The Coach: a Web-based Resource for Improving the Writing Skills of Engineering Students

Dr. David M. Beams, University of Texas, Tyler

David Beams earned the B.S. and M.S. degrees in electrical engineering from the University of Illinois at Urbana-Champaign in 1974 and 1977, respectively, and spent 16 years in industrial practice as a design engineer before returning to graduate study. He earned the Ph.D. from the University of Wisconsin-Madison in 1997 and joined the founding faculty of the School of Engineering of the University of Texas at Tyler. Dr. Beams has authored or co-authored a number of papers in engineering education and conference publications in electrical engineering. He is a holder or co-holder of four patents and is a licensed professional engineer in Texas and Wisconsin.

Dr. Luke Niiler, University of Alabama

Dr. Niiler is a member of the English faculty and Director of the Writing Center at the University of Alabama.

Dr. Beth Todd, University of Alabama

Dr. Beth Todd is an associate professor in the Department of Mechanical Engineering at the University of Alabama. She is interested in improving students’ abilities in design and professional skills.

Dr. Marcus Brown, University of Alabama

Prof. Garry W. Warren, University of Alabama

Prof. Warren is in the Dept. of Metallurgical & Materials Engineering and has been teaching for 34 years. His research expertise is primarily in the area of processing, thermodynamics, corrosion and electrochemistry. He has been interested in improving the writing and communication skills of students for many years.
The Coach: a Web-based Resource for Improving Writing Skills of Engineering Students

Abstract

While the opportunities to create writing-intensive courses in engineering are limited, the need for stronger writing curriculum within engineering courses is great. This need has been shown in studies such as the Engineering Writing Initiative (EWI), which tracked the development of engineering students’ writing skills at the University of Texas at Tyler over a four-year period. In that study, the presenters identified two key deficiencies in engineering students’ written communications: rhetorical skills (awareness of audience, purpose, and message) and visual communications (graphs, figures, etc.).

The work begun by EWI continues with The Coach, a collaborative, NSF-funded project at three institutions: the University of Alabama (UA); the University of Texas at Tyler (UT-Tyler), a state-supported regional university and a component of the University of Texas System; and Bevill State Community College (BSCC), a state-supported community college in Alabama. The Coach is developing a series of web-based writing instruction modules and will help students learn to write for audiences of engineers through sequences of writing samples, prompts, and heuristics. By emphasizing writing as a design process, The Coach is intended to provide engineering faculty with a valuable resource for developing students’ rhetorical skills. The Coach’s development is founded upon the understanding engineering curricula are the most-appropriate venue for building stronger engineering writing skills.

Technical issues prevented the planned launch of The Coach in 2011, but roll-out took place on all three campuses in fall, 2012. This paper describes in detail the state of development, curricular introduction, and preliminary assessment of The Coach. Implementation of The Coach is ongoing in spring 2013, and a summary and final evaluation of The Coach will follow once all data have been collected. Preliminary results show a measurable improvement in the quality of engineering writing through use of The Coach; however, this conclusion must be embraced with reservation until further data analysis has been completed.

Background

The recognition that engineers need to be adept at verbal and written communication is not a new insight. In 1955, a report in the Journal of Engineering Education stated that “a high level of performance in the oral, written, and graphical communication of ideas” should be a key part of the engineering curriculum. The immediate genesis of The Coach can be traced to the Engineering Writing Initiative (EWI), a four-year (2004-2008) longitudinal study of the development of writing skills of a cohort of engineering students at UT-Tyler. Reports of work-in-progress and a summary report were made to the Annual Conference of ASEE. EWI, in its turn, was conceived as a response to curricular use of the Electrical Engineering Laboratory Style Guide at UT-Tyler. The chief deficiencies identified by EWI were in rhetoric (awareness of audience, purpose, and message) and in graphical communication. Experience at UT-Tyler with the Style Guide found that it was a useful template but not an
effective teacher; it appeared to be of limited value in addressing the problems identified by EWI.

*The Coach* (named for legendary Alabama football coach Paul “Bear” Bryant) was conceived as an on-line tool to meet the need for improving the rhetorical skills of engineering students. As described in the proposal to NSF, *The Coach* was designed to be a “pilot attempt to create, utilize, and test for efficacy a website that will use a scaffolded and rhetorically-sound sequence of writing samples, writing heuristics, and writing prompts to teach writing to engineering students.”

A brief tour of *The Coach*

Users of *The Coach* enter through an opening page which is reproduced in Fig. 1 below. Although *The Coach* was (and is) intended to support multiple types of written engineering communication, its initial implementation has been limited to the laboratory report (which is the only choice currently listed under “Services”; future versions of *The Coach* will enable instructors to customize tools according to their specific needs). Clicking “Lab Report” brings up a dialog box which allows the user to choose between continuation of an existing report or the creation of a new report.

![Image of The Coach opening screen]({filename})

Fig. 1. Opening screen of *The Coach*.

The sections of a laboratory report composed with *The Coach* are as follows:
Figure 2 shows what the user sees when entering the Cover Page section. The user receives a popup window called “Coach’s Advice” when entering and when navigating away from a new section; this repetition is intended to reinforce strong writing practices. The popup window in Fig. 2 shows the “Coach’s Advice” for the Cover Page.

The content of the “Coach’s Advice” popups was generated through a series of discussions with participating Engineering and writing faculty, who were asked to indicate their preferences for each section of the report. No concern was deemed inconsequential. The Cover Page popup, for example, reminds students to limit titles to two lines; center and bold titles; and use 12-point Times New Roman font.

Prompts are not limited to mechanical concerns, of course. Some prompts are heuristics, or means of encouraging students to speculate and solve the design problems that written communications can present. For example, the Introduction and Background popup reminds students to indicate the reason for the experiment, the conditions in which the experiment took place, and any additional information that may be required for a reader to understand the report.

Popups, in addition, are “voiced” with the character of a coach, who encourages and exhorts through the use of imperatives (“Tell me”), first-person pronouns (“I want to learn”; “I want information”), and second-person pronouns (“You’re almost done!”). The idea here is to provide students with a conception, however tentative, of an active, inquisitive reading audience—an audience to whom they are accountable.

The user may close the popup but may recall it at any time by clicking the “Coach’s Advice” button of the graphical user interface (GUI). The layout shown in Fig. 2 is representative of that of the other sections. The user may navigate from any section to any other; however, the user must visit all sections of the report before the “Submit” button becomes active.
Clicking the button “Coach, what do you think so far?” launches a utility to perform certain syntactical checks of the user’s writing. At present, the principal function of this utility is to look for certain pronouns (particularly first- and second-person pronouns in the accusative and possessive cases). The user’s text is mirrored in a space to the right of the text window and questionable text is highlighted in red. Suggestions for improvement are given above the replicated text. An example of this utility is shown in Fig. 3. Future developments will include what the researchers have termed ‘trigger phrases,’ or redundant constructions (‘red in color’), vague terms (‘very,’ ‘obviously’), and colloquialisms. Such terms and phrases will be highlighted, with alternative forms suggested.

*The Coach* also features sample lab reports from Mechanical Engineering and Electrical Engineering. Users can click on the appropriate buttons to find materials that adhere to the writing standards shown throughout the website.

*The Coach* saves the user’s work on a server at UA when the user moves from one page to another. The user may also save his or her work to this server at any time. There is presently no provision in *The Coach* for local storage. Each laboratory report is given a unique 32-character key when the report is created. This key may be displayed (using the “View Lab Report Key” button) or e-mailed to the user (using the “E-Mail Lab Report Key” button). A typical response of *The Coach* to a request to e-mail a key is depicted in Fig. 4. Recalling a saved report requires this key; at the present time, there is no means in *The Coach* for retrieval of a lost key.
Fig. 3. Use of the “Coach, what do you think so far?” feature.

Fig. 4. Popup window for e-mailing the report key to the user. The user may display the key without e-mailing with the “View Lab Report Key” button. The report key buttons may be accessed from any page at any time.

Files created by *The Coach* are in .html format. A utility for conversion to a Word .doc format is included on the “Submit” page, as shown in Fig. 5. Files are saved to the user’s
local machine with the default file name “Savertf.doc.” Such files may be re-opened locally and printed with Word.

**Submission Information**

Version 0.7.25

Fig. 5. “Submit” screen, including provision for saving the report as a .doc file.

Images added to *The Coach* must presently be in the form of .jpg files. An equation editor is included in *The Coach*, although its capabilities at present are more limited than those of standalone equation editors.

Evaluation methodologies

*The Coach* was first introduced to curricular use in the fall semester of 2012 at UA, UT-Tyler, and BSCC. Representative student work in the form of laboratory reports was collected at all three schools from students who had given written informed consent. This work was presented to a panel of three raters (two at UA and one at UT-Tyler) who evaluated the work via ‘analytic scoring,’ or ‘multiple trait scoring’: that is, according to a five-point rubric. Wolcott and Legg explain this scoring method as an “[examination] the particular—or at least salient—features of compositions in terms of their audience and purpose.” 6 Elbow notes that multiple trait scoring involves “looking hard and thoughtfully at a piece of writing in order to make distinctions as to the quality of different features or dimensions.” 7 And Hamp-Lyons suggests that this scoring method constitutes an acknowledgement of the “complexity of writing” 8 (as cited 6). Evaluation of the effectiveness of *The Coach* was judged from statistical analysis of the data provided by the raters.

Each lab report was evaluated in a double-blind read by three trained, graduate-level evaluators: two from UA, and one from UT-Tyler. To prepare for these evaluations, evaluators were trained, calibrated, or ‘normed’ via double-blind reads of several lab reports organized in the same manner as shown in the Coach: Abstract, Background, Methods and Materials, Results, Discussion, Conclusion, References, Appendices. Raters read according
to the four-point rubric shown below, and were determined to be in ‘agreement’ if their evaluations in any category (a) were identical or (b) differed by a factor of one. A fourth rater (in this instance, one of the co-PIs on the project) assigned a binding evaluation in the event of any discrepancy of two or greater. (It should be noted that the raters from UA were in agreement over 90% of the time, and never varied by a factor of greater than two in any rating).

Throughout the fall semester of 2012, raters at UA met periodically to re-calibrate, in the interest of assuring the consistency and integrity of the rating process. These raters were provided with sample lab reports for evaluation. After rating each draft, each rater was asked to defend her choice of rating in each category shown below. This discussion often proved fruitful, as it pointed out the difficulty of drawing clear distinctions between rating categories. For example, does a poorly designed and executed figure merit a lower rating in terms of quality of detail or quality of tables and figures? Should misplaced (but vital) information be understood as a problem with organization or details? Should the “language” category be understood to include only matters related to field-specific discourse (i.e., command of technical terms, avoidance of jargon), or should it also include more general concerns such as grammar, punctuation, and mechanics? It should be noted that such interpretive concerns (and the attendant discussions addressing them) are par for the course with regard to multiple trait scoring. Faigley et al. have noted the same phenomenon. Wolcott and Legg, further, have noted the difficulty of creating precise categories for evaluation: Broad, generous categories preclude sharp distinctions, while sharp, highly-focused categories create undue complexity. Given these contingencies, regular re-calibration and discussion of ratings remains key to a robust and reliable evaluation process.

In addition to the analysis of laboratory reports, data collection for The Coach included questionnaires at both the beginning and the end of the semester that attempted to assess engineering students’ perceptions and expectations of the role of writing in their curricula and careers. The end-of-semester questionnaire included a section for reactions and comments of those who had used The Coach. Informed consent was not sought for these anonymous questionnaires.

Data collection

At UA and BSCC, various courses tested The Coach. At BSCC, students in sophomore-level physics classes were required to write one report without The Coach (to serve as a control group) and one subsequent report with The Coach (serving as an experimental group). At UA, reports were obtained from 149 students in a large multi-section freshman-level introduction to mechanical engineering. Three sections served as the control group while the remaining three sections formed the experimental group. Informed consent on the part of students was necessary, however, for their work to be used in evaluation of The Coach.

The Coach was introduced at UT-Tyler in ENGR 1201, a freshman-level introduction to the engineering profession that includes simple laboratory exercises in principles of mechanical engineering, civil engineering, electrical engineering, and computer science. At UT-Tyler, students were recruited for either an experimental group that would use The Coach for all
laboratory exercises, or for a control group whose members would write reports as they would if *The Coach* did not exist. All students involved in either group were eligible to receive a synopsis of the evaluations of their work by the raters. The experimental group students were also eligible for a drawing for iTunes gift cards at the end of the fall semester, but no material inducements were offered to the control group. Ten students took part in the experimental group; 22 were involved in the control group.

Faculty members teaching ENGR 1201 accommodated the trial use of *The Coach* by agreeing to mandate the laboratory report format of *The Coach* for all assignments and to require reports to be submitted in hard-copy form. Instructors in ENGR 1201 turned over all laboratory reports to the faculty member responsible for *The Coach* at UT-Tyler, but only those reports submitted by participants were scanned to .pdf form. All reports were then returned for normal grading. This process was meant to preserve the anonymity of those participating in *The Coach* and to prevent any potential bias in grading. Scanned reports were redacted to remove names and to assign each report a random code. The scanned work was then distributed to the raters.

A similar process of removing names and assigning random codes to each report was carried out at UA and BSCC.

The data-collection process at UA and BSCC had an advantage compared to that of UT-Tyler in that the same cohort of students was involved in producing reports with and without *The Coach*. The method of UT-Tyler, however, had an advantage in that reports were received from both control and experimental groups at the same time, removing experience as a possible confounding factor in the interpretation of the data.

Evaluation rubric

Each rater scored each of the following traits on each report using a scale of 1 (strong disagreement) to 5 (strong agreement):

- **Language:** The writer demonstrates a command of professional language.
- **Tables and Figures:** The writer uses tables and figures appropriately.
- **Organization:** The writer's draft is properly organized.
- **Detail:** The writer uses an adequate level of detail.

These criteria represent a version of the rubric used at UT-Tyler as part of the Engineering Writing Initiative. Similar criteria have been used by engineering programs at the University of Arizona, the University of South Florida, and the University of Washington.¹⁰, ¹¹, ¹²

Preliminary results: evaluation of laboratory reports

At this writing, data analysis for the fall semester of 2012 is incomplete; data for UA have not yet been compiled; the following results reflect the work of UT-Tyler and BSCC. A total of 70 control-group reports and 26 experimental-group reports were received from UT-Tyler; 17 control-group reports and 13 experimental-group reports were received from BSCC.
The charts of Figs. 6–9 below show the mean ratings for each of the four traits assessed by each of the evaluators. The columns are paired by raters; the leftmost pair reflects the evaluations of the UT-Tyler rater, while the center and rightmost pair reflect the evaluations of UA raters 1 and 2, respectively. The label “(Ctrl)” denotes evaluation of control-group reports; “(Exp)” denotes experimental group. The dark band in each column reflects the mean; the extent of each column reflects ± one standard deviation about the mean. The ordinate range is from 1 to 5; columns were truncated at 5.0 if the sum of the mean and one standard deviation exceeded 5.0.

Figure 6 shows the results for professional language. The column “UTT( Exp)” is truncated at 5. All raters’ results indicate improvement in the experimental group; the rater from UT-Tyler shows an improvement in the mean of 0.56; raters 1 and 2 from UA showed an improvement of 0.79 and 0.61, respectively.

![Language Chart](image)

Fig. 6. Evaluation of control- and experimental-group laboratory reports for professional language. A rating of 5 indicates strong agreement that reports reflect a professional command of language; 1 indicates strong disagreement.

Evaluation for appropriate use of tables and figures is shown in Fig. 7. All raters’ results indicate improvement in the experimental group; the rater from UT-Tyler shows an improvement in the mean of 0.35, and both raters from UA showed an improvement of 0.48.
Fig. 7. Evaluation of control- and experimental-group laboratory reports for appropriate use of tables and figures. Both columns of the rater from UT-Tyler have been truncated at 5. A striking difference is seen between the UT-Tyler rater and the UA raters.

Figure 8 gives the results for organization. The rater of UT-Tyler shows an improvement of 0.64 from control to experimental groups; rater 1 of UA recorded an improvement of 0.44, while rater 2 recorded an improvement of 0.49.

Fig. 8. Evaluation of control- and experimental-group laboratory reports for proper organization. Both columns of the rater from UT-Tyler have been truncated at 5.
Figure 9 gives results for an adequate level of detail. The rater of UT-Tyler shows an improvement in the mean of 0.11 from control to experimental groups; rater 1 of UA found an improvement of 0.37, while rater 2 recorded an improvement of 0.53.

Fig. 9. Evaluation of control- and experimental-group laboratory reports for adequate level of detail. The control-group column of the rater from UT-Tyler has been truncated at 5.

Table 1 below summarizes results of all traits and all raters. The control group consisted of 88 independent laboratory reports with a total of 261 evaluations. The experimental group comprised 39 independent reports with a total of 104 evaluations. The experimental group showed improvement in all categories, but further statistical tests to assess significance of these differences is pending at this writing.

Table 1. Summary of all traits and all raters for control and experimental groups.

<table>
<thead>
<tr>
<th>Trait</th>
<th>Language Mean</th>
<th>Tables and figures Mean</th>
<th>Organization Mean</th>
<th>Detail Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean, control</td>
<td>2.88</td>
<td>2.05</td>
<td>2.86</td>
<td>2.81</td>
</tr>
<tr>
<td>Mean, experimental</td>
<td>3.47</td>
<td>2.88</td>
<td>3.55</td>
<td>3.38</td>
</tr>
<tr>
<td>Δ (experimental – control)</td>
<td>+0.59</td>
<td>+0.83</td>
<td>+0.69</td>
<td>+0.57</td>
</tr>
<tr>
<td>σ control</td>
<td>1.21</td>
<td>1.32</td>
<td>1.28</td>
<td>1.20</td>
</tr>
<tr>
<td>σ experimental</td>
<td>1.07</td>
<td>1.20</td>
<td>0.98</td>
<td>1.12</td>
</tr>
</tbody>
</table>

Discussion

The data presented in Figs. 6-9 show a striking degree of inter-rater reliability between the two raters of UA. They also show a consistent positive bias on the part of the rater from UT-Tyler. Table 2 presents comparative aggregate statistics for the three raters for all reports.
The degree of consistency between the two raters of UA is confirmed in these data, as is the bias of the UT-Tyler rater.

Table 2. Comparative statistics of raters for all reports. The high degree of inter-rater reliability of the two University of Alabama raters is evident.

<table>
<thead>
<tr>
<th></th>
<th>Language</th>
<th>Tables and figures</th>
<th>Organization</th>
<th>Detail</th>
</tr>
</thead>
<tbody>
<tr>
<td>UT-Tyler, Mean</td>
<td>3.57</td>
<td>3.69</td>
<td>3.88</td>
<td>3.65</td>
</tr>
<tr>
<td>UA1, Mean</td>
<td>2.82</td>
<td>2.26</td>
<td>3.01</td>
<td>2.95</td>
</tr>
<tr>
<td>UA2, Mean</td>
<td>2.81</td>
<td>2.13</td>
<td>2.96</td>
<td>2.94</td>
</tr>
<tr>
<td>UT-Tyler, σ</td>
<td>1.52</td>
<td>1.49</td>
<td>1.46</td>
<td>1.35</td>
</tr>
<tr>
<td>UA1, σ</td>
<td>0.93</td>
<td>0.84</td>
<td>1.02</td>
<td>1.07</td>
</tr>
<tr>
<td>UA2, σ</td>
<td>0.95</td>
<td>0.89</td>
<td>0.92</td>
<td>1.00</td>
</tr>
</tbody>
</table>

The improvement in tables and figures of the experimental group relative to the control group is somewhat curious given that The Coach offers no explicit guidance related to these. It is also the trait that showed the greatest divergence between the rater of UT-Tyler and those of UA. “Appropriate use of tables and figures” may be open to quite-different interpretations. As an example, one report that the rater from UT-Tyler rated as 5 (strong agreement that appropriate use was made of tables and figures) was rated as 1 by both raters from UA. Review of this report showed two photographs without captions or annotation. It is possible that the rater from UT-Tyler thought that any use of photographs was appropriate, but in the judgment of the experiments, the UA rating was closer to the truth. Analysis of other similar discrepancies may give some insight into the thinking of the raters; further guidance to the raters has been provided prior to the start of data collection for Spring, 2013.

It may be asked why more improvement in organization was not noted given that The Coach is perhaps best suited to model this trait. However, as noted above, the basic laboratory report structure of The Coach was required for all reports in this study so that the raters could not distinguish a priori reports from the control group from those of the experimental group. The imposition of this basic structure on the control group may have reduced the evident effect on organization of The Coach.

The scope of The Coach is presently limited to laboratory reports. However, the repertoire of document types must be expanded for The Coach to achieve its intended impact. Manion and Adams indicated that the laboratory report is an atypical document; it does not successfully replicate the type of document that students will encounter in industry [6].

Preliminary results: student surveys

Students were surveyed at cooperating schools regarding their experiences with The Coach at the end of the semester. They were presented with a number of statements and propositions about The Coach and asked to indicate the degree to which they agreed with each by assigning a response of 1 to 5, with 1 indicating strong disagreement and 5 indicating strong
agreement. Table 3 below summarizes 31 individual responses from the three participating schools.

Table 3.  Student perceptions of *The Coach* from the experimental groups for Fall, 2012. The evaluation scale is 1 (strong disagreement) to 5 (strong agreement).

<table>
<thead>
<tr>
<th>Statement</th>
<th>Median score</th>
</tr>
</thead>
<tbody>
<tr>
<td>It helped me to write a better report than I would have written without it.</td>
<td>4</td>
</tr>
<tr>
<td>The instructions on how to start were clear and easy to understand.</td>
<td>5</td>
</tr>
<tr>
<td>It provided me with useful suggestions on how to write reports.</td>
<td>3</td>
</tr>
<tr>
<td>The information provided by <em>The Coach</em> on how to write was helpful.</td>
<td>4</td>
</tr>
<tr>
<td>The comments and suggestions got in the way of completing the writing assignments.</td>
<td>2</td>
</tr>
<tr>
<td>The program was &quot;glitchy.&quot;</td>
<td>3</td>
</tr>
<tr>
<td>I learned some writing techniques that will help me in future writing assignments.</td>
<td>4</td>
</tr>
<tr>
<td>I would prefer to NOT use <em>The Coach</em> in future writing assignments for engineering classes.</td>
<td>2</td>
</tr>
<tr>
<td><em>The Coach</em> could be a good program to use in other engineering courses that have writing assignments.</td>
<td>3</td>
</tr>
<tr>
<td>I found the writing samples provided by <em>The Coach</em> to be useful.</td>
<td>3</td>
</tr>
</tbody>
</table>

A few observations may be made about these data. There was overall a propensity to agree that *The Coach* improved their ability to construct a laboratory report and a clear agreement to the proposition that the instructions were clear and easy to use. There was, however, evidence of carelessness (or malice) on the part of two individuals taking the survey who returned surveys in which all responses were 1 (strongly disagree). It seems incongruous that an individual could strongly disagree with “it helped me to write a better report” while also disagreeing strongly with “I would prefer to NOT use *The Coach* in future writing assignments for engineering classes.”

It is expected that a clearer picture of student reactions to *The Coach* will emerge when final data from all become available after the Spring, 2013 evaluation cycle. In addition to responses to the prompts shown above, these data will include control and experimental group responses to a “Writing Attitude Questionnaire,” based in part on the work of McAndrew and Reigstad\(^\text{13}\). In this questionnaire, students are asked about their writing attitudes and practices both prior to and after using the Coach.

Continuing work

Data collection is continuing at all three cooperating schools in Spring, 2013. UT-Tyler has hired three new raters. UT-Tyler has changed its data-collection procedures to mimic those of UA and BSCC; all ENGR 1201 students will participate in writing both control reports and experimental reports.

Conclusions
Preliminary analysis of The Coach appears to show improvements in the quality of written reports, but further data collection and analysis are required before a firm conclusion may be made. Ongoing improvements to The Coach will continue to be made in response to student experience and in conformity with its goals.

Bibliography


8 L. Hamp-Lyons, “Uncovering Possibilities for a Constructivist Paradigm for Writing Assessment (Review).” *College Composition and Communication*, 46, 446-455.


