

Exploring Alternative Methods of Evaluating Written Communication Skills in the Admission Process

Abstract

Purpose: With large numbers of candidates applying to physician assistant (PA) programs, it is essential that the process be streamlined while maintaining the highest standards. To this end, a prospective, cross-sectional, comparative study investigating alternative methods that may be used to evaluate candidates' written communication skill was conducted. **Methods:** Candidates were required to write an onsite essay. The scores of the onsite essay and an essay previously submitted by the candidate to the Central Application Service for Physician Assistants (CASPA) were compared. Community PAs evaluated the CASPA essays. Onsite essays were evaluated by Microsoft Readability Studies, automated essay scoring (AES) by Vantage Learning's IntelliMetric® systems, and community PAs. Further, the different methods of evaluation were also compared to pre-admission grade point average (GPA). **Results:** The level of statistical significance was set at $p < 0.05$. There was a difference between the scores of the CASPA and onsite essays evaluated by community PAs using a programmatic rubric ($z = -5.025, p < 0.01$). A difference was found between the scores of the onsite essays evaluated by the IntelliMetric® AES system and community PAs utilizing a hardcopy of the IntelliMetric® rubric ($z = -7.542, p < 0.01$). IntelliMetric® AES evaluations positively correlated to pre-admission GPA ($R = 0.307, p < 0.01$). The CASPA essay scores inversely correlated to preadmission GPA ($R = -0.260, p < 0.05$). **Conclusion:** Onsite essays should be used to evaluate candidates' written communication skills. In order to adhere to admission process deadlines and due to the results of this study, automated essay scoring should be considered as an option when evaluating candidates' writing skills.

Evaluation of Candidates' Written Communication Skills

Introduction

Physician assistant (PA) faculty have responsibility for selecting a limited number of candidates from a typically large, highly qualified applicant pool for admission into PA programs.¹ Selection decisions have broad implications. This is true not only for the applicants and for the success of the program, but for physicians who later supervise these individuals as well as the patients at large, on whose care they will depend.

Cognitive or quantitative variables such as pre-professional grade point average (GPA) and standardized test scores are known predictors of success for applicants seeking admission into and graduating from many healthcare programs.²⁻⁶ Non-cognitive abilities, such as oral and written communication skills, are less consistent in predicting the same success. In large part, the degree to which the non-cognitive factors identify a successful candidate depend on the outcome being considered (e.g., attrition, deceleration from program, board examination passing rates, or functioning as a successful practitioner after graduation).⁷⁻⁹ Additionally, few PA programs have specific writing courses to help correlate more directly pre-admission writing samples. Notwithstanding, many admission committees believe that both cognitive and non-cognitive factors are important in the selection of applicants and in the future success of the candidates in their chosen healthcare field.^{4, 10-13}

Currently, to gain admission to the Butler University Physician Assistant Program ("Program"), individuals are required to submit an application to the Central Application Service for Physician Assistants (CASPA) that includes the candidate's demographics, academic record, experience in healthcare, and an essay on a topic that was developed by the Physician Assistant Education Association. According to CASPA, this topic does not

Evaluation of Candidates' Written Communication Skills

change on a yearly basis (A. Donnelly, Manager, Application Services, [adonnelly@liaison-intl.com], e-mail, June 6, 2008). While many standardized tests include a separate graded writing sample (e.g., Graduate Record Exam), the Program does not utilize a standardized test for admission. Hence, the essay submitted to CASPA is evaluated as the applicant's writing sample.

The Program recruits a pool of community PAs to participate in the review and evaluation of the candidates' CASPA essays. The program strives to have two PAs evaluate each essay utilizing a Program-developed Likert scale rubric. Nevertheless, scores derived from the evaluation of a single PA is considered acceptable. The number of essays that a PA evaluates each year differs based on the number of applications received and number of community PAs recruited. This past admission cycle, 521 essays were evaluated and twenty-six volunteers recruited. The effort required to coordinate and evaluate the essay review is significant. With the increasing number of applications received each year, the current process is becoming daunting.

Because the essays are prepared in advance by the applicants and submitted to CASPA, it is not known with complete certainty whether an applicant independently writes the essay. Further, since the CASPA essay topic largely remains unchanged year to year, recycled essays may be passed from candidate to candidate and resubmitted to the program. One study assessing the personal qualities of applicants who matriculated into medical schools concluded that over 44% of applicants had structural input from others and 51% had input on content development of their personal essays.¹⁴ This raises an important question: to what extent does this system actually assess the *applicant's* writing ability?

Evaluation of Candidates' Written Communication Skills

The current process is staff and time intensive and depends on recruitment and use of busy physician assistants in the community. Compounding this issue is the large number of applications received immediately prior to the deadline. This is a significant problem for the Program and for applicants who are making travel arrangements— in some cases from very far distances— to participate in the interview on very short notice. Changing deadlines for the application process is possible, albeit difficult, and likely would not result in the realistically needed increased time to adequately review and receive the increased number of essays back from community PAs.

The Program has been successful in selecting capable students who regularly achieve above average national board scores and who, in the past have received positive reviews from physicians employing them (D. Frosch unpublished programmatic data, 2008). However, to date the Program has identified only GPA as an outcome marker for success for the established admission criteria (J. Holycross, unpublished data, 2001). Assessing optimal ways to evaluate a candidate's written communication skills continues to be worthy of examination.

Computerized Scoring

By 1997, Microsoft Office[®] incorporated grammar checking as a tool for users of the software.¹⁵ One of the pioneers in grammar checking formulas, specifically readability tests, was Rudolf Flesch. His work has become a standard component in various word processing programs including Microsoft Word[®], which has both the Reading Ease and Flesch-Kincaid Grade Level tests. A Flesch Reading Ease score rates text on a 100-point scale. Essays that are easier to read are rated higher (closer to 100), while harder to read essays are rated lower (closer to 0). Table 1 ranks the ease of reading

Evaluation of Candidates' Written Communication Skills

associated with a given score. The Flesch-Kincaid Grade Level rates text on a U.S. school grade level (e.g. a score of 6.0 means that a sixth grader can understand the document).¹⁶ Both of these rudimentary evaluation tools base both ratings on formulas that utilize the average number of syllables per word and words per sentence.

Table 1 Flesch Reading Ease Score¹⁶

Reading Ease Score	Description of Style
0-30	Very Difficult
30-50	Difficult
50-60	Fairly Difficult
60-70	Standard
70-80	Fairly Easy
80-90	Easy
90-100	Very Easy

In the article “Trait Ratings for Automated Essay Grading,” it is reported that grammar-checking software effectively evaluates essays between 500 to 1000 words covering a wide range of topics. The authors go on to state that the “relationship between essay length and essay ratings tends not to be linear but rather curvilinear;” i.e., length is important up to a point but then negligible.¹⁷

Advances in Software Technology

While some forms of writing may not be effectively analyzed by software, “approximately 90 percent of required writing in a typical college classroom can be evaluated using Automated Essay Scoring (AES).”¹⁸ In addition to spelling and grammatical standards, AES utilizes professionally trained writers to develop a series of essays that represent each applicable grading level. The series of essays serve as multiple “model answers” to the software. The writing sample of each applicant is compared to these standards and a score for the candidate is determined.¹⁹ The automated system applies criteria uniformly and mechanically, avoiding the fluctuations found in untrained

Evaluation of Candidates' Written Communication Skills

graders. Although mechanical uniformity has appeal, the value is limited by the extent to which the resulting scores accurately measure the quality of the essay. According to company-supported research, AES is superior to human raters because its test retest reliability is perfect, and it may avoid biases that occur with human evaluators.²⁰

Additional benefits of the AES systems may include the ability to overcome time constraints in evaluating essays, the costs associated with the process, reliability in the outcome, and generalizability issues in writing assessments.²¹

Vantage Learning's IntelliMetric[®] system is one AES product that "is grounded in a cognitive model" with "information processing and understanding" to evaluate essays.²² The IntelliMetric[®] system makes judgments incorporating mathematical models that report a holistic score as well as individual scores for five domains (Table 2) using a 6-point Likert scale. The holistic score is not simply a summation of the five domains but a computer generated score evaluating the essay's overall performance.

Table 2. Vantage Learning IntelliMetric[®] Rubric Domains

Domain ¹⁸	Area of Evaluation ²¹
Focus and Unity	Is there a main idea, and is it consistently supported?
Development and Elaboration	Are the supporting ideas varied, well developed, and elaborative?
Organization and Structure	Does the essay logically transition ideas from introduction, supporting paragraphs, and conclusion?
Sentence Structure	Is there syntactic complexity and variety?
Mechanics and Conventions	Does the essay follow rules of standard American English?

There are some common criticisms of AES software such as IntelliMetric[®]. First, it is possible to respond to an essay question using appropriate keywords and synonyms

Evaluation of Candidates' Written Communication Skills

but the essay may still lack a comprehensible answer.¹⁷ It has been countered that if an individual is capable of writing an essay with appropriate keywords and synonyms in an organized, grammatically correct fashion, they are likely capable of producing a comprehensible essay. A second criticism of AES software is that it requires a great deal of effort to write multiple model answers to essay topics to “train” the software that properly grades the writing samples.²¹ This is a well-founded concern, as each newly developed essay topic requires significant numbers of essays to serve as models at each grading level and to test the software to ensure reliability. In addition, some critics wonder if it is possible for a computer to “artificially think” to generate reliable evaluations of writing samples. A number of studies do suggest the reliability of the AES systems; however, only a few of these studies are independent of company support. The use of advanced technology to assist faculty in evaluating essays in admission decisions should be studied in an unbiased manner.

Null Hypotheses

This study seeks to consider the following null hypotheses:

1. Utilizing a programmatic rubric, there is no difference between the scores of the CASPA and onsite written essays.
2. Utilizing the IntelliMetric[®] rubric, there is no difference between the scores of onsite essays evaluated by the AES system and community PAs.
3. There is no correlation between the candidates' scores evaluated by the seven methods of onsite essay evaluation and GPA.

Methods

Evaluation of Candidates' Written Communication Skills

Of the 521 applicants in the most recent admission cycle, the top 90 were selected for interviews using the Program's standard evaluation process. A twenty-five minute, onsite written essay was then required of each candidate as part of the interview process. The applicants drafted the essay on a university computer. Unbeknownst to the candidates, the essay was not taken into account in the ranking process; it was used solely to facilitate the study of the comparison of human-scored essays and machine-scored essays. The topic chosen for the onsite essay was non-medical and pre-developed by Vantage Learning to avoid the significant time, effort, and cost required to set up new standard answer models for the IntelliMetric[®] system. The study protocol was reviewed by Butler University's institutional review board involving human subjects and approved as exempt. Ten CASPA essays were evaluated only once by a community PA, and therefore were not included in the CASPA essay evaluation. Two of the 90 candidates did not submit an onsite essay. Completed onsite essays were reviewed by a faculty member to excise any identifying names or dates, and were assigned random identification numbers. To ensure uniformity, all essays were reduced to single-spaced documents.

To evaluate the effectiveness of community PAs and the AES system, six fabricated essays were developed by the authors of this study and used as controls. These fabricated essays included: two essays that were well-written but in response to a different essay topic; one essay with simple repetition of the topic; one essay consisting of four sentences written in response to the essay topic and then simply repeated in subsequent paragraphs in different sequences; one essay with the initial half consisting of a well-written response to the essay topic, and the second half consisting of a simple restatement of the essay topic (not a true response to the topic); and one well-written essay that responded to

Evaluation of Candidates' Written Communication Skills

the topic. The single-spaced fabricated essays were about the average length of the candidates' essays and randomly assigned identification numbers in the same manner, making them indistinguishable in appearance from the candidates' essays. While the IntelliMetric® license fee was reduced, the study was conducted independent of Vantage Learning, the licensor of IntelliMetric®.

For consistency, the PAs who assessed the onsite and fabricated essays were from a group of community PA volunteers who had reviewed CASPA essays in the past. Each onsite essay was evaluated by two PAs using the programmatic rubric and by two other PAs using a hard copy of the Vantage Learning IntelliMetric® rubric. Each community PA used only one rubric type: either the programmatic rubric or the IntelliMetric® rubric. Permission was obtained from Vantage Learning to utilize the IntelliMetric® rubric for purposes of this study. If the scores of the two evaluators differed by 50% or more of the total score, a third evaluator was used and the score adjusted to be the average of the three scores.

As a means of rudimentary comparative analysis, onsite and fabricated essays were evaluated by Microsoft Word® version 2003 to obtain the Flesch Reading Ease, Flesch-Kincaid Grade Level, and Word Count. De-identified, random numbered, onsite and fabricated essays were electronically submitted for automated scoring utilizing the IntelliMetric® system, which provided domain and holistic scores for each essay. The holistic scores were not used in this study since the exact nature of how these scores are obtained is based on complex artificial intelligence interactions. Further, a holistic score was not requested of the community PAs. For each essay, domains were totaled together to obtain a final score. These summative values were used for statistical analysis because

Evaluation of Candidates' Written Communication Skills

they more closely approximated the scoring of the programmatic rubric that was used with the CASPA essays. Once results were received, data were maintained on an Excel[®] spreadsheet and statistical analyses performed using Statistical Package for the Social Sciences (SPSS), version 15.

The seven methods of evaluation are listed in table 3. A post hoc power analysis was performed. To determine if there was a statistically significant difference between the ranked difference scores between the CASPA and onsite essays utilizing the programmatic rubric and community PAs, a Wilcoxon Signed Rank test was utilized. To determine if there was a statistically significant difference in scoring totals between the onsite essays evaluated by the AES and community PAs using the IntelliMetric[®] rubric, a Wilcoxon Signed Rank test was performed. Further, a Spearman Rank correlation coefficient value was calculated comparing GPA and the seven methods of evaluation.

Table 3. Methods of Essay Evaluation

Essay Origin	Evaluator	Evaluation Tool Utilized
CASPA	Community PA	Programmatic Developed Rubric
Onsite	Microsoft Word [®]	Flesch Reading Ease
Onsite	Microsoft Word [®]	Flesch Kincaid Level
Onsite	Microsoft Word [®]	Word Count
Onsite	Community PA	Programmatic Developed Rubric
Onsite	Community PA	IntelliMetric [®] Rubric
Onsite	Vantage Learning	IntelliMetric [®] Rubric

Results

The descriptive statistics of the methods of evaluation for the candidates' essays were calculated (Table 4). A post hoc power analysis was performed utilizing the Wilcoxon Signed Rank and Spearman Rank correlation tests (Table 5). One hundred

Evaluation of Candidates' Written Communication Skills

percent power analyses with large effect sizes were obtained for these studies, thereby decreasing the chance of a Type II error.

Table 4. Descriptive Statistics for Methods of Evaluation, N = 88

Method of Evaluation	Possible Range	Mean	Median	S.D. (+/-)	Range
CASPA Essay [^]	0 - 10	8.48	8.5	1.26	2 - 10
Word Count	∞	357.93	349	107.61	142 - 687
Flesch Reading Ease	0 - 100	58.90	59.75	9.27	40.4 - 78
Flesch Kincaid Level	Grade Level	9.46	9.3	1.93	5.9 - 14.2
AES	5 - 30	15.44	15.00	4.11	5 - 25
Onsite Community PA Programmatic Rubric	0 - 10	7.15	7.5	1.61	3 - 10
Onsite Community PA IntelliMetric [®] Rubric	5 - 30	21.57	21.50	3.86	8 - 30

[^] N = 78

Table 5. Post Hoc Power Analysis

	N =	Power %	Effect Size
Wilcoxon			Cohen's d
CASPA vs. Onsite Community PA Programmatic Rubric	78	100	0.92
AES vs. Community PA IntelliMetric [®] Rubric	88	100	1.54
Spearman (vs. GPA)			Correlation r
CASPA Essay	78	100	0.94
Word Count	88	100	0.92
Flesch Reading Ease	88	100	0.97
Flesch Kincaid Level	88	100	0.91
AES	88	100	0.90
Onsite Community PA Programmatic Rubric	88	100	0.84
Onsite Community PA IntelliMetric [®] Rubric	88	100	0.96

There was a statistically significant difference between the ranked difference scores between the CASPA and onsite essays by utilizing the programmatic developed rubric and community PAs ($z = -5.025$, $p < 0.01$). The Wilcoxon Signed Rank test showed a significant difference in the range of score frequencies; therefore, the hypothesis of no difference is rejected.

Evaluation of Candidates' Written Communication Skills

There was a statistically significant difference of the totaled scores between the onsite essays evaluated by the community PAs utilizing the IntelliMetric[®] rubric and the AES outcome using the Wilcoxon Signed Rank test with a $z = -7.542$, $p < 0.01$; therefore, the hypothesis of no difference is rejected. A Spearman Rank correlation coefficient was calculated for the seven methods of evaluation compared to GPA (Table 8). The CASPA essay was inversely correlated and the AES and word count positively correlated at a statistically significant level compared to GPA; therefore, the hypothesis of no relationship between these methods of evaluation and GPA is rejected. For the remaining four methods of evaluation, the null hypothesis was retained.

Table 8. Spearman Rank Correlation Coefficient of Essay Scores Evaluated by Different Methods and GPA, N = 88

Correlation	Spearman Coefficient	Significance
CASPA Essay [^]	-0.260	0.022*
Community PA Programmatic Rubric	0.076	0.479
Community PA IntelliMetric Rubric	0.170	0.112
AES Scoring	0.307	0.004*
Word Count	0.237	0.026*
Flesch Reading Ease	-0.067	0.536
Flesch Kincaid	0.122	0.257

[^] N = 78; *p is significant < 0.05

Discussion

Due to the large number of essays received prior to the deadline for invitations to candidates to participate in interviews, only one community PA evaluated ten CASPA essays in the top 90 candidates. This decreases the reliability of those essay scores.

Therefore, in our study, these scores were removed from the evaluation of the CASPA essays. Two candidates completed the interview process without having submitted onsite essays. It is unclear how this might have occurred; however, because all students who

Evaluation of Candidates' Written Communication Skills

participated in the onsite essays signed and submitted a statement testifying the essay was their own, we can only surmise the students did not “sit” for the timed essay.

The five of six fabricated essays were identified in every instance by the IntelliMetric® System; they were “thrown out” and not evaluated. The sixth fabricated essay that was well written and responded to the topic was scored as expected. The same was not true of the community PA evaluators. One community PA, utilizing the programmatic rubric, evaluated an essay as 7/10, even though it was one of the fabricated essays written using four sentences on topic with subsequent repetition of those four sentences in a different sequence. This was within one and a half points of the mean of all evaluated essays, raising concern of validity among the community PAs' evaluations.

There may have been a few reasons why a statistically significant difference was found between the summative scores on candidates' essays from CASPA and those written onsite by utilizing the programmatic developed rubric and community PAs. The students may have been incapable of composing a written response to the onsite essay as they had for the essay prepared in advance for CASPA because they felt pressured or constrained by time. However, timely or pressured communications are often the nature of a PAs professional responsibility; therefore, should we not select for this quality in our applicant as well? In addition, it is unclear (as found in previous studies) if or to what extent candidates received help in developing the prepared essay's content, grammar, or spelling. The onsite essay significantly eliminates doubt regarding the origin of the essay and is an essential step in actually assessing the *applicant's* writing ability.

There was a statistically significant difference of the scores between the onsite essays evaluated by the community PAs utilizing the IntelliMetric® rubric and the AES

Evaluation of Candidates' Written Communication Skills

scores by the Wilcoxon Signed Rank test with a $z = -7.542$, $p < 0.01$. Not only were the scores different, the community PAs' mean average rating was higher in 82 of the 88 essays compared to the AES system.

The Spearman Rank correlation was to evaluate a possible relationship between GPA and the candidates' individual totaled essay scores. GPA was chosen, as it is the only admission criterion with a known positive predictive outcome for the Program with a first time pass on the national certification examination (J. Holycross, unpublished data, 2001). Automated Essay Scoring ($R = 0.307$, $p < 0.01$) and Word Count ($R = 0.237$, $p < 0.05$) proved statistically significant and correlated positively with GPA. The CASPA essays negatively correlated with GPA ($R = -0.260$, $p < 0.05$). As previously reported, essay length is important to a certain number of words so that concepts and ideas may be developed; however, beyond this point, the essay length does not add to the positive outcome of the essay. The candidates likely did not exceed the reported curvilinear point of an effective essay length; therefore, those candidates with larger word count essays were likely able to develop better essays. It seems reasonable to assume that an individual who has a higher GPA is likely able to write an essay more effectively than one with a lesser GPA. While this assumption may not always be true, the AES system demonstrated a statistically significant positive correlation with those candidates and GPA whereas with the community PA evaluations, the statistically significant opposite was true, which is a point concern.

There are many limitations associated with this study. Because other institutions were not used in this study, the generalizability of the findings is limited to the Program. In an attempt to validate this study, the analysis compares the AES scoring from

Evaluation of Candidates' Written Communication Skills

IntelliMetric[®] to an admittedly imperfect, yet successful process using community PA evaluators. Even though standardized writing tests do a fairly good job of predicting first year GPA in all ethnic groups and particularly for women in undergraduate institutions (ranging from .31 to .51)²³, limited studies exist evaluating graded writing samples used in the admission process and writing success in graduate level programs.²⁴ While pre-admission GPA is a predictor of a positive outcome in passing the national certification exam for the Program, it is unknown if it is valid to assume GPA positively correlates with the candidates' writing ability. Comparing the pre-admission essay to a writing course in the PA curriculum would be ideal; however, no such class exists where writing capability is solely evaluated within the Program's curriculum.

Future studies should consider challenging all of the methods of evaluation for intrarater reliability by submitting two of the same essays with different identification numbers to determine if the grading outcome would be the same. This would allow the test retest capabilities to be analyzed. Future studies should determine the relationship between the onsite essay AES score and successfully completing the PA program without deceleration and first pass on the national certification examination. Further, other studies could compare various methods of evaluation against trained evaluators to validate the method.

Further, the future studies should consider utilizing two, twenty-five minute timed essays for reasons of reliability and construct validity.²⁴ If onsite essays and automated essay scoring are implemented, directions to encourage candidates to write essays 500 – 1000 words in length should be recommended. Finally, the studies should consider

Evaluation of Candidates' Written Communication Skills

investigating students' comfort levels and test anxiety between computerized and paper-and-pencil writing test by age, gender, and ethnicity.

Conclusion

The purpose of this study is to support that there may be a more effective and reliable way to evaluate the writing skills of candidates for admission to the PA program than the utilization of non-trained volunteer community PAs. Questions exist as to whether the current, labor-intensive process of essay review by these PAs is a reliable process. There is uncertainty about the source of the essay itself. Incorporation of AES into the admission process may reduce the time normally spent waiting for community PAs to evaluate the essays and may potentially increase the reliability of the essay scoring.

Acknowledgments

The authors wish to thank the Mr. Jeffrey McGregor, Information Resources, Butler University, for technical support that was given to make the drafting of onsite essays possible. The authors would also like to thank the community PAs who participated in the evaluation of the onsite essays.

Evaluation of Candidates' Written Communication Skills

References

1. Accreditation Review Commission on Education for the Physician Assistant Standards of Accreditation A2.05b. <http://www.arc-pa.org/Standards/standards.html>. Accessed July 7, 2008.
2. Campbell A, Dickson C. Predicting student success: a 10-year review using integrative review and meta-analysis. *J Prof Nurs*. 1996; 12(1): 47 – 59.
3. Platt L, Turocy P, McGlumphy B. Preadmission criteria as predictors of academic success in entry level athletic training and other allied health educational programs. *Journal of Athletic Training*. 2001; 36(2): 141 – 144.
4. Sandow P , Jones A , Peek C, Courts F, Watson R. Correlation of admission criteria with dental school performance and attrition. *J Dent Educ*. 2002; 66(3): 385 – 392.
5. Hardigan P, Lai L, Arneson D, Robeson A. Significance of academic merit, test scores, interviews, and the admission process: a case study. *American Journal of Pharmaceutical Education*. 2002; 65: 40 – 43.
6. Salvatori P. Reliability and validity of admissions tools used to select students for the health professions. *Advances in Health Sciences Education*. 2001; 6:159 – 175.
7. Sadler J. Effectiveness of student admission essays in identifying attrition. *Nurse Education Today*. 2003; 23(8): 620 - 627.
8. Ferguson E, James D, O'Hehir F, Sanders A. Learning in practice. *BMJ*. 2003; 326: 429 – 432.

Evaluation of Candidates' Written Communication Skills

9. Kulatunga-Moruzi C, Norman G. Validity of admissions measures in predicting performance outcomes: the contribution of cognitive and non-cognitive dimensions. *Teaching and Learning in Medicine*. 2002; 14(1): 34-42.
10. Dieter P, Carter R, Rabold J. Automating the complex school admission process to improve screening and tracking of applicants and decision making outcomes. *Perspective on Physician Assistant Education*. 2000; 11(1): 25 – 34.
11. Skaff K, Rapp D, Fahringer D. Predictive connections between admissions criteria and outcomes assessment. *Perspective on Physician Assistant Education*. 1998; 9(2): 75-78.
12. Hanson M, Dore K, Reiter H, Eva K. Medical school admissions: revisiting the veracity and independence of completion of an autobiographical screening tool. *Acad Med*. 2007; 82(10): S8 - S11.
13. Chestnut R, Phillips C. Current practices and anticipated changes in academic and nonacademic admission sources for entry-level PharmD programs. *American Journal of Pharmaceutical Education*. 2000; 64: 251-259.
14. Albanese M, Snow M, Skochelak S, Huggett K, Farrell, P. Assessing personal qualities in medical school admissions. *Acad Med*. 2003; 78(3): 313 – 321.
15. Bill Gates 1997 Annual report letter to shareholders.
http://www.microsoft.com/msft/reports/ar97/bill_letter/bill_letter.htm. Accessed July 7, 2008.
16. Flesch R. A new readability yardstick. *J Appl Psychol*. 1948; 32(3): 221 – 233.
17. Shermis M, Koch C, Page E, Keith T, Harrington S. Trait ratings for automated essay grading. *Educational and Psychological Measurement*. 2002; 62(5): 5 – 18.

Evaluation of Candidates' Written Communication Skills

18. Shermis M, Barrera F. Automated essay scoring for electronic portfolios. *Assessment Update*. 2002; 14(4): 1-4.
19. Rudner L, Garcia V. An evaluation of IntelliMetric[®] essay scoring system. *Journal of Technology Learning and Assessment*. 2006; 4(4): 1 – 21.
20. Shermis M, Burstein J, Leacock C. Applications of computers in assessment and analysis of writing, chapter 27. In: *Handbook of Writing Research* Guilford Press, c2006: 403-416.
21. Dikli S. An overview of automated scoring essays. *Journal of Technology, Learning, and Assessment*. 2006; 5(1): 4-35.
22. Vantage Learning. IntelliMetric[®] scoring accuracy across genres and grade levels. 2006. www.vantagelearning.com. Accessed July 7, 2008.
23. Korbin J. Forecasting the predictive validity of the new SAT I writing section. Available at the College Board Webpage www.collegeboard.com/prod_downloads/sat/newsat_pred_val.pdf Accessed June 15, 2008
24. Breland H, Bridgeman B, Fowles M. Writing assessment in admission to higher education: review and framework. College Entrance Examination Board and Educational Testing Service, 1999 New York