The Long and the Short of It: Correlations between Word Count and Paper Score

by John Fallon
Quick Exercise: You have been given a handout with a sample paper and directions on how to score the paper. Please take 3 or 4 minutes to score the paper.

BACKGROUND INFORMATION
The writer of the paper is a high school junior who had 25 minutes to read the prompt; pre-write as desired; and compose the essay (same constraints as SAT).

FIVE TRAITS TO CONSIDER AS YOU READ THE PAPER:

Critical Thinking—How thoughtful is the paper?

Organization—To what extent is the paper structured?

Development—What is the quality of details in the paper?

Style—How clear and sophisticated is the language used in the paper?

Mechanics—To what extent is there correct punctuation and grammar?
1. “Cracking the code” of the SAT

2. Dr. Les Perelman

“I have never found a quantifiable predictor in 25 years of grading that was anywhere near as strong as this one is.”
Were the code cracker and Perelman correct?

Just write more to score higher?
• We needed someone to write the essay
  42 HS juniors from Ohio Hi Point HS

• We needed someone to grade the essay
  8 college English faculty
  12 HS language arts teachers

Note: All the evaluators were English teachers.
All 20 evaluators scored all 42 papers. (That’s 840 paper readings!)

Did longer papers receive higher scores?

Yes.
## Results: Scores by Word Count

<table>
<thead>
<tr>
<th>Word Count</th>
<th>Average Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>51-159 Words</td>
<td>15</td>
</tr>
<tr>
<td>160-267 Words</td>
<td>21</td>
</tr>
<tr>
<td>268-375 Words</td>
<td>24</td>
</tr>
<tr>
<td>376-477 Words</td>
<td>26</td>
</tr>
</tbody>
</table>
The more the points cluster tightly about the line, the higher the magnitude of the correlation between word count and score.
** We have a tight cluster!
<table>
<thead>
<tr>
<th>Paper #1 (long)</th>
<th>Paper #2 (short)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**NEWSFLASH:** Paper #1 and paper #2 are the same paper!
Interpretation of the difference in scores on the exercise paper:

If a paper even “looks” longer, it will receive a higher grade.
Another interesting factoid about the paper you just scored (paper #9)

Four teams without rubric training scored this paper: 27.85 (out of a possible 40)

Six teams with rubric training scored this paper: 22.75 (out of a possible 40)

Teams without rubric training graded higher, on average, for ALL THE PAPERS (not just paper #9)
### Mean score by team

<table>
<thead>
<tr>
<th>Mean paper score</th>
<th>Team</th>
</tr>
</thead>
<tbody>
<tr>
<td>26</td>
<td>HS Untrained</td>
</tr>
<tr>
<td>24</td>
<td>HS Untrained</td>
</tr>
<tr>
<td>19</td>
<td>C Untrained</td>
</tr>
<tr>
<td>19</td>
<td>C Untrained</td>
</tr>
<tr>
<td>18</td>
<td>HS Trained</td>
</tr>
<tr>
<td>17</td>
<td>C Trained</td>
</tr>
<tr>
<td>16</td>
<td>C Trained</td>
</tr>
<tr>
<td>16</td>
<td>HS Trained</td>
</tr>
<tr>
<td>15</td>
<td>HS Trained</td>
</tr>
<tr>
<td>15</td>
<td>C Trained</td>
</tr>
</tbody>
</table>

**Average paper score:**

- **Untrained Teams:** 22
- **Trained Teams:** 16
This chart looks at the strength of the relationship between word count and paper score by evaluation group.

Levels of Magnitude

.5 and below is weak

.8 and above is strong

We have between .59 and .71 (moderate magnitude)
Interpreting Pearson’s

CORRELATION COEFFICIENT MAGNITUDE: 0.5 LOW AND 0.8 HIGH

COEFFICIENT OF DETERMINATION (\(r^2\)) TO WHAT EXTENT DOES WORD COUNT EXPLAIN PAPER SCORE?

- HT: 0.599, 36%
- CT: 0.615, 38%
- HU: 0.714, 51%
- CU: 0.639, 41%
Perelman and the SAT code cracker were right. Paper grade is explained by word count, but also by rubric training:

- 36% HS teachers trained
- 51% HS teachers untrained

- 38% College teachers trained
- 41% for College teachers untrained
The Long and the Short of It

1. As a WAC leader at your school, consider training teachers to use a rubric rather than just providing them with a rubric.

2. Remind your students to develop their essay answers—word count matters to their paper scores.

Questions / Comments?
Thank you