Anonymous Marking @ Kent

Dr. Edd Pitt

Centre for the Study of Higher Education
Why Anonymous Marking??

- 2008 NUS ‘mark my words, not my name’ campaign.
  - “reduces both the fear and likelihood of discrimination”
- 1999 survey, “44% of students’ unions believed that discrimination and bias played a part in the way that students’ work was assessed and addressed.”
- NUS - AM
  - Protects staff from potential accusations of prejudice; without activating prior knowledge of a student’s past performance, a marker cannot be biased by prejudgements (Fleming, 1999).
Bias in the Grading of Work

- Assessment processes should be equitable (Brennan, 2008).
- Potential exists for discrimination on grounds of race, age personal characteristics (Dennis & Newstead, 1994)
- High school - Science markers - higher marks - perceived a blind piece of work written by male. (Goddard-Spear, 1984)
- HE: Less experienced 2nd markers graded work that was not anonymised (no knowledge of mark)
  - Higher grades than the first marker to males
  - Females were similar for both the first and second markers. Bradley (1984)
Bias in the Grading of Work?

- Not replicated
  - Bias attributable - marker’s prior knowledge of student, & not gender (Dennis and Newstead, 1994).

- No consistent evidence of gender bias
  - Secondary school (Baird, 1988)

- Exams marked anonymously & non-anonymously
  - No evidence of gender bias in marking.
    —(Owen, Stefaniak and Corrigan, 2010)
Bias in the Grading of Work?

- AM University of Wales
- Women 1\textsuperscript{st} Degrees up by 13%
- Prior knowledge creates bias in mind (Brennan, 2008 p. 43)
- Bias minimised if marker has no knowledge of the students’ previous performance, grades, or gender (Malouff, Emmerton & Schutte, 2013)
- Little conclusive evidence that anonymous marking eradicates gender bias (Krawczyk, 2017).
Hinton and Higson (2017) analysed 32,000 student records from a 12-year period.

Mean performance on AM exams & coursework & non AM oral presentations

Little evidence this reduces attainment gaps

AM Exams - attainment gap reduced by
- 0.6% (Gender)
- 1.5% (BME)

Oral presentations, attainment gap narrowed,
- Not attributable to anonymous assessment.

AM not reduced between-group performance differences.
Change to AM – Potential Effect

- 2nd marker no prior knowledge – AM not needed (Dennis, 2007)

- AM negative commentary on accepted assessment accuracy within the sector.
  - Cynicism - previous marking practices
  - Distrust between lecturers & students (Owen, et al, 2010)
Arguments Against of Anonymous Marking

• Critics
  • Erode trust in the assessment process
  • Depersonalises teaching
  • Undermines the development aspect of feedback
  – (see, Baty, 2007; Southee, 2009; Price et al, 2010; Beals, 2012; MacDonald Ross, 2012).
The Effect of Anonymous Marking on Feedback

- Feedback on an anonymous basis is potentially problematic. (Whitelegg, 2002)
- AM disrupts the feedback loop - individualisation of feedback comments removed
  - Distance between staff and students increased.
- Monologic feedback paradigm - passive recipients of feedback rather than active engagers with feedback dialogue (Winstone & Pitt, 2017)
- Brennan (2008) professional work context employees are rarely appraised anonymously.
The Effect of Anonymous Marking on Feedback

- Depersonalised feedback recognised (Whitelegg, 2002; Crook, Gross, and Dymott, 2006).

- Feedback effect upon student emotions (Rowe, Fitness & Wood, et al., 2013; Pitt & Norton, 2016; Pitt, 2017)

- Objective, depersonalised feedback fosters feelings of not belonging and detachment from the learning situation (Pitt, 2017)
The Present Study

1. Explore whether students perceive anonymous marking as fairer than non-anonymous marking, and whether perceptions of fairness differ according to student gender and ethnicity.
   - Test for statistical differences between student marks for anonymously and non-anonymously marked work.

2. Whether students perceive non-anonymous marking as fostering a stronger relationship with their lecturer than anonymous marking

3. Whether feedback on non-anonymously marked work is perceived as more helpful to learning than feedback from anonymously marked work
Design

• Across four first year UK undergraduate subjects (Business (n=87); Politics (n=60); Pharmacy (n=40); French (n=8)).

• 442 students –
  • 1 Module AM Coursework
  • 3 Non AM Coursework

• 195 responded
  • Male n = 98, $M_{age} = 19.48$, $SD = 2.69$
  • Female n = 97, $M_{age} = 19.16$, $SD = 1.14$

• White (m=54, f=48),

• Black (m=27, f=35)

• Asian (m=17, f=14).
1. **Questionnaire**

- **Anonymous/Non Anonymous marked work**
  - Bias, fairness, transparency & confidence in the process

- **Feedback & Learning**
  - Content, future use, clarifying gaps in knowledge, Use in next assessment, seeking clarification, increasing motivation

- **Relationships**
  - Learning progress recognition, relationship with lecturer, effort recognition
### Student Performance (M, SD) AM & AM Coursework, by Gender & Ethnicity.

<table>
<thead>
<tr>
<th></th>
<th>Anonymously marked coursework</th>
<th>Non-anonymously marked coursework</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>M (SD)</td>
</tr>
<tr>
<td><strong>Male</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>24</td>
<td>62.42 (11.42)</td>
</tr>
<tr>
<td>Black</td>
<td>40</td>
<td>58.88 (10.24)</td>
</tr>
<tr>
<td>White</td>
<td>120</td>
<td>61.40 (11.15)</td>
</tr>
<tr>
<td><strong>Female</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>26</td>
<td>59.88 (13.09)</td>
</tr>
<tr>
<td>Black</td>
<td>44</td>
<td>59.95 (11.47)</td>
</tr>
<tr>
<td>White</td>
<td>77</td>
<td>64.66 (9.21)</td>
</tr>
</tbody>
</table>
Results – Performance Data

- 2 (Anonymity; anonymous and non-anonymous) x 2 (Gender) x 3 (Ethnicity) Mixed ANOVA with Anonymity as the repeated measure.

- Students’ performance did not differ significantly across coursework that was marked anonymously and non-anonymously ($F(1, 319) = .40, p = .49, \eta^2_p = .002$).

- None of the interactions including the factor of Anonymity statistically significant (all $p$s > .05)

- AM did not advantage or disadvantage particular groups of students.

- Main effect of ethnicity ($F(2, 319) = 5.13, p = .006, \eta^2_p = .03$);
  - Regardless of whether work was marked anonymously or non-anonymously,
  - White ($M = 62.81, SD = 8.43$) higher than
  - Black ($M = 59.81, SD = 7.70$) ($p = .006, d = 0.36$).
## Mean (SD) Survey Data AM & Non AM

<table>
<thead>
<tr>
<th>Format</th>
<th>Domain</th>
<th>α</th>
<th>Ethnicity &amp; Gender</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>White</td>
<td>Black</td>
<td>Asian</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Male N = 54</td>
<td>Male N = 27</td>
<td>Male N = 17</td>
<td>Female N = 48</td>
<td>Female N = 35</td>
<td>Female N = 14</td>
</tr>
<tr>
<td>AM</td>
<td>Fairness</td>
<td>.86</td>
<td>3.24 (.68)</td>
<td>3.87 (.66)</td>
<td>3.13 (.86)</td>
<td>3.55 (.90)</td>
<td>3.74 (.90)</td>
<td>3.50 (.89)</td>
</tr>
<tr>
<td></td>
<td>Feedback &amp;</td>
<td>.77</td>
<td>3.16 (.58)</td>
<td>3.46 (.64)</td>
<td>3.06 (.70)</td>
<td>3.22 (.54)</td>
<td>3.36 (.61)</td>
<td>3.29 (.96)</td>
</tr>
<tr>
<td></td>
<td>Learning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Relationships</td>
<td>.72</td>
<td>3.02 (.74)</td>
<td>3.29 (.59)</td>
<td>2.68 (.68)</td>
<td>3.08 (.77)</td>
<td>3.14 (.66)</td>
<td>3.00 (.90)</td>
</tr>
<tr>
<td>Non AM</td>
<td>Fairness</td>
<td>.85</td>
<td>3.46 (.65)</td>
<td>3.03 (.72)</td>
<td>3.50 (.88)</td>
<td>3.22 (.91)</td>
<td>3.25 (.94)</td>
<td>3.36 (.79)</td>
</tr>
<tr>
<td></td>
<td>Feedback &amp;</td>
<td>.73</td>
<td>3.53 (.50)</td>
<td>3.58 (.55)</td>
<td>3.56 (.71)</td>
<td>3.45 (.49)</td>
<td>3.36 (.51)</td>
<td>3.63 (.63)</td>
</tr>
<tr>
<td></td>
<td>Learning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Relationships</td>
<td>.67</td>
<td>3.33 (.78)</td>
<td>3.52 (.71)</td>
<td>3.51 (.76)</td>
<td>3.33 (.72)</td>
<td>3.51 (.61)</td>
<td>3.74 (.53)</td>
</tr>
</tbody>
</table>
**Perceptions of Fairness in Marking**

- 2 (Anonymity: Anonymous; Not Anonymous) x 2 (Gender) x 3 (Ethnicity) mixed ANOVA anonymity as the repeated measure, and perceived fairness in marking as the DV.
- No significant difference in perceptions of fairness ($F(1, 189) = 3.09$, $p = .08, \eta^2_p = .02$)
- No significant gender or ethnicity differences - perceived fairness of marking ($F(1, 189) = .57, p = .45, \eta^2_p = .003$, and $F(2, 189) = 1.46, p = .24, \eta^2_p = .02$).
- Significant interaction- AM & Gender, $F(1, 189) = 4.02, p = .046, \eta^2_p = .02$.
- Males – No difference in the perceived fairness of AM & NonAM ($t(97) = -.95, p = .35, d = 0.09$),
- Females perceived AM significantly fairer than NonAM, $t(96) = 3.63, p < .001, d = 0.37$.
- Neither the Anonymity x Ethnicity interaction, nor the three-way interaction, were significant ($p < .05$).
Perceptions of Feedback

• 2 (Anonymity: Anonymous; Not Anonymous) x 2 (Gender) x 3 (Ethnicity) mixed ANOVA, anonymity as the repeated measure, and perceptions of feedback as the DV.

• Feedback on NonAM perceived by students to have greater potential for learning than feedback on AM \( F(1, 189) = 15.15, p < .001, \eta_p^2 = .07 \).

• No significant effects of gender \( (F(1, 189) = 1.95, p = .17, \eta_p^2 = .01) \) nor ethnicity \( (F(2, 189) = 1.35, p = .26, \eta_p^2 = .01) \) on the perceived learning value of feedback.

• No significant two - or three-way interactions (all \( ps > .05 \)).
Perceptions of Relationships

- 2 (Anonymity: Anonymous; Not Anonymous) x 2 (Gender) x 3 (Ethnicity) mixed ANOVA, anonymity as the repeated measure, and perceptions of relationships as the DV.
- Significant main effect of anonymity, $F(1, 188) = 26.32, p < .001 \quad \eta^2_p = .12$.
- Students perceived a stronger relationship with the marker where work had not been marked anonymously, than when it had been marked anonymously.
- Main effect of gender ($F(1, 188) = 2.95, p = .09, \eta^2_p = .02$), nor the main effect of ethnicity ($F(2, 188) = 2.73, p = .07, \eta^2_p = .03$) significant.
- There were no significant two- or three-way interactions (all $ps > .05$).
Conclusions – Performance Data

- No statistical difference between students’ marks for AM and NonAM.
  - Pre-existing knowledge of the student’s identity does little to influence grading of work (Krawczyk, 2017).

- Student ethnicity and student performance data
  - Support Hinton and Higson’s (2017) anonymising student work does little to address already well-documented attainment gaps.

- Students not advantaged or disadvantaged by NonAM
Conclusions - Gender Bias

- No significant difference in perceived fairness of marking
- Female students perceived AM significantly fairer than NonAM
  - Performance - fairness of marking concerns not founded.
  - Do not provide support for biased marking when the identity of the student is known (Worsley & Knight, 1998; Brennan, 2008; Malouff, Emmerton & Schutte, 2013)
  - Align with those of Owen et al. (2010) in failing to demonstrate significant gender bias in marking.
- Open, transparent communication with students about marking practices, moderation, second marking and the role of the external examiner.
Conclusions - *Perceptions of Anonymous Feedback*

- Feedback NonAM perceived more helpful to learning
- Personalised feedback prefered (Handley et al., 2008; Tuck, 2012; Birch et al., 2016).
- Feedback on AM not reference students’ previous performance, their use of feedback from previous assessments or be emotionally sensitive
- NonAM fosters a stronger relationship with their lecturer than is possible with anonymous marking.
  - Enhance students’ internal well-being & positive mind-set (Price et al, 2010)
- Birch et al. (2016), AM may undermine feedback to foster positive relationships between lecturers & students.
Decisions Decisions!!

- November 2017 Education Board
  - Anonymous marking project discussed.

- No evidence that requiring further and broader introduction of anonymisation of marking would be advantageous

- Schools requested communicate benefits of NonAM coursework

- Address worries some students have.
WHAT HAPPENS TO MY WORK AFTER I SUBMIT?

YOUR WORK IS MARKED BY A MEMBER OF THE MODULE DELIVERY TEAM

MARKS ARE CHECKED FOR ACCURACY & CONSISTENCY. WORK SAMPLES ARE MODERATED BY AN EXPERIENCED MARKER

MARKING IS REVIEWED AND MAY BE ADJUSTED

IS MARKING CONFIRMED AS ACCURATE AND CONSISTENT?

YES

PROVISIONAL MARKS & FEEDBACK ARE RELEASED

EXTERNAL EXAMINER VERIFIES MARKS & ACADEMIC STANDARDS BY REVIEWING SAMPLES OF WORK

FINAL MARKS ARE CONFIRMED AT THE EXAM BOARD!

NO

IF THERE ARE STILL CONCERNS, WORK IS REFERRED TO THE CHAIR OF THE BOARD OF EXAMINERS

IF MARKERS STILL DISAGREE, MARKING IS REVIEWED BY THE EXTERNAL EXAMINER

ASK YOUR SCHOOL FOR A DETAILED TIMELINE

FOR MORE INFO, SEE HTTPS://WWW.KENT.AC.UK/TEACHING/DOCUMENTS/QUALITY-ASSURANCE/CREDITS-FRAMEWORK/CREDITS-INFORMATION-XX.X.PDF
References


