even more detailed analysis of the tasks that would be required of students in order to master the objectives.
Keywords: Student analysis of error

Error Analysis

Error analysis is a term used by many different disciplines, and that has multiple meanings. A recent search of Google Scholar indicates tens of thousands of articles with *error analysis* in their keywords. And, while a generic definition of *error analysis* might suggest a simple determination of how or when errors occur, *error analysis* in the sciences may examine the degree to which a measurement may vary around a mean or how two estimated values may change if parameters are changed. In the field of education, error analysis has a long history of application in mathematics (e.g., Radatz, 1979) and in second language acquisition (e.g., Corder, 1967, 1978). More recently, Narciss and Huth (2004) explored error analysis in the context of developing informative tutoring feedback. In this paper, we propose a unique approach to using error analysis in the field of instructional design – that is, a way to integrate ethnography of errors, if you will, into the design and development of remedial learning objects.

The Project: Customization of Online Instruction in a Business Statistics Course

The purpose of the larger project was to implement customized online learning modules for undergraduates in a business statistics course. These modules were to act as intelligent tutors, providing highly individualized instruction and remediation as needed by the learners. A single topic, the normal distribution, was selected. The full accounting of the project is the subject of Kunzler's (2012) dissertation research. In this current paper, we would like to address a specific part of the project: how student errors were identified and deconstructed as part of task analysis.

After the development of eleven specific objectives in the Analysis Phase, a task analysis determined the sequence of steps that a learner would need to complete in order to accomplish each objective. Table 1 displays Objective Four and its task analysis as an example. Eighteen different steps are needed to complete the task and demonstrate mastery of Objective Four.

Table 1

Task Analysis for Objective #4 (from Kunzler, 2012)

Objective #4: Given a normally distributed dataset, a computer spreadsheet, and a numerical value, the student will correctly calculate the Z-statistic that corresponds to the numerical value.

Objective #4 Task Analysis	**Task Attributes (Hi, Med, Lo)**			**Required Tools**		**Domain(s)**			
Task	**Difficulty**	**Duration**	**Remediation Likelihood**	**Computer**	**Microsoft Excel®**	**Cognitive**	**Motor**	**Affective**	**Motivation**
1. Open the dataset in a Microsoft Excel® spreadsheet	L	L	L	X	X	X			
2. Identify the appropriate column or row of values that comprise the normally distributed dataset	L	L	L	X	X	X			
3. Select a blank cell on the spreadsheet	L	L	L	X	X	X			
4. Enter the first part of the formula to compute the mean value by typing “=average(” in the cell	L	L	M	X	X	X			
5. Use the mouse cursor to select the column or row of values identified in Task #2 above	L	M	M	X	X	X			
6. Type “)” into the cell after selecting the values	L	L	M	X	X	X			
7. Finish the mean formula by pressing the enter key	L	L	M	X	X	X			
8. Select another blank cell on the spreadsheet	L	L	L	X	X	X			
9. Enter the first part of the formula to compute the standard deviation by typing “=stdev(” in the cell	L	L	M	X	X	X			
10. Use the mouse cursor to select the column or row of values identified in Task #2 above	L	M	M	X	X	X			
11. Type “)” into the cell after selecting the values	L	L	M	X	X	X			
12. Finish the standard deviation formula by pressing the enter key	L	L	M	X	X	X			
13. Select another blank cell on the spreadsheet	L	L	L	X	X	X			
14. Enter the given value into the cell	L	L	L	X	X	X			
15. elect another blank cell on the spreadsheet	L	L	L	X	X	X			
16. Option A: Calculate the Z-statistic by entering the formula “=(x -m)/s”, where x, m, and s are entered by clicking on the cells that contain the given value, the calculated mean, and the calculated standard deviation, respectively	M	M	M	X	X	X			
17. Option B: Calculate the Z-statistic by entering the formula “=standardize(x,m,s)”, where x, m, and s are entered by clicking on the cells that contain the given value, the calculated mean, and the calculated standard deviation, respectively	M	M	M	X	X	X			
18. Round the Z-statistic to the nearest hundredth, tenth, or whole number (as specified)	M	L	M	X	X	X			

In order to design and develop the customized learning modules, it was critical to understand what types of errors students might make in completing this task and in attempting to meet Objective Four. An analysis of student errors was undertaken in stages – the analysis first asked WHAT errors the students made, then WHERE in the task sequence the errors were made, and finally, WHY students made those errors. While the WHAT and WHERE questions are more typical of the Design Phase, it was the qualitative study of WHY that provided the best approach to designing the instruction itself.

WHAT errors did the students make?

To identify which errors students might make while attaining the objectives, a 10-item achievement assessment (which had been determined to be aligned with the objectives) was administered to the undergraduate business statistics class in the semester prior to the implementation of the study (Kunzler, 2012). These students had completed instruction on the topic, the normal distribution. Student responses were recorded and categorized by response. The frequency with which any given response occurred was calculated. Figure 1 shows the frequency of each response for Question 7, which was aligned with Objective 4. Question 7 was a free-response (not multiple choice) item. As can be seen, only six students (9%) returned the correct answer; 91% of students responded incorrectly.

Although the identification of WHAT errors occurred was useful in identifying preliminary patterns of incorrect responses., it was not adequate to facilitate the design and development of the customized instructional modules. Knowing WHAT errors occurred (and their frequency) primarily highlighted the most common errors and indicated that most students were unable to master the objective. In a routine ADDIE process, knowing that most students missed this question and could not meet the objective might provoke a design response of repetition and additional practice. However, merely adding more problems would not meet the goals of the project, which was to create highly individualized instruction.

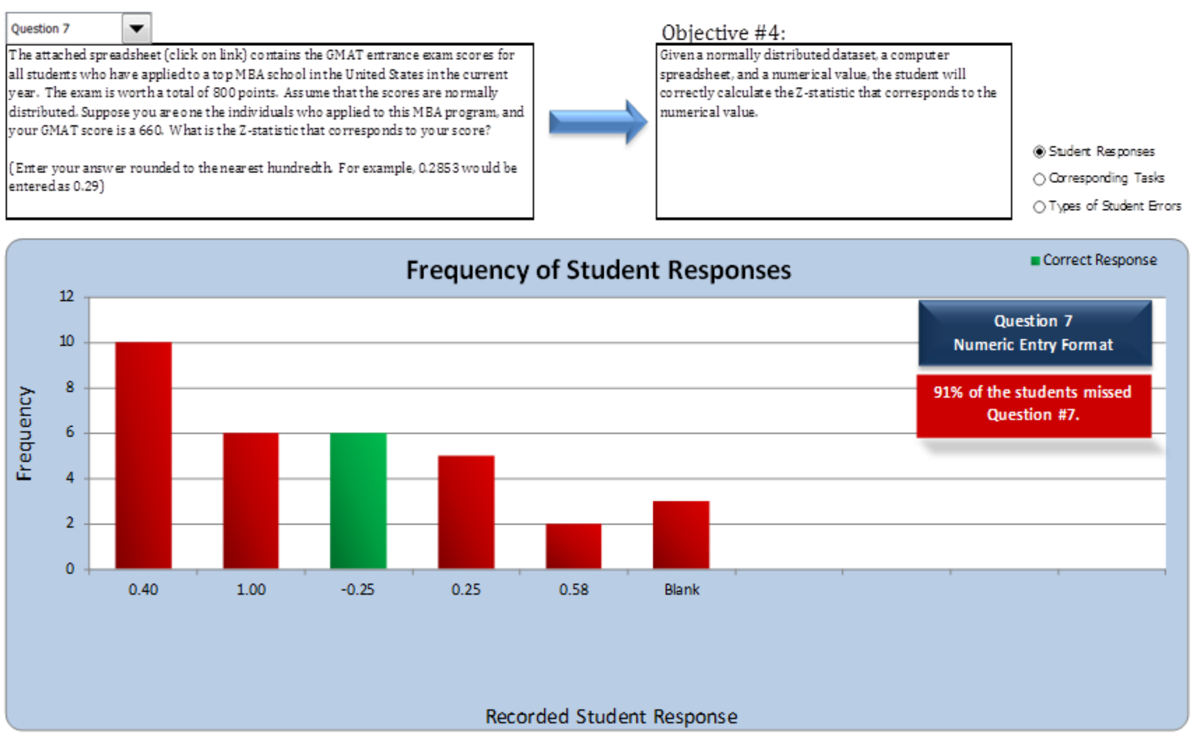

Figure 1. Sample Frequency Chart Showing Frequency of Given Student Responses.

WHERE did students make their errors?

The next attempt to analyze errors in a way that would inform the design process was to map the errors against the elements of the task. In this step, students were asked to describe their error, and then each students' descriptions were compared to the specific steps in the task. In the example for Question 7 below, Table 2 contains a few of the students' descriptions, and Figure 2 shows how all the students' descriptions are summarized *vis à vis* the tasks pertaining to Objective Four.

As can be seen in Figure 2, almost all of the student errors were related to Task 4, #17: *Calculate the Z-statistic by entering the formula "=standardize"*.

Table 2.

Sample of Student Responses and Errors for Question #7.

Shaded = Incorrect

Question #7 Student Responses and Reasons for Errors

Response	Correct?	Self-Explanation for Reason
0.40	N	I didn't use the right formula I didn't get the right Z, I also rounded the inputs. I used Normdist instead of standardize
	N	Didn't answer because I didn't know how to do it
1.00	N	I didn't know which function to use so I got it wrong. I calcutlated normsdist of 660 to get 1.
0.25	N	I did it right but I put a positive .25, when it is a negative .25.
-0.54	N	I understood the concept but must have entered some info wrong I put -.54.
-0.25	Y	I got it correct
0.40	N	In this problem, you should use the "STANDARDIZE" function. I used the "NORMDIST" function instead.
1.00	N	I put 1 because I did the formula normsdist and put 660 and it gave me one. I didn't really understand what I was looking for.
0.40	N	I got this wrong because I used the NormDist formula instead of using the standardize formula
0.35	N	I guessed. I really had no idea what to do.
0.61	N	I got it wrong and I used the normdist formula which was the wrong formula. Should have used the standardize function
0.40	N	I did the mean and the standard deviation but I used it in the wrong formula. I did the NORMDIST instead.
0.40	N	I couldn't remember how to get the z score I used the normdist function and put that as the answer.
0.25	N	I didn't know if it needed to be written as positive or negative so I wrote it out as a positive number in the quiz answer.

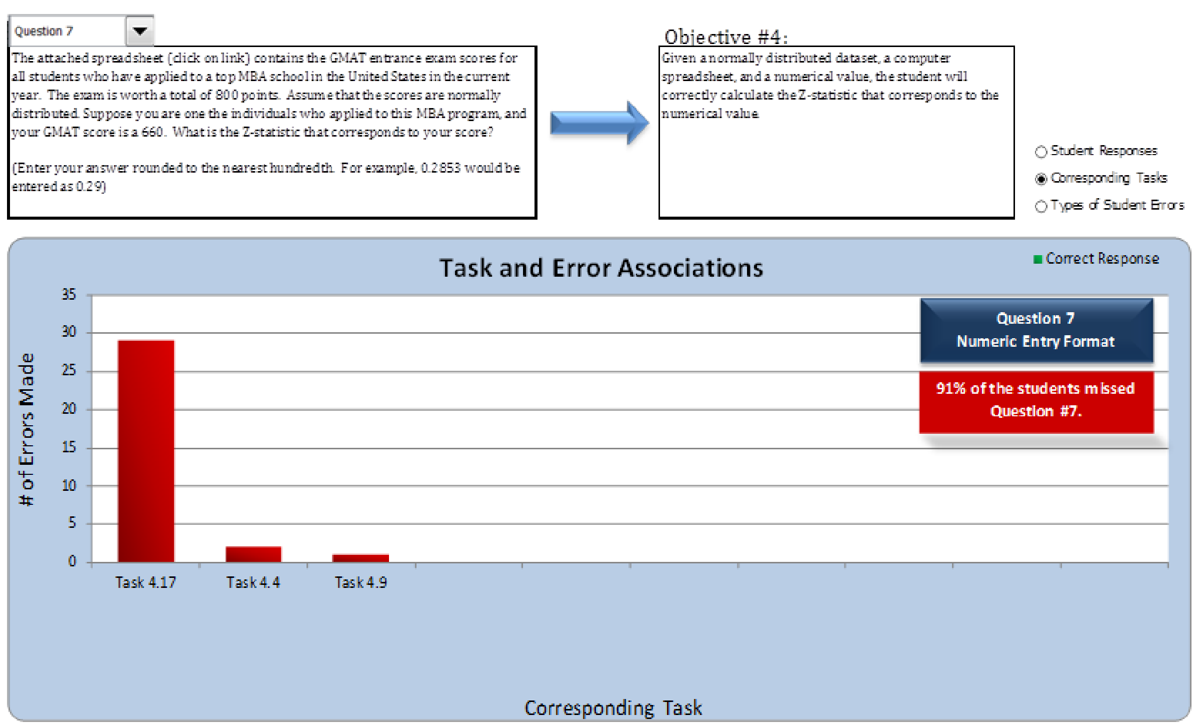

Figure 2. Sample Frequency Chart Showing Frequency of Errors According to Task.

However, once again, using this method to identify potential topics for the customized instructional modules proved to be insufficiently helpful: there is not a fine enough discrimination between the student descriptions when those descriptions are mapped onto the task list. Therefore, combining an error analysis with the task analysis could not form an adequate basis for the design and development of the customized instructional modules.

WHY did students make their errors?

Although the task analysis itself was important in designing the instructional programs, categorizing student errors by corresponding task proved to only be marginally useful because it resulted in insufficient differentiation among the student error categories. Comparing the WHAT category of error (Figure 1) with the WHERE category of error (Figure 2) indicates that only three steps within the task analysis were involved. Five different incorrect responses were given by two or more students, and several other incorrect responses were given by individual students. The five different incorrect responses were made by 91% of the students! If customized remedial instructional modules are to be successfully designed and developed, the source of student error must be more fully understood. Therefore, a final and different approach to the problem was attempted: understanding WHY students made the errors they did.

In this approach, again using only Question #7 as an example, each of the six responses (one correct and five incorrect) were analyzed in conjunction with the students' self-stated reasons for their answers. Students who had arrived at the same (incorrect) response had different reasons for coming to that response; and students who had arrived at different (incorrect) responses had actually committed the same error. Individual student self-reported reasons were analyzed and coded into error categories. The coded error categories were phrases that summarized specific student behavior that led to the error. As indicated in Figure 3, although five incorrect responses were given by two or more students, there were six different error categories identified by the researcher.

As shown in the sample data in Figure 3, the most common student error on Question #7 occurred when students mistakenly used the *normdist* function instead of the *standardize* function. Another fre-

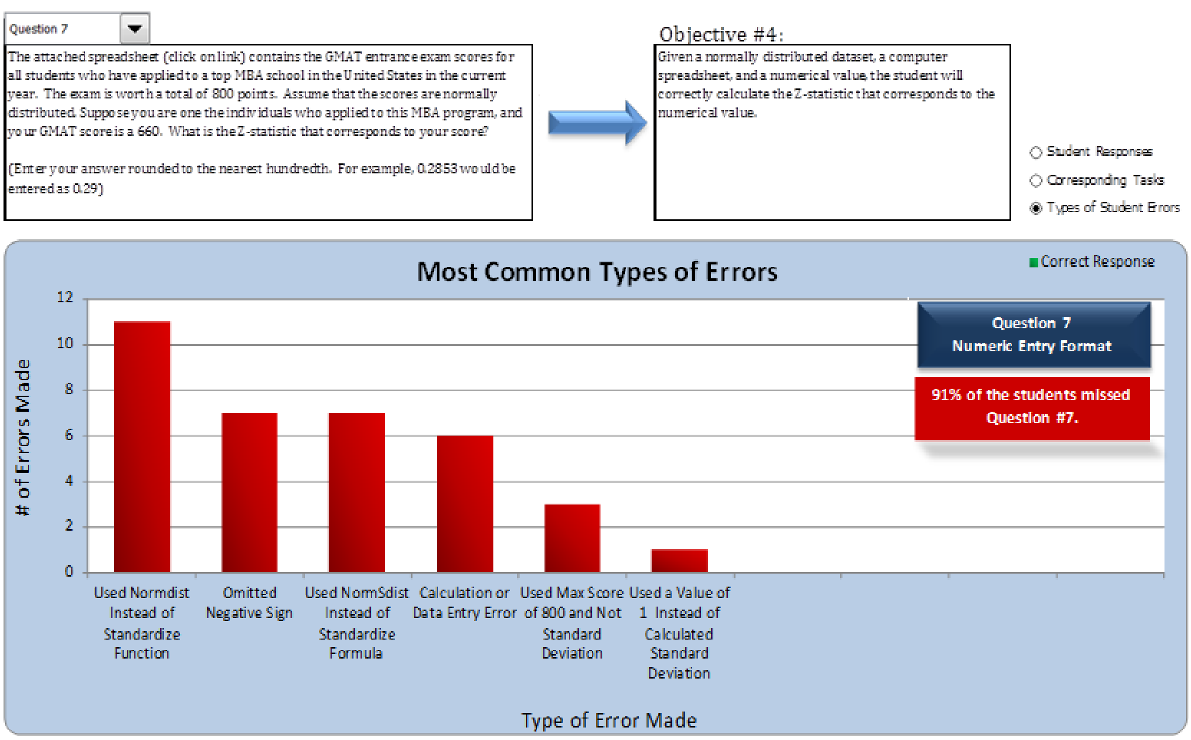

Figure 3. Frequency Chart Showing Frequency of Errors by Error Type.

quently occurring error for this question was the omission of the negative sign. This step in the process allowed the designer to identify the most common errors and to develop customized instructional modules based on the identified reasons for common errors. It also allowed the designer to determine if the cost of developing customized instructional modules would be too high for uncommon errors.

The process exemplified in the tables and figures in this paper for Question 7 and Objective Four were followed for each of the ten questions of the assessment, for each of the eleven instructional objectives, and for each of the elective task lists. In every case, the final qualitative consolidation of self-reported student error analysis (the WHY error analysis) proved to be the most useful in designing and developing the customized instructional modules. Only the final WHY error analysis provided critical discrimination among error types and supported the development of the final customized instructional modules.

The Application in Support of the Error Analysis

In a final note, among the unique aspects of this student error analysis is the fact that all the analyses themselves – the tabulation of correct and incorrect responses, the WHAT, WHERE, and WHY analyses – were all conducted within a single Microsoft Excel spreadsheet utility (Kunzler, 2012). The senior author was able to program Microsoft Excel to analyze the data and to display those data in a series of dashboard graphics enabling the user to move from one question to another, and from one analysis to another, by using drop down lists and option buttons. Programming these features in Microsoft Excel appears to be a new contribution to the functions of the spreadsheet application in error analysis.

Conclusion

The analysis of student errors was a very helpful and necessary input into the process of designing and programming the customized instruction. The results of the analysis were used to predict the errors students will most likely make in the future, thus enabling the development of logical algorithms for launching customized instructional content when student errors exist. The error prediction then warranted the development of a specific piece of customized instruction to be launched for students who made an error during one step of the process. In the *custom* module developed for this larger project (Kunzler, 2012), these pieces of customized instruction were delivered to individual students as video content. A unique piece of video content was developed for the purpose of instructing students for each type of error made during instruction. The customized modules would have been much more limited if WHAT student errors only had been identified. Instead, by also understanding WHERE and WHY student errors occured, the resulting modules were completely customizable to student needs.

References

Corder, S. P. (1967) The significance of learners' errors. *International Review of Applied Linguistics 5*: 161-9

Corder, S.P. (1978). Simple Codes and the Source of the Second Language Learner‟s Initial Heuristic Hypothesis. *Studies in Second Acquisition.* 1, 1-10

Kunzler, J. (2012). *Exploring Customization in Higher Education: An Experiment in Leveraging Computer Spreadsheet Technology to Deliver Highly Individualized Online Instruction to Undergraduate Business Students*. Unpublished PhD dissertation, Graduate Department of Educational Leadership & Instructional Design, College of Education, Idaho State University, Pocatello.

Narciss, S. & Huth, K. (2004). How to design informative tutoring feedback for multimedia learning. In H. Niegemann, R. Brünken and D. Leutner (Eds.), *Instructional design for multimedia learning* (pp. 181-195). Münster, Germany: Waxmann.

Radatz,. (1979). Error Analysis in mathematics education. *Journal for Research in Mathematics Education*, 10: 163-172.