

Evaluating and Improving the Assessment of Undergraduate Student Writing in a School of Business at a Large Regional University

David Glew, Tracy Meyer, Becky Sawyer, Pete Schuhmann, and Barry Wray¹
UNC Wilmington, Wilmington, North Carolina, 28403

Abstract

Business schools are often criticized for the inadequate writing skills of their graduates. Improving writing skills involves first understanding the current skill level of students. This research attempts to provide insights into the effectiveness of the current method of assessing writing skills in a school of business at a large regional university and to provide insights relative to how the assessment process can be improved. Three studies were conducted. Study 1 is a quantitative analysis of the full set of data based on the use of *Criterion*®ⁱ, an online writing evaluation service developed by Educational Testing Service (ETS). Study 2 is a quantitative analysis of a coded subset of the data and Study 3 involves a qualitative data collection effort consisting of comments gathered from assessment proctors. While the overarching purpose of this research was to determine the effectiveness of the assessment tool utilized towards understanding the current status of students' writing skills, the results also offer recommendations to improve the assessment process such that a clear picture of writing weaknesses can be more effectively addressed.

Keywords: Writing skills, assessment, teaching effectiveness.

Accreditation by the Association to Advance Collegiate Schools of Business (AACSB) is the highest certification a business school in the United States can attain. In order to achieve and maintain this accreditation a school must develop a systematic process to “develop, monitor, evaluate, and revise the substance and delivery of the curricula of degree programs and to assess the impact of the curricula on learning.”ⁱⁱ A very important component of this curriculum is student communication abilities.ⁱⁱⁱ Therefore business schools wishing to be accredited by the AACSB must ensure all students acquire the necessary skills to be effective oral and written communicators. In order to accomplish the goal of ensuring effective writing skills, one business school at a large regional university uses an automated online evaluation system that requires students to write a timed essay and then provides a holistic score as well as sub-scores for grammar, usage, mechanics and style for each student. This research examines the effectiveness of this online student writing skills assessment process and provides recommendations for improving the process.

The business school has stated as one of its undergraduate learning goals:

¹ Corresponding author's email: wrayb@uncw.edu

“Our students will be able to conceptualize a complex issue into a coherent written statement and oral presentation, demonstrated with the effective use of technology.”

To measure the extent to which students are able to conceptualize a complex issue into a coherent written statement, the following process was implemented:

“To effectively measure written communication, the business school has adopted an ETS instrument developed through AACSB. This instrument is administered in the capstone classes of each major every semester. Faculty are trained in administration of the writing assessment.”

This learning goal and corresponding assessment process are part of the broader business school Assurance of Learning Plan. *Criterion*², an online writing evaluation service developed by Educational Testing Service (ETS), was selected to measure students' writing proficiency. *Criterion* requires students to write a timed essay on a given subject which is then evaluated by the ETS computerized *e-rater*¹ scoring engine and *Critique*¹ writing assessment tools. Business school writing assessments began in spring 2008 and have been conducted in each department's capstone courses. As of the start of fall 2009, approximately 800 students had completed the writing assessment.

A research team composed of one member from each department of the business school was formed to analyze the results to date from the *Criterion* assessments. Specifically, the research team was charged with “making the ‘first cut’ at interpreting the data, focusing on the areas that should be improved and identifying ways they can be improved.”

Research Question

The question for the research reported in this article is “how effective is the current method of assessing writing skills in the business school?” Three studies were conducted to analyze the *Criterion* assessment data: 1) a quantitative analysis of the full set of *Criterion* data, 2) a quantitative analysis of a coded subset of data, and 3) a qualitative analysis of comments gathered from assessment proctors.

Literature Review

The current research on assessment of writing skills in business schools is limited. Warrnock (2009) described a pilot effort for an accreditation-driven writing assessment in a business school and detailed the pilot's logistics and methods. Bacon, Paul, Johnson and Conley (2008) and Bacon and Anderson (2004) researched assessing and improving writing skills of marketing students by using an internally developed assessment tool. They developed a computer-scored measure of students' ability to edit for writing mechanics. The authors tested the method with a pretest and posttest design and found that when stu-

² *Criterion*® and *e-rater*® are registered trademarks of Educational Testing Service (ETS). *Critique*™ is a trademark of ETS.

dents were held accountable for improvement (with 5% of the course grade dependent on improvement) there was a significant increase in students' ability to edit. Without that accountability, improvement did not occur. Wolf, Connelly, and Komare (2008) and Wolf and Stevens (2007) examined the role of rubrics in assessing teaching and student learning. They found different assessment rubrics can be used to improve both courses and programs in different fields and disciplines. In addition, Quible, Zane, and Griffin (2007) examined possible reasons for writing deficiencies for business writers. They offer evidence that a modified context-based approach, the glossing approach, and consistent error marking can reduce the number of sentence-level errors students make. Fraser et al. (2005) developed a multilevel assessment of business writing using both an online objective test and in class writing exercise. Zhao and Alexander (2004) identified short- and long-term impacts of business communication education on students' writing skill developments and performance outcomes. Earlier research in assessing business writing was done by Varner, Iris, and Pomerence (1998). They determined instructors need to be aware of issues such as developing clear goals from the beginning to achieve effective reliability and validity when conducting a writing assessment.

Methodology

Two studies were conducted to analyze the *Criterion* assessment data:

- Study 1: A quantitative analysis of the full set of *Criterion* data
- Study 2: A quantitative analysis of a coded subset of data

In addition, a third study was conducted to more closely examine the assessment process:

- Study 3: A qualitative analysis of comments gathered from assessment proctors

Study 1

Sample. Over the period of spring 2008 to summer 2009, 860 business school seniors took the *Criterion* writing assessment in their capstone course. Due to incomplete data, 37 observations were dropped leaving a sample size of 823.

Assessment. Upon logging in to the *Criterion* online assessment system, students were given these instructions: "Present your perspective on [the following issue], using relevant reasons and/or examples to support your views. *'It is important for higher education to challenge established traditions and values'*" (italics added). Students were required to submit their essay within 45 minutes.

Variables. The *Criterion* assessment data includes the following variables (a detailed explanation of these variables is provided in Appendices A and B):

1. Time to complete the essay
2. Holistic score (ordinal scale from low of 1 to high of 6)
3. Number of grammar errors

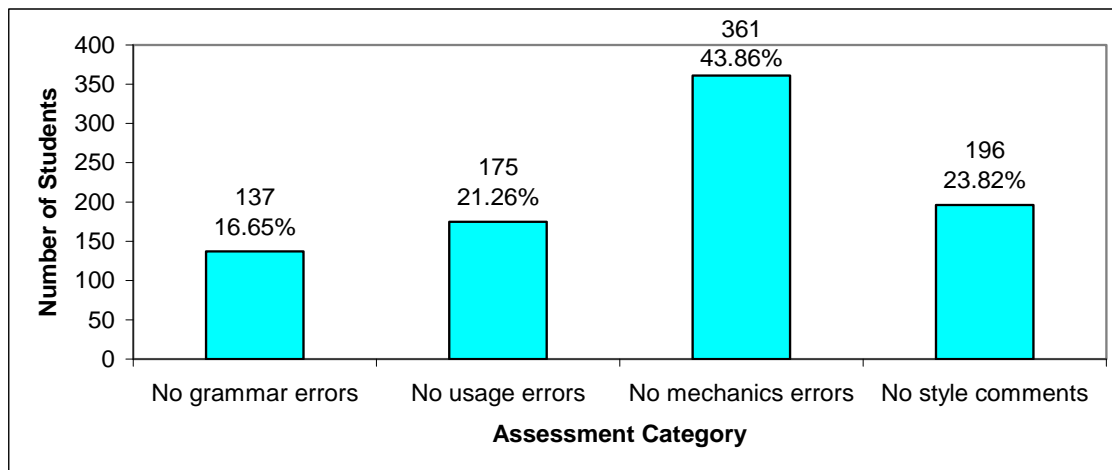
- 4. Number of usage errors
- 5. Number of mechanics errors
- 6. Number of style comments

Analyses. Descriptive statistics and inter-item correlations are provided in Tables 1 and 2. For ease of interpretation, the distribution of holistic scores is depicted in Figure 1. The mean and standard deviation of the number of errors or comments for grammar, usage, and mechanics is listed in Table 1. To further illustrate the prevalence or absence of errors or comments, the number of essays with no errors or comments in each category is shown in Figure 2

Table 1. Study 1 Descriptive Statistics

	Minimum	Maximum	Mean	Std. Deviation
Holistic Score	1	6	3.77	1.10
Time (minutes)	.03	45.00	29.49	11.26
Grammar Errors	0	19	2.50	2.31
Usage Errors	0	12	2.08	1.88
Mechanics Errors	0	57	1.57	3.53
Style Comments	0	89	13.36	14.50

Figure 2. Study 1 Student Essays with No Errors or Style Comments



A more in-depth investigation of these data was conducted by performing an ordered logit multiple regression analysis of holistic score on other variables (see McKelvey & Zavoina, 1975 for details). Since essays would likely contain more or fewer errors simply because of their length (e.g., longer essays would be expected to have more errors) variables 3 – 6 above were standardized by dividing the number of errors or comments in

each category by the time taken to complete the essay. Hence three “errors per minute” and one “comments per minute” variable were created.

The ordered logit multiple regression analysis produced the following results:

1. While controlling for major (using 7 indicator variables with POM as the baseline), the number of grammar errors per minute *increases* the holistic score. This result is highly statistically significant (p -value < 0.001).
2. While controlling for major (using 7 indicator variables with POM as the baseline), the number of usage and mechanics errors and style comments per minute *decreases* the holistic score. These results are highly statistically significant (p -values < 0.005).

These results must be viewed with caution however due to the existence of multicollinearity in the independent variables (Table 2).

Table 2. Study 1 Inter-item Correlations

	2	3	4	5	6
1. Holistic Score	.45**	.31**	.27**	-.05	-.13**
2. Time	–	.20**	.29**	.14**	-.12**
3. Grammar Errors		–	.33**	.37**	.09*
4. Usage Errors			–	.19**	-.03
5. Mechanics Errors				–	.06
6. Style Comments					–

*Correlation is significant at the 0.05 level (2-tailed).

** Correlation is significant at the 0.01 level (2-tailed).

Study 2

Sample. A stratified random sample of 119 essays was drawn from the larger data set of 823 essays used in study 1. Twenty essays for each holistic score were included (except for holistic score 6 for which only 19 essays were available).

Coding and Variables. The same error and comment variables described in study 1 were available for the 119 essays. In addition, variables related to the length and content of the essays were calculated (e.g., key words from the essay topic "It is important for higher education to challenge established traditions and values" were counted). The variables that were calculated were:

1. Number of words
2. Average word length

3. Number of sentences
4. Average sentence length
5. Instances of the phrase “higher education”
6. Instances of the word “challenge”
7. Instances of the word “established”
8. Instances of the word “values”
9. Instances of the word “tradition”
10. Instances of the word “importance”

Analyses. Descriptive statistics and inter-item correlations are provided in Tables 3 and 4.

Table 3. Study 2 Descriptive Statistics

	Minimum	Maximum	Mean	Std. Deviation
Holistic Score	1	6	3.48	1.71
Grammar Errors	0	14	2.69	2.87
Usage Errors	0	9	1.87	1.85
Mechanics Errors	0	43	1.85	4.62
Style Comments	0	89	16.64	15.13
Total Words	53	1057	364.96	243.94
Avg. Word Length	4.11	5.94	4.99	.34
Total Sentences	3	53	17.60	10.96
Avg. Sentence Length	10.33	39.15	20.79	4.76
“Higher Education”	0	14	4.19	3.14
“Challenge”	0	20	4.09	3.12
“Established”	0	10	2.60	2.45
“Values”	0	17	5.73	3.83
“Traditions”	0	19	4.96	3.35
“Important”	0	9	1.88	1.55

An ordered logit multiple regression of holistic score on grammar, usage, mechanics and style errors per word was conducted. In study 1, grammar, usage, mechanics and style variables were standardized to account for different amounts of time spent on the essays. In study 2, a similar standardization was performed to convert these values into three “errors per word” and one “comments per word” variables.

The ordered logit multiple regression analysis in study 2 produced the following results:

1. The coefficient on grammar errors per word was *positive* and highly statistically significant (i.e., higher holistic scores were associated with more grammar errors).
2. The coefficients on mechanics and style comments per word were *negative* and statistically significant (i.e., higher holistic scores were associated with fewer mechanics errors and fewer style comments).

3. The coefficient on usage errors per word was not statistically significant.

These results must be viewed with caution however due to the existence of multicollinearity in the independent variables (Table 4).

Study 3

Sample. The sixteen business school faculty members who proctored the *Criterion* assessment in their courses were asked to briefly respond the questions shown below. Eleven (69%) faculty members responded.

Qualitative Content Analysis. The eleven responses to each question were examined for similarities and differences. The summaries of these responses are based on the subjective judgment of the research team.

Q1. How do you fit the assessment into your course (e.g., set aside one day of the semester, ask students to take the assessment outside of class)?

Four (36%) respondents used class time to give the assessment. Seven (64%) asked students to take the assessment outside of class.

Q2. Do you make the assessment a course requirement, give extra credit, or in any other way link the assessment to students' course grades? Do you use any other incentives?

Five (45%) respondents made the assessment a course requirement, but include no other incentives. Three (27%) respondents linked taking the assessment to the students' course grades, but did so variously by offering a nominal amount of extra credit or telling students they would lose points if they did not complete it. Two respondents (18%) reported giving extra credit.

Overall, responses to this issue varied from providing incentives to separating the assessment as much as possible from the course. In addition, some respondents approached the assessment differently at the time they answered the questionnaire than they had in the past. At least one respondent mentioned that the writing assessment was unpopular with students and had negatively affected course evaluations.

Q3. Besides the technical information of how to login to the ETS *Criterion* website, what other information do you give the students? For example, do you explain the purpose of the assessment? Do you play up or downplay the importance of the assessment? Do you encourage a high level of effort or do you try to remain neutral and let students determine their own level of effort?

Most respondents encouraged students to do their best. They explained the importance of the assessment for business school accreditation purposes, and highlighted the potentially

Table 4. Study 2 Inter-item Correlations

	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. Holistic Score	.46 **	.41 **	-.11	.14	.91 **	-.04	.85 **	.27 **	.51 **	.52 **	.48 **	.58 **	.48 **	.38 **
2. Grammar Errors	–	.47 **	.38 **	.36 **	.60 **	-.39 **	.48 **	.35 **	.41 **	.31 **	.17	.33 **	.29 **	.39 **
3. Usage Errors		–	.34 **	-.03	.46 **	-.01	.40 **	.20 *	.32 **	.39 **	.22 *	.33 **	.30 **	.38 **
4. Mechanics Errors			–	.27 **	.02	-.25 **	-.01	.14	.12	-.02	-.06	-.07	-.06	.05
5. Style Comments				–	-.04	-.34 **	-.01	-.08	-.00	.10	-.00	-.04	.07	.14
6. Total Words					–	-.23 *	.91 **	.33 **	.51 **	.49 **	.44 **	.56 **	.45 **	.40 **
7. Avg. Word Length						–	-.27 **	.01	.00	-.01	.25 **	.03	.10	-.10
8. Total Sentences							–	-.02	.44 **	.51 **	.41 **	.53 **	.45 **	.38 **
9. Avg. Sentence Length								–	.21 *	.08	.11	.13	.06	.14
10. Higher Education									–	.39 **	.39 **	.40 **	.29 **	.41 **
11. Challenge										–	.45 **	.54 **	.60 **	.43 **
12. Established											–	.62 **	.59 **	.39 **
13. Values												–	.83 **	.40 **
14. Traditions													–	.42 **
15. Important														–

* Correlation is significant at the 0.05 level (2-tailed).

** Correlation is significant at the 0.01 level (2-tailed).

valuable diagnostic feedback students would receive. Other respondents remained neutral or informed students that they would review the time students spent writing their essay and base course credit on their apparent effort.

Q4. What is your impression of how seriously the students take the assessment?

Responses to this question varied widely. Some felt most students took the assessment seriously because they were interested, wanted to get a good score, or wanted the busi-

ness school to look good. Others stated that students did not take the assessment seriously because of a lack of consequences or because the essay question was too abstract. Some respondents believed their efforts at encouraging students, including changing the way they included the assessment in their courses, had increased students' motivation.

Q5. In your opinion, is the ETS *Criterion* online writing assessment a good way to measure writing ability? Is there a different approach you would suggest we take?

Faculty members were divided on whether *Criterion* is a good way to measure writing ability. Some pointed to efficiency and consistency as the primary advantages of an automated scoring system. Other respondents noted that the system could not evaluate content, so students could write their essay on an unrelated topic and still receive good scores. This was especially true given the topic students were asked to address. The lack of comparison data from a national sample was also lamented. A third group of respondents were simply unsure whether *Criterion* was a good way to assess writing ability.

Recommendations for alternative approaches included offering a business communications course, hiring a trained/professional writer to evaluate students' work, and assessing writing within a course based on existing or additional written assignments.

Respondents were invited to make other comments

Additional comments included:

- Recommending and offering support for a business communications class
- Recommending a stronger focus on writing in the freshman English course
- Reporting an increased focus on writing within the respondents' own courses
- Requiring students to re-write poor quality work and to visit the writing center
- Asking where business school students actually learn to write
- Noting that writing ability should have been addressed in admissions testing and in English courses

Interpretation of Results

Quantitative Data Analyses

The business school has adopted the following self-defined labels for three categories of holistic scores:

<u>Holistic Score</u>	<u>Label</u>
1-2	Needs a lot of work
3-4	Needs some work
5-6	Doing ok

At first glance, the quantitative data analyses suggest business school students, with an average holistic score of 3.77, need some improvement in their writing, but not much. Average error rates for grammar (2.50 errors per essay), usage (2.08) and mechanics (1.57) do not appear to indicate a crisis-level situation. Moreover, a substantial proportion of students made no errors in grammar (16.65% made no grammar errors), usage (21.26%), and mechanics (43.86%).

A more in-depth analysis of the data reveals several unexpected results. Holistic score had a strong significant *positive* correlation with grammar and usage errors. In fact, grammar errors were the best predictors of holistic scores in the ordered logit regression analyses. In other words, the essays receiving the highest holistic scores were the ones with the most grammar errors. However, there was multicollinearity in the logistic regression model as a result of strong correlations between the independent variables (tables 2 and 4). The existence of multicollinearity inflates the variances of the parameter estimates. The inflated variances may result in lack of statistical significance of some individual independent variables while the overall model was strongly significant. Multicollinearity may also result in wrong signs and magnitudes of regression coefficient estimates and, consequently, incorrect conclusions about relationships between independent and dependent variables.

These unexpected results may also be explained in part by examining the relationship between holistic scores and two measures of essay length: time spent on the essay and number of words. Of these two measures, number of words is more accurate because some students wrote their essays outside of *Criterion*, and then cut and pasted the text into the online system (thus the recorded time spent by some students of mere seconds). Holistic score and time spent on the essay are correlated at $r = .45$. Holistic score and number of words are correlated at $r = .91$. This means the length of the essay alone explains approximately 83% of the variance in holistic scores.

Taken together, these relationships indicate that students who wrote longer essays made more grammar and usage errors, but also received higher holistic scores. One tentative conclusion that could be drawn from these results is that the *Criterion* evaluation system appears to give more weight to the length of an essay than to the appropriate use of grammar.

Style comments had a significant but negative association with holistic score. *Criterion* does not assign style “errors”, but rather counts items such as the repetition of certain words and sentences beginning with coordinating conjunctions. Each instance of these items creates a style comment (e.g., if a student used the word “we” fifteen times, the essay would receive fifteen style comments). Style comments appear more subjective than the grammar, usage and mechanics errors identified by *Criterion*, and do not necessarily indicate a problem with the student’s writing. For instance, the essay that received the highest number of style comments – 89 – also received the highest holistic score possible – 6. Due to these factors, style comments may offer little benefit in assessing student writing.

Mechanics errors also had a significant negative association with holistic score. However, nearly one-half of all students' essays (43.86%) had no mechanics errors, so less improvement appears needed in this area than in the other error categories.

Holistic score was significantly correlated with the incidence of all six key words taken from the essay prompt: "higher education," "challenge," "established," "tradition," "values," and "important." The average correlation between holistic score and these words was $r = .49$.

Qualitative Data Analysis

A fundamental tenet of scholarly research is that data are no better than the procedure used to gather them. Comments by the faculty members who proctored the *Criterion* assessment reveal substantial differences in the way the assessment was administered. Across the various capstone courses, students received inconsistent instructions regarding the importance of the test and inconsistent incentives for successful completion. This is the case across proctors in a given semester as well as for some proctors across time.

Such inconsistency in the administration of the assessment creates a situation where otherwise strong writers could produce low-scoring essays. This induced variability in the variable we are attempting to explain (holistic score) makes any analysis of cause-and-effect suspect and likely unreliable.

Moreover, because the *Criterion* system grades essays using strictly mechanical means (word counts, error counts, etc) it is also highly likely that otherwise nonsensical essays could produce high holistic scores. This possibility was confirmed by a quick review of some of the actual essays in the dataset. They frequently drifted off topic and sometimes were completely unrelated to the essay prompt (e.g., one was written on Kobe Bryant and the LA Lakers; another begins "I really do not want to do this. This is probably the dumbest thing I have ever heard of."; a third ends with "I love you.").

ETS acknowledges^{iv} these sorts of weakness in *Criterion*:

"Can students trick the Criterion service?"

Yes. Since the e-rater engine cannot really understand English, it will be fooled by an illogical, but well-written, argument."

"Will the Criterion service catch cheating or plagiarism?"

No. The Criterion service simply evaluates the essay."

In short, interpreting the results of the quantitative analyses described is severely limited by the variability in how the writing assessment has been administered thus far in the business school, and the inability of *Criterion* to accurately recognize the content of the essays. At best, these results should be interpreted with a great deal of caution.

Recommendations

Two general sources of information about business school student writing exist: 1) faculty anecdotes and 2) the *Criterion* assessment data addressed in this report. Faculty anecdotes bring to mind egregious instances of poor student writing, but are not a systematic evaluation of student skills. The *Criterion* assessment system comes closer to a systematic evaluation, but the varied procedures by which it has been administered severely compromises the reliability of its results.

These two sources of information about student writing, however, raise at least six issues (listed in Table 5). Each issue (column A), in turn, is related to a potential problem (column B), which then requires one or more solutions (column C). With the exception of inadequate assessment (which has been addressed to some extent in this report) the problems are described as “potential” because they have yet to be systematically documented in this sample of students. However, anecdotal evidence suggests they do exist.

The issues, problems and categories of solutions listed in Table 5 are not mutually exclusive, but they are arranged in the sequence by which the research team recommends they be addressed. For example, the problems of inadequate assessment and apathy must be addressed before it will be possible to accurately determine whether low skill level is truly a problem.

Table 5. Issues, Problems and Solutions

A. Writing Issues →	B. Potential Problems →	C. Categories of Solutions
1. Skill assessment	Inadequate assessment	Improve assessment
2. Motivation	Apathy	Increase motivation
3. Skill level	Low skill level	Develop skills
4. Recency of skill acquisition	Old or stale skills	Refresh skills
5. Feedback frequency	Infrequent feedback	Provide frequent feedback
6. Practice frequency	Few opportunities for practice	Increase opportunities for practice

The research team’s strongest recommendations at this point focus on improving the assessment of student writing. However, recommendations related to the other issues have also been made.

Improving Assessment

Recommendation 1: Standardize the assessment procedure.

The primary weakness in the current assessment procedure is that individual faculty members take different approaches to administering *Criterion*. To improve reliability, the procedure should be identical for all students. This means all students who complete the

Criterion assessment should receive the same instructions, take the assessment under the same physical conditions, use the same word processing software, have the same requirement to take the assessment, and have the same incentive to do so.

The research team recommends that an assessment team (possibly comprised of research team members), rather than course instructors, be in charge of administering the assessment. This approach will standardize the instructions students receive. One class session during the semester should be designated for the assessment, and students should meet in the computer lab on that day. Students should write their essays entirely in *Criterion* rather than cutting and pasting them from other programs. Writing the essay entirely in *Criterion* has at least three advantages: it provides a better assessment of whether students have mastered writing skills (vs. rely on technology to catch errors); it reduces the chance that students will paste in an essay they did not write; and it reflects the actual time students spend writing the essay – a variable that may prove valuable in future analyses. Completing the assessment should be a requirement (for the course or for graduation) for all students. Any additional incentives (e.g., any impact on a course grade) should be the same for all students.

Some instructors' are legitimately concerned that the assessment may have a negative impact on students' perceptions of their teaching, so consideration should be given to ways the assessment might be disconnected from the capstone courses. As one possibility, the capstone courses could simply serve as a tracking mechanism to verify that students have completed the assessment. An "incomplete" grade might be assigned to students who do not complete the assessment, but there would be no other impact on the course grade.

Recommendation 2: Add a human evaluator.

The primary limitation of *Criterion* is that it cannot distinguish whether the content of an essay is related to the designated essay prompt. Adding a human evaluator who, at a minimum, simply determines whether the essays are on topic will substantially enhance the validity of the assessment. This arrangement is noted by ETS^v:

“How do institutions use the Criterion service for assessment purposes?”

Some institutions use the Criterion scores for exit testing — combining a Criterion score with the score from a faculty reader....”

“Educators can stop students from deliberately trying to fool the Criterion service by announcing that a random sample of essays will be read by independent readers.”

A human evaluator could also overcome other limitations in *Criterion*. For instance, twenty students have taken the assessment but have not received scores because *Criterion* was unable to identify certain organizational elements, such as thesis statements and conclusions. Upon closer review, these elements were “missing” because the students had

failed to double-space between paragraphs. A more valuable contribution of a human evaluator would be to provide insights related to specific writing skills that appear weak.

Using a human evaluator in addition to *Criterion* would raise the cost of the assessment, and would provide at least some redundant information. Consideration should be given, however, to the possibility of using a human evaluator *instead of Criterion*. A qualified graduate student, faculty member, or writing professional could provide much richer information regarding students' writing skills, at a potentially lower cost. Developing uniform scoring guidelines would eliminate some, but not all, of the variability across human evaluators.

Increasing Motivation

Recommendation 3: Create incentives for high performance.

Currently, students have little, if any, external incentive to do well on the assessment. Adding a positive incentive to demonstrate a high level of writing ability (e.g., a holistic score of 5 or better) will result in a more accurate assessment of students' true ability levels. Possible incentives include extra credit, a certificate and corresponding resume line item, a "Dean's List", a scholarship for the best essay, entry into a drawing for a prize such as an iPhone, and countless others. Possible negative incentives include an incomplete grade in the course, delayed graduation, required remedial instruction, or a requirement to re-take the assessment until a higher score is achieved.

Recommendation 4: Develop an intrinsically interesting essay prompt.

Students and faculty have commented that the current essay prompt of the role of higher education in modern society is too abstract and reduces students' motivation. A more intrinsically motivating prompt, such as "How will graduating with a business school degree affect your future career?" will create greater focus on the question and ownership of the answer. The possibility of developing a customized essay prompt is outlined on the ETS website:

"...when educators want students to write on a topic not available in the Criterion library, they can create and assign their own prompt for a student assignment. Although essays written on educator-created topics do not receive the holistic score, all of the features of diagnostic feedback will be reported when the essay is submitted. Colleges and universities can also work with ETS to create new topics tailored to their needs."^{vi}

Developing and Refreshing Skills

Recommendation 5: Include systematic writing instruction as part of the business school curriculum.

One of the most commonly suggested ways to improve student writing is to add a business communications course to the business school core. The research team recognizes that developing and offering such a course would require substantial resources – ones that appear unavailable, at least in the foreseeable future. However, teaching writing skills is not an all-or-nothing proposition, and there is no reason to wait until a business communications course is a viable option to begin.

The research team does not make a formal recommendation here as to how such instruction should be provided. The appropriate level of instruction depends on a more accurate assessment of students' existing skills than is currently available. In the meantime, the following ideas are offered to promote future discussions on how writing instruction might be implemented in the business school:

- Add a one-credit writing skills lab to the existing business school writing intensive courses
- Teach core writing skills modules in one or more sessions of the business school writing intensive courses, by the instructor or by a guest
- Develop a required online writing skills course, or adopt an existing course
- Promote self-study by requiring a minimum score on a standardized writing exam (e.g., a test similar to the GMAT verbal section) prior to admission to the business school
- Encourage passive learning by posting grammar questions/answers and “tips of the day” on the hallway monitors

Providing Frequent Feedback and Increasing Opportunities for Practice

The number of opportunities students have for deliberate, informed writing practice reflects the value of effective writing in the business school culture. If at the core of this culture the assumption is “writing is *their* problem” (“they” being the students, the English Department, the secondary education system, or anything else) then the level of writing feedback and the number of opportunities for practice are unlikely to change. If, in contrast, the core belief is “it is *our* responsibility to produce effective writers,” (“we” being the business school faculty) then steps will be taken to provide students with meaningful feedback and practice. No specific solution is recommended in this category; however, all organizations do well to understand their own culture.

Here are a few ways to increase feedback and opportunities for practice:

- Assign more short writing assignments rather than one long assignment
- Require first drafts and subsequent revisions
- Develop a writing scoring rubric to simplify grading. Potentially, a business school-wide scoring rubric could be developed that includes the most common errors in student writing. A companion document could be distributed to students.
- Utilize qualified business school graduate students (or graduate students outside business school) to assist with scoring grammar, usage and mechanics in written assignments. The instructor could still grade for content.

- Establish a business school writing lab

Conclusions

There is no doubt effective writing skills are among the most valuable tools business school students can take with them when they graduate. Assessment is a key to determining the current level of students' writing skills. However, assessment alone leaves several key questions unanswered.

The question of "Why?"

Assessment alone does little to improve students' writing skills. The *Criterion* system provides students with a limited amount of feedback on a short essay (an average length of eighteen sentences), but no chance to make revisions, and in practical terms, no incentive to even review the feedback. Improving the assessment of student writing (the principle recommendations of this report), although important, will do little to directly improve students' skills.

The questions of "Where?" and "When?"

Measuring whether students are able to "conceptualize a complex issue into a coherent written statement" is necessary to assess where the business school stands relative to its formal learning goal. Although an attempt is being made to assess writing skills, these skills are not formally taught in business school courses. Where are they taught? One common assertion is that students should have these skills before they enter the business school. Anecdotal evidence suggests at least some students arrive at the business school without these skills.

Moving the assessment to the beginning of students' coursework, instead of, or in addition to, performing an assessment at the end of their program would allow time for skill deficiencies to be corrected. Some of the recommendations given in this report suggest deficiencies can be corrected in the business school. Alternatively, a minimum level of writing proficiency could be added as a business school admission requirement (a practice not uncommon at other universities). In that case, skill deficiencies could be addressed outside the business school.

The question of "Who?"

Does measuring progress towards the business school learning goal require an assessment of all students? If no individual intervention is planned for low performers, assessing the full population of business school students is an expensive endeavor. The average level of student writing skill can be measured with a much smaller sample. Likewise, if not every student needs remedial writing instruction, requiring all students to participate in a skill development program would be an inefficient use of their time.

This article highlights some of the challenges one business school encountered while addressing the matter of student writing. Several recommendations are offered, but a variety of questions have yet to be answered. To be sure, others across universities face similar

issues. We look forward to further investigation and insight on this important and complex topic.

References

- Bacon, D. R., & Anderson, E. S. (2004). Assessing and enhancing the basic writing skills of marketing students. *Business Communication Quarterly*, 67 (4), 443-454.
- Bacon, D. R., Paul, P., Johnson, C., & Conley, T. (2008). Improving writing through the marketing curriculum. *Journal of Marketing Education*, 30 (3), 217-225.
- Fraser, L. L., Harich, K. R., Norby, J. M., Brzovic, K., Rizkallah, T., & Loewy, D. (2005). Diagnostic and Value-Added Assessment of Business Writing. *Business Communication Quarterly*. 68 (3) 290-305
- McKelvey, R. D., & Zavoina, W. (1975). "A Statistical Model for the Analysis of Ordinal Level Dependent Variables", *Journal of Mathematical Sociology* (4), 103-120.
- Quible, Z. K., & Griffin, F. (2007). Are writing deficiencies creating a lost generation of business writers? *Journal of Education for Business*, 83 (1), 32-36.
- Varner, I. I., & Pomerence, P. J. (1998). Assessing Competency in Business Writing. *Business Communication Quarterly*. 61 (4), 83-91.
- Warnock, S. (2009). Methods and Results of an Accreditation-Driven Writing Assessment in a Business College. *Journal of Business and Technical Communication*, 23 (1), 83-107.
- Wolf, K., Connelly M., & Komare A. (2008). A Tale of Two Rubrics: Improving Teaching and Learning Across the Content Areas through Assessment. *The Journal of Effective Teaching*, 8(1), 21-32.
- Wolf, K., & Stevens E. (2007). The Role of Rubrics in Advancing and Assessing Student Learning. *The Journal of Effective Teaching*, 7(1), 3-14.
- Zhao, J. J., & Alexander, M. W. (2004). The impact of business communication education on students' short- and long-term performances. *Business Communication Quarterly*, 67, 24-40.

Appendix A: *Criterion Diagnostic Feedback Categories*^{vii}

Grammar Errors

- fragment or missing comma
- run-on sentences
- garbled sentences
- subject-verb agreement
- ill-formed verbs
- pronoun errors
- possessive errors
- wrong or missing word
- proofreading errors

Usage Errors

- wrong article
- missing or extra article
- confused words
- wrong form of word
- faulty comparisons
- preposition error
- nonstandard word form
- negation error

Mechanics Errors

- spelling
- capitalize proper nouns
- missing initial capital letter in a sentence
- missing question mark
- missing final punctuation
- missing apostrophe
- missing comma
- hyphen error
- fused words
- compound words
- duplicates

Style Comments

- repetition of words
- inappropriate words or phrases
- sentences beginning with coordinating conjunctions
- too many short sentences
- too many long sentences
- passive voice

Organization and Development (These items are used to help determine the holistic score, but a separate score for organization and development is not assigned. An essay may not receive a holistic score if *Criterion* is unable to identify these items.)

- Thesis statement
- Main ideas
- Supporting ideas
- Conclusion
- Introductory material
- Other
- Transitional words and phrases
- Repetition of ideas
- Topic relationships and technical quality

Appendix B Criterion Scoring Rubric^{viii ix x}

Score of 6: essay presents a cogent, well-articulated analysis of the complexities of the issue and conveys meaning skillfully. A typical essay in this category

- presents an insightful position on the issue
- develops the position with compelling reasons and/or persuasive examples
- sustains a well-focused, well-organized analysis, connecting ideas logically
- expresses ideas fluently and precisely, using effective vocabulary and sentence variety
- demonstrates facility with the conventions (i.e., grammar, usage, and mechanics) of standard written English but may have minor errors

Score of 5: essay presents a generally thoughtful, well-developed analysis of the complexities of the issue and conveys meaning clearly. A typical essay in this category

- presents a well-considered position on the issue
- develops the position with logically sound reasons and/or well-chosen examples
- is focused and generally well organized, connecting ideas appropriately
- expresses ideas clearly and well, using appropriate vocabulary and sentence variety
- demonstrates facility with the conventions of standard written English but may have minor errors

Score of 4: essay presents a competent analysis of the issue and conveys meaning adequately.

- A typical essay in this category
- presents a clear position on the issue
- develops the position on the issue with relevant reasons and/or examples
- is adequately focused and organized
- expresses ideas with reasonable clarity
- generally demonstrates control of the conventions of standard written English but may have some errors

Score of 3: essay demonstrates some competence in its analysis of the issue and in conveying meaning but is obviously flawed. A typical essay in this category exhibits ONE OR MORE of the following characteristics:

- is vague or limited in presenting or developing a position on the issue
- is weak in the use of relevant reasons or examples
- is poorly focused and/or poorly organized
- has problems in language and sentence structure that result in a lack of clarity

contains occasional major errors or frequent minor errors in grammar, usage, or mechanics that can interfere with meaning

Score of 2: essay demonstrates serious weaknesses in analytical writing.

- A typical essay in this category exhibits ONE OR MORE of the following characteristics:
- is unclear or seriously limited in presenting or developing a position on the issue
- provides few, if any, relevant reasons or examples is unfocused and/or disorganized
- has serious problems in the use of language and sentence structure that frequently interfere with meaning
- contains serious errors in grammar, usage, or mechanics that frequently obscure meaning

Score of 1: essay demonstrates fundamental deficiencies in analytical writing. A typical essay in this category exhibits ONE OR MORE of the following characteristics:

- provides little or no evidence of the ability to understand and analyze the issue
- provides little or no evidence of the ability to develop an organized response
- has severe problems in language and sentence structure that persistently interfere with meaning
- contains pervasive errors in grammar, usage, or mechanics that result in incoherence

ⁱ 2010. ETS *Criterion* webpage.

<http://www.ets.org/portal/site/ets/menuitem.435c0b5cc7bd0ae7015d9510c3921509/?vgnextoid=b47d253b164f4010VgnVCM10000022f95190RCRD> (Accessed 03/17/11).

ⁱⁱ 2011. Standard 15 of the Association to Advance Collegiate Schools of Business.

<http://www.aacsb.edu/accreditation/business/standards/aol/standard15.asp>, (Accessed 03/17/11)

ⁱⁱⁱ Ibid.

^{iv} 2011 ETS *Criterion* webpage

http://www.ets.org/criterion/higher_ed/about/faq/ (Accessed 08/04/11).

^v Ibid.

^{vi} 2010. ETS *Criterion* webpage.

<http://www.ets.org/portal/site/ets/menuitem.435c0b5cc7bd0ae7015d9510c3921509/?vgnextoid=b47d253b164f4010VgnVCM10000022f95190RCRD> (Accessed 03/17/11).

^{vii} 2007. The *Criterion* Teaching Guide.

http://www.ets.org/Media/Resources_For/Higher_Education/pdf/Criterion_Teacher_Guide_web_6487.pdf (Accessed 03/17/11).

^{viii} 2003. Burstein, J., Chodorow, M., & Leacock, C. *Criterion* Online Essay Evaluation: An Application for Automated Evaluation of Student Essays. Published proceedings of the fifteenth annual conference on innovative applications of artificial intelligence, held in Acapulco, Mexico. http://www.ets.org/Media/Research/pdf/erater_iaai03_burstein.pdf (Accessed 03/17/11).

^{ix} 2004. Burstein, J., Chodorow, M., & Leacock, C. Automated Essay Evaluation: The *Criterion* Online Writing Service *AI Magazine* 25(3), 27-36.

<http://www.aaai.org/ojs/index.php/aimagazine/article/viewFile/1774/1672> (Accessed 03/17/11).

^x 2010. ETS *Criterion* webpage. http://www.criterionwriting.com.au/6_Point_Scale.html (Accessed 03/17/11).