WRITING IN CYBERSPACE
A Pilot Study of WritePlacer Plus with Intellimetric Scoring

RAMAPO COLLEGE OF NEW JERSEY

Office of Testing & Placement Program
Ramapo College of New Jersey

Wanda G. Kosinski, Testing Coordinator

©2001, Office of Institutional Research and Planning, Ramapo College of New Jersey

This work was supported in whole by a grant from the Ramapo College Foundation.
## CONTENTS

I. Introduction .............................................................................................................. 1
II. Background ............................................................................................................ 1
III. About WritePlacer Plus .......................................................................................... 2
IV. How a Computer Scores an Essay .......................................................................... 2
V. Study Design .......................................................................................................... 3
VI. Results .................................................................................................................... 4
VII. Students' Reactions to WritePlacer Plus .............................................................. 7
VIII. Conclusions and Recommendations ..................................................................... 8
Introduction

This report summarizes the results of a pilot study of WritePlacer Plus with electronic scoring conducted at Ramapo College on March 31 and April 7, 2001. WritePlacer Plus is the direct writing assessment component of the College Board's ACCUPLACER online system that is currently used at Ramapo College for placement testing. Funding for the pilot study was made possible in whole by a grant from the Ramapo College Foundation.

The goal of the pilot study was to simplify and make more efficient the students' placement testing process by providing them with a complete online testing experience that includes direct writing assessment, administered through the Internet and scored immediately with accurate and reliable results. The purpose of this research was to compare the performance of computer scoring of essays and traditional scoring of essays by human expert readers to insure that students are receiving reliable and accurate, as well as timely test results.

Background

In March 2000, Ramapo College instituted the ACCUPLACER system of computerized placement tests. Rather than pencil and paper tests used in the past that were sent out to be scored, the computer tests enabled students to receive scores and course placements at the conclusion of the testing. However, for those students who were required to answer an essay question, the score was available one to two weeks later. This meant that students had to wait to register for writing courses until they received the additional score report in the mail.
In March 2001, the wait time for essay score and writing course placement was eliminated with the addition of WritePlacer Plus with electronic scoring. At the conclusion of the testing, Ramapo College students now leave with all test scores including the essay and have all the information needed for appropriate advisement and registration.

**About WritePlacer Plus**

WritePlacer Plus is a direct measure of student writing skills offered as part of the College Board's ACCUPLACER Program. Examinees are asked to provide a writing sample in response to a specific prompt. This assessment measures writing skill at the level expected of an entering level college student. The test is used to assess the writing skill of first-year college students in order to determine the need for developmental instruction before enrolling in college-level English courses.

**How a Computer Scores an Essay**

IntelliMetric is an artificial intelligence scoring system that was developed by Vantage Learning for College Board. WritePlacer Plus uses IntelliMetric to read and score writing samples. The IntelliMetric system models the process used by humans to read and score essays. The computer is trained to evaluate the overall effectiveness of an essay based on the following writing features:

- Focus
- Organization
- Development and Support
- Mechanical Conventions

The computer then scores each essay and returns the score, along with feedback, immediately. Scores are reported on a scale of 2 to 12 and reflect the sum of two
traditional readers' scores. "A score of zero indicates that the response was off topic, in a
language other than English, too short to score, or in some other way not scorable."¹

In other words, the computer is trained to score essays in a way that is conceptually
equivalent to the training received by human readers who conduct holistic scoring. This
equivalent training results in the computer's achievement of impressive strong agreement
with the scores assigned by human expert readers. In addition to the numerical essay score,
WritePlacer Plus also provides students with additional feedback on their writing strengths
and developmental needs (see Individual Score Report in Appendix).

Study Design

On March 31 and April 7, 2001, ACCUPLACER was administered to 154 students.
Of these, 105 students were required to take the English portion of the test that includes
reading comprehension and sentence skills. Sixty-five of these students achieved a score of
less than 98 on the sentence skills section and were required to take the essay test,
WritePlacer Plus.

A total of 65 student essays were included in the WritePlacer Plus pilot study - 30
eyoars from the March 31 test and 35 essays from the April 7 test. On the day of the test,
written responses were electronically scored and students received their score report and
writing course placement. At the conclusion of the test session, a copy of each student's
essay was printed. In order to eliminate scoring bias, the student's name and identification
number were concealed and each essay was then assigned a code number. The essays
along with a score roster were submitted to faculty in Basic Studies at Ramapo College for

scoring. After the on-campus scoring, the essays were sent to the College Board for a second expert human scoring.

Results

Rate of Agreement between Scorers

Table A in the Appendix shows the frequency of score differences between the computer scoring and the two scorings by human experts. The scores were examined in terms of the rate of agreement between Computer and the Human Experts. The computer agreed with the College Board (CB) graders within one point in 92% (60) of the cases. The computer agreed with the Basic Studies (BS) graders within one point in 43% (28) of the cases. Human expert readers agreed with each other within one point in 54% (35) of the cases. These results are presented in the table below.

<table>
<thead>
<tr>
<th>SCORERS</th>
<th>AGREEMENT</th>
<th>NO.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computer v. College Board</td>
<td>92%</td>
<td>60</td>
</tr>
<tr>
<td>Computer v. Basic Studies</td>
<td>43%</td>
<td>28</td>
</tr>
<tr>
<td>College Board v. Basic Studies</td>
<td>54%</td>
<td>35</td>
</tr>
</tbody>
</table>

Score Outliers

Table B in the Appendix illustrates the frequency of essay score outliers - that is to say, where the scorers differed by 2 (or more) points. For this sample of essays, the computer and College Board graders differed in 8% (5) of the cases. The computer and Basic Studies graders differed in 57% (37) of the cases. The human experts differed with
each other in 46% (30) of the cases. These results are presented in the table below.

<table>
<thead>
<tr>
<th>SCORERS</th>
<th>DISAGREEMENT</th>
<th>NO.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computer v. College Board</td>
<td>8%</td>
<td>5</td>
</tr>
<tr>
<td>Computer v. Basic Studies</td>
<td>57%</td>
<td>37</td>
</tr>
<tr>
<td>College Board v. Basic Studies</td>
<td>46%</td>
<td>30</td>
</tr>
</tbody>
</table>

Correlations

Figure 1 graphically depicts the relationship between the computer scores and College Board human expert scores for the sample essays. A visual scan of the data shows a clear, positive relationship between the two sets of scores that is supported by statistical tests. The correlation coefficient (Pearson's $r$) between computer score and College Board human score is .92 and is significant at the 0.01 level (2-tailed) (see Table C in Appendix).²

² Correlation is significant at the 0.01 level (2-tailed) means that one can be 99% confident that the results are due to a real relationship, not chance factors (or sampling error).
The correlations between essay scores for Basic Studies and the computer and between the two human expert scorers are also positive and supported by statistical tests, but not as strong. The correlation coefficient (Pearson's r) between Basic Studies and the computer is .64 and between College Board and Basic Studies is .60. Both are statistically significant at the .01 level (2-tailed) (figures 2 and 3 below).

Figure 2: Computer v. Basic Studies

Figure 3: College Board v. Basic Studies
SAT scores

The essay scores for all three readings are strongly correlated with SAT-Verbal scores (see Table C in the Appendix). These results are presented in the table below.

<table>
<thead>
<tr>
<th>SAT-Verbal Score</th>
<th>Pearson’s $r$</th>
<th>Significance Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>with Computer</td>
<td>.51</td>
<td>.01</td>
</tr>
<tr>
<td>with College Board</td>
<td>.42</td>
<td>.01</td>
</tr>
<tr>
<td>with Basic Studies</td>
<td>.47</td>
<td>.01</td>
</tr>
</tbody>
</table>

Students' Reaction to WritePlacer Plus

To date, 172 students have been tested with WritePlacer Plus and had their writing samples electronically scored. At the conclusion of testing students are asked to answer a short reaction survey and to add any comments that they would like to make about their testing experience. Students were asked if they had enough time to plan and write their essay and the majority (85%) 'strongly agreed' or 'agreed' that they had enough time for the written portion of the test. When asked if they would rather have taken the test in a paper and pencil version, 86% indicated their preference for taking the test on computer.

Here's a sample of some of the additional comments that students added to the survey:

My testing experience was good. I like the fact that the test was on the computer. This was very helpful when it comes to writing the essay. Because I can type much faster then I write. (Student survey 449)
I liked the ACCUPLACER test much better then pencil and paper. I felt it was less stressful. (Student survey 306)

I thought it was well organized and everyone was very polite and friendly. I really, really liked that it was all done by computer. (Student survey 303)

The test was very comfortable to take and it was not at all time consuming. I wish in high school we could have taken stand[a]rdized tests this way. (Student survey 317)

It was very quiet and comfortable. The proctors are very sweet and I love this college. Keep up the terrific work ☺. (Student survey 336)

**Conclusions and Recommendations**

Accurate and efficient writing skills assessment is important. The benefits of assessing students' writing skills with WritePlacer Plus with computer scoring are obvious:

- Assurance of reliable and accurate scoring
- Diagnostic score report and writing course placement returned within seconds of test completion
- Scoring that is free of bias and prejudice
- Choice of field tested prompts
- Easy access to view or print copies of essay responses
- Complete online testing experience that continues the college wide efforts to support and fully utilize available technology

The results of this study and the positive feedback received from our students suggest that using WritePlacer Plus with IntelliMetric scoring enhances the students' testing experience. It provides a better way to service the students enrolled at the College.

The comprehensive study of testing and placement conducted recently at Ramapo College included the recommendation to provide students with more information about
their strengths and weaknesses in the skills areas covered by the placement test. The
diagnostic reporting included in the dimensional score component of WritePlacer Plus is a
positive step towards meeting this goal.

WritePlacer Plus Electronic Scoring provides students with a paragraph that
includes a general explanation of the essay score they received. In addition to the overall
score and explanation, each of the individual areas on which the essay is assessed is
discussed and assigned its own score (see samples of Individual Score Reports in
Appendix). The diagnostic score report includes information on the following writing
skills considered important for college-level students:

- Focus - the clarity with which the main idea or point of view is stated and
  maintained
- Development - The amount, depth, and specificity of the supporting details
- Organization - The clarity of writing and the logical sequence of ideas
- Sentence Structure - The effectiveness of the sentence structure and the extent to
  which the writing is free of errors in sentence structure.
- Conventions - The ability to spell common words and the use of the conventions of
capitalization and punctuation. The extent to which the writing is free of errors in
usage and shows care and precision in word choice.

**Recommendation #1: Ramapo College should continue to use WritePlacer Plus
with Intellimetric Scoring.**

**Recommendation #2: Faculty teaching developmental writing courses should
receive a blind copy of the diagnostic score report that could be used to tailor
courses to incoming students' needs.**

---

Planning, Ramapo College of New Jersey.
We believe that the continuation of WritePlacer Plus will allow us to provide our students with state of the art placement testing, improve the validity and reliability of our placements, and provide students and faculty with helpful information to enhance the process of teaching and learning.

END