Self- vs. Peer- Evaluation
Which one is more accurate?

Jon Guest – Aston University
Robert Riegler – Coventry University
Background

• Importance of developing the independent/self directed learning skills of students in HE
• To become effective independent learners students need to acquire good evaluation skills
• To what extent can they judge the quality of a piece of work?
We have previously focussed on self-evaluation - Guest and Riegler, 2017

- Students display high level of over-confidence

- Do students have a better understanding of the standards than their self evaluation estimates suggest?

- Do they find it particularly difficult to objectively evaluate the quality of their own work?

- Might they find it easier to accurately judge the quality of work produced by their peers?
Literature

• Falchikov and Goldfinch (2000)
  • Meta analysis

• Ashenafi (2017)
  • "corroborates the findings of Falchikov and Goldfinch"

• Limited number of papers compare self, peer and tutor scores/grades
  • Sunol, Arbat, Pujol, Feliu, Fraguell and Llado (2016)
Some research design issues

• Inter-marker reliability
• Gatekeeper activity
• Impression management bias
• Difficult task
  • Mark incentive
Research Design

• Essay title plus assessment criteria were released 1 month before deadline.

• Essay deadline:
  • Self-evaluation form submitted.

• 1 day after essay deadline:
  • Essays allocated randomly to students using “The Workshop” tool in Moodle.

• 1 week after essay deadline:
  • Peer-evaluation form submitted
Data

- 131 students, 110 students agreed to participate in the study.
- 77 Male, 33 Female students
- 76 students have UK Education background (A levels or BTEC)
- 8 students repeated the module
- 11 students are ERASMUS/Direct Entry
Results (First Look)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
<th>t test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tutor Mark</td>
<td>110</td>
<td>56.51</td>
<td>9.83</td>
<td>3</td>
<td>78</td>
<td></td>
</tr>
<tr>
<td>Self- Evaluation</td>
<td>110</td>
<td>59.64</td>
<td>8.23</td>
<td>0</td>
<td>72</td>
<td>p = 0.01</td>
</tr>
<tr>
<td>Tutor Mark</td>
<td>110</td>
<td>55.54</td>
<td>8.97</td>
<td>25</td>
<td>80</td>
<td></td>
</tr>
<tr>
<td>Peer- Evaluation</td>
<td>110</td>
<td>56.35</td>
<td>12.31</td>
<td>0</td>
<td>87</td>
<td>p = 0.57</td>
</tr>
</tbody>
</table>
SE, PE and Tutor Grade differences

Student Self Evaluation Grade

Student Peer Evaluation Grade

Tutor Grade
The charts show a comparison between student self-evaluation and peer-evaluation against tutor grades. Each chart represents a different type of evaluation:

- **Self-Evaluation** graph on the left:
  - Y-axis: Student Self-Evaluation
  - X-axis: Tutor Grade
  - Legend: Student
  - Dotted line: Tutor Grade = Self-Evaluation

- **Peer-Evaluation** graph on the right:
  - Y-axis: Student Peer-Evaluation
  - X-axis: Tutor Grade
  - Legend: Student
  - Dotted line: Tutor Grade = Self-Evaluation

Note: The graphs illustrate the relationship between student and tutor evaluations across different grades. The dotted line indicates the self-evaluation against tutor grades, showing alignment in certain segments.
Chi$^2$ Test: Pr = 0.496
Conclusion

- Students are on average more accurate in evaluating the work of their peers than their own work.
- Data on Peer-evaluation is more dispersed than on self-evaluation.
- Students are very heterogenous with respect to SE and PE precision.
- Trade off between emotional attachment and cognitive workload.
- Surprise about quantity of feedback provided.
What's next?

• 94 student provided feedback on peer-essay (voluntarily)

• Analyse impact of student ability on evaluation accuracy.

• Are results sensitive with respect to timing?