Study of the Impact of Feedback on Math Instructor Performance

Peg Hohensee, Ph.D.
Leslie Johnson, M.S.
Kaplan University (Purdue NewU)
What is Your Current Instructor Feedback Plan?

- Annual Evaluation
- Review/Observation each term
- Online courses
Prior to mid 2012

- Typically Limited
- Primarily Corrective
- Yearly Formal Evaluation

Mid 2012 to 2014

- Increased Frequency
- Positive and Opportunistic
- Yearly Formal Evaluation

Kaplan Math Monitoring Background
New Plan for Monitoring Faculty

- Purposes:
  - Provide Feedback to Instructors
  - Verify Quality Instruction
  - Content for Yearly Adjunct Instructor Evaluation
- Frequency - Typically Once/Term
- Scoring for Leadership and Research Purposes Only – Instructors not provided “scores” (may not be aware of scores)
2014 Research Questions

1. Can providing a more holistic review paired with a frequent comprehensive feedback program improve instructor performance in online classes?

2. Will the program translate to improved student satisfaction with the instructor?
2014 Study Sample:

• KU Math Dept: 200+ full time and adjunct faculty

• 117 adjunct faculty subjects: Taught 2 or more pre- and post-program terms
Instructor Comment

Thanks so much for the email! I'm always happy to get an outside perspective on what I am doing well and what needs improvement. I feel terrible about the Unit 9 grading this term - I do take the grading deadlines very seriously but I've been holed up in bed with sciatica for the past few weeks and haven't been anywhere near 100%. All my final grades are posted now.

Thanks again for your feedback!
Research Questions

• Can providing a more holistic review paired with a frequent comprehensive feedback program improve instructor performance in online classes?
## 2014 Results

### Matched Pairs t-Test for Instrument Scores

<table>
<thead>
<tr>
<th></th>
<th>Post-Program</th>
<th>Pre-Program</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>19.55</td>
<td>15.41</td>
</tr>
<tr>
<td>Variance</td>
<td>4.23</td>
<td>4.90</td>
</tr>
<tr>
<td>Observations</td>
<td>117</td>
<td>117</td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>0.50</td>
<td></td>
</tr>
<tr>
<td>df</td>
<td>116</td>
<td></td>
</tr>
<tr>
<td>t Stat</td>
<td>20.91</td>
<td></td>
</tr>
<tr>
<td>( P(T \leq t) ) one-tail</td>
<td>0.0000</td>
<td></td>
</tr>
</tbody>
</table>

Cohen’s d (effect size): 3.88241 – Very Large
Research Questions

• Will the program translate to improved student satisfaction with the instructor?
### 2014 Results

Matched Pairs t-Test for End of Term Survey Scores

<table>
<thead>
<tr>
<th></th>
<th>Post-Program</th>
<th>Pre-Program</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean</strong></td>
<td>4.61</td>
<td>4.55</td>
</tr>
<tr>
<td><strong>Variance</strong></td>
<td>0.04</td>
<td>0.05</td>
</tr>
<tr>
<td><strong>Observations</strong></td>
<td>117</td>
<td>117</td>
</tr>
<tr>
<td><strong>Pearson Correlation</strong></td>
<td>0.27</td>
<td></td>
</tr>
<tr>
<td><strong>df</strong></td>
<td>116</td>
<td></td>
</tr>
<tr>
<td><strong>t Stat</strong></td>
<td>2.57</td>
<td></td>
</tr>
<tr>
<td><strong>P(T&lt;=t) one-tail</strong></td>
<td>0.0058</td>
<td></td>
</tr>
</tbody>
</table>

Cohen’s d (effect size): 0.47676 - Medium
2014 Study


• Available via the website
Longitudinal Study
Prior to mid 2012
• Typically Limited
• Primarily Corrective

Mid 2012 to 2014
• Increased Frequency
• Positive and Opportunistic

2014 to 2016
• Same Frequency
• Improved Instrument

Kaplan Math Monitoring Background
"The best data science, in fact, is surprisingly intuitive."

"While the methodology of good data science is often intuitive, the results are frequently counterintuitive."

*Everybody Lies: Big Data, New Data, and What the Internet Can Tell Us About Who We Really Are*

•-- Seth Stephens-Davidowitz
2017 Research Questions

1. Can providing a more holistic review paired with a frequent comprehensive feedback program improve instructor performance in online classes?

2. Will the program translate to improved student satisfaction with the instructor?

3. Is student performance correlated to instructor score?
2017 Study Instrument

- Google Forms
- Boolean question format
- 22 scoreable items
- 2 negatively scored items
- Scoring range: (-2, 20)
- Based on University Best Practices
- Same reliability and validity arguments as 2014
- Available via the website
2017 Study Sample:

- KU Math Dept: ~100 full time and adjunct faculty
- 1000+ Available data points
Number of Monitored Classes Each Year

![Bar chart showing the number of monitored classes each year.
- 2014: 406 classes
- 2015: 383 classes
- 2016: 270 classes]
Research Questions

• Can providing a more holistic review paired with a frequent comprehensive feedback program improve instructor performance in online classes?
2012: 15.41 for initial study baseline

<table>
<thead>
<tr>
<th>Year</th>
<th>Average Monitoring Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014</td>
<td>17.72</td>
</tr>
<tr>
<td>2015</td>
<td>18.82</td>
</tr>
<tr>
<td>2016</td>
<td>18.81</td>
</tr>
</tbody>
</table>
Research Questions

• Will the program translate to improved student satisfaction with the instructor?
2012:
4.55 for initial study baseline

Average EOT

- 2014: 4.63
- 2015: 4.64
- 2016: 4.63

Not Necessarily!
Research Questions

• Is student performance correlated to instructor score?
Likely!
### 2014-2016 Monitoring and Urate

#### t-Test: Paired Two Sample for Means

<table>
<thead>
<tr>
<th></th>
<th>Average Monitoring Score</th>
<th>Average Urate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>18.45</td>
<td>0.212466667</td>
</tr>
<tr>
<td>Variance</td>
<td>0.3997</td>
<td>5.59633E-05</td>
</tr>
<tr>
<td>Observations</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>-0.99914589</td>
<td></td>
</tr>
<tr>
<td>Hypothesized Mean Diff</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>df</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>t Stat</td>
<td>49.3804699</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) one-tail</td>
<td>0.000204924</td>
<td></td>
</tr>
<tr>
<td>t Critical one-tail</td>
<td>2.91998558</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) two-tail</td>
<td>0.000409848</td>
<td></td>
</tr>
<tr>
<td>t Critical two-tail</td>
<td>4.30265273</td>
<td></td>
</tr>
</tbody>
</table>

**EXTREMELY LIKELY!**
Limitations and Considerations

- Using averages for each year
- Unaccounted for variables
Where to go from here?

Would you have documents to use for basis for performance standards?
Contact

Peg Hohensee, Ph.D.
phohensee@kaplan.edu

Leslie Johnson, M.S.
ljohnson6@kaplan.edu