Institution: University of Essex

Unit of Assessment: 16

Title of case study: Reducing Gender Bias in Teaching Evaluations

Period when the underpinning research was undertaken: 2012 - 2018

Details of staff conducting the underpinning research from the submitting unit:

Name(s):	Role(s) (e.g. job title):	Period(s) employed by submitting HEI:
Friederike Mengel	Professor of Economics	2012- present

Period when the claimed impact occurred: 2014 – 2020

Is this case study continued from a case study submitted in 2014? N

1. Summary of the impact

Professor Mengel's research on gender bias in teaching evaluations has led to a change in teaching evaluation practices in Higher Education across Europe, Canada and the United States. The impact has had a wide reach. We can evidence significant impact in 14 institutions across these regions. It has led to changes in teaching evaluation practices, in teacher behaviour, in policy and to wider public recognition and debate around gender bias in teaching evaluations. We estimate that these changes have directly affected approximately 25,000 teachers and 425,000 students as well as many more indirectly.

The research has found evidence of bias in teaching evaluations against female teachers and generated a huge amount of attention both in academia and beyond. Its Altmetric attention score of 1257 is in the top 0.02% of all outputs tracked on Altmetric.

2. Underpinning research

The relevant research [R1] is part of Professor Mengel's research agenda on social identity and discrimination and how it can affect chances of success in the labour market. This includes research on in-group bias, gender differences in networking and gender bias in teaching evaluations [R2-R4].

The main reference [R1] studies student evaluations of teachers at a top Business School in the Netherlands. The research exploits a unique data set of almost 20,000 teaching evaluations filled in by students who are randomly assigned to teachers. The data further contains information on teacher performance allowing us to judge whether gender differences in teaching evaluations received can be tracked back to performance differences between male and female teachers. The research found evidence of bias in evaluation against female teachers and no evidence of performance differences. The effect is driven by male students and affects junior teachers in particular.

The research is unique in that it (i) shows causal, not just correlational evidence and (ii) allows to largely rule out performance differences between male and female teachers. The research also shows that effect sizes are substantial enough to induce a significant chance that teachers move from what is considered an acceptable score to a score that induces a performance monitoring process.

The bias has potentially harming effects on junior women's careers both directly (via reduced chances to be tenured or promoted) as well as indirectly by affecting their confidence and forcing them to reallocate resources from research to teaching.

This research has generated a huge amount of attention both in academia and beyond. Its

Altmetric attention score of 1257 (January 2021) is in the top 0.02% of all outputs tracked on Altmetric and the paper is the output published in JEEA with the highest attention record ever.

3. References to the research [can be provided by the HEI upon request]

[R1] Mengel, F., Sauermann, J. and Zoelitz, U. 'Gender Bias in Teaching Evaluations'. *Journal of the European Economic Association* 17(2) (2019), 535-566. https://doi.org/10.1093/jeea/jvx057

[R2] Chen, Y. and Mengel, F. "Social Identity and discrimination: Introduction to the special issue". *European Economic Review* 90, (2016), 1-3. http://dx.doi.org/10.1016/j.euroecorev.2016.10.002

[R3] Currarini, S. and Mengel, F. "Identity, Homophily and In-Group Bias". *European Economic Review* 90 (SI: Social Identity and Discrimination) (2016), 40-55. https://doi.org/10.1016/j.euroecorev.2016.02.015

[R4] Mengel, F. Gender differences in networking. *Economic Journal 130* (2020), 1842-1873. https://doi.org/10.1093/ej/ueaa035

4. Details of the impact

Essex research has led to improved teaching evaluation practices in Higher Education with evidence of change from 14 institutions across Europe, Canada and the United States. Specifically, at the University of Alberta, University of Birmingham, Brown University, University of Colorado, Hertie School of Governance Berlin, Maastricht University, McGill University, Oregon University, University of Rochester, Sheffield University, Tilburg University, Toulouse School of Economics and Union College. Based on the number of teachers and students at these 14 institutions, we estimate that the resulting changes have directly affected approximately 25,000 teachers and 425,000 students as well as many more indirectly.

The research [R1] has led to (i) changes to how teaching evaluations are conducted at several institutions including use of the research within teacher training, (ii) changes to how junior teachers approach teaching and view teaching evaluations, (iii) informing policy debate around teaching evaluation and public debate through press coverage and online communication of the research to lay public readerships.

Changing how teacher evaluations are conducted in Higher Education Institutions

The research has impacted on a wide variety of Universities across the regions discussed in terms of the establishing of task forces and reviews and changing institutional policies and practices.

The research was instrumental in establishing a task force to consider teaching evaluations at Oregon University and at Brown University (both United States), as a result, teaching evaluation practices have changed to counter gender bias.

At Brown University the committee's final recommendations include specific measures to mitigate the potential bias including changing questions from more holistic questions to questions about specific aspects of teaching quality and asking students to back up their scores with examples. Essex research [R1] is explicitly referenced in the final report including in the recommendations section [S1] and in the current teaching resources on how to interpret your course feedback [S1a]. Brown University piloted their new course feedback in spring 2018 and subsequently adopted these recommended changes in the 2019-2020 academic year [S1b].

In 2019, based on this research [R1], the University of Oregon (UO) set up a task force and subsequently introduced a holistic new teaching evaluation system that does more than simply replace problematic evaluation instruments so that they can help the UO community more effectively define, develop, evaluate, and reward teaching excellence [S2]. The task force originally convened to address the issues, set up a website with their results [S2a], and the Chair of the task

force confirmed that Essex research was crucial for the establishment of the task force, as well as their results.

> *'Without this evidence [R1] we would not have been able to help our University Senate to understand the potential harm that is being caused to female faculty when evaluation of teaching relies predominantly on potentially biased student ratings.'* [S2b]

The faculty council at Rochester University, Unites States, have decided to collect evidence from stakeholders (Deans, Heads of Departments) as to how they use teaching evaluations in direct response to the research [R1]. The chair of the curriculum committee at Rochester confirmed that the research was and remains crucial in this process.

> *"the Faculty Council at the University is going to establish a committee to study what teaching evaluations are actually used for by various stakeholders, such as deans, department chairs, faculty, and students."* [S3]

This research has been raised in connection with a motion to the General Faculties Council at the University of Alberta [S4], it has been extensively discussed in committees at Maastricht University and at Tilburg University. At Maastricht University the Dean for Education has asked for proposals of how to mitigate the problem [S4a].

At Hertie School of Governance a review is ongoing which was triggered by the research [S4b]. The McGill Tribune at Mc Gill University in Canada has published a piece urging their University to act upon our findings [S4c]. The magazine of Toulouse School of Economics in France also discussed our research [S4d].

Changing how teaching evaluations are approached and viewed

Mengel's research has been used in teacher training and has also featured in a "MicroCPD" video and webpage developed for educational training at the University of Birmingham [S5, S5a]. It was used in a teaching programme at University of Colorado, Colorado Springs in 2017 [S6]. Junior teachers at the University of Birmingham, at Maastricht University, the University of Sheffield and Tilburg University confirm that the evidence made them reflect on both their teaching style as well as procedures in their University regarding teaching [S7].

Informing political and public debate around gender bias in teaching evaluations

The research has been used to support the case made by the Norwegian Academy of Young Researchers for changes to criteria for teaching competence at all Norwegian educational institutions. The research was cited specifically to support their recommendation that assessment of teaching quality should not give too much importance to student feedback [S8]. The German Economic Association has discussed the research in an email sent to its members and has set up a panel to discuss the issue of gender bias in economics where Mengel presented the research [S9].

Professor Mengel's research has led to an increased understanding of gender bias generating a considerable amount of public discussion. The research has been tweeted about by more than 2000 tweeters (January 2021) located in Europe, North and South America, Asia, Australia and Africa with an estimated **6,506,927 of followers (Altmetric, January 2021)**. It has also been discussed in news articles in several major national and international newspapers and radio programmes with a combined readership of over 2.5 million including in The Economist, Sueddeutsche Zeitung (a German national newspaper), in a radio-show on Radio Bavaria (Germany), the Wire (India), Scientific America, Forbes, World Economic Forum , New York Times and others [S10].

5. Sources to corroborate the impact

[S1] Brown University Report on the Student Course Evaluation Instrument (page 8, 11)
[S1a] Brown University: Interpreting Your Course Feedback
[S1b] 2017-18 Review of Brown's Student Course Feedback Form
[S2] Office of the Provost: Revising University of Oregon's Teaching Evaluations
[S2a] Website with results from the Oregon Teaching Evaluations Task Force
[S2b] Testimonial from Oregon University stating that the research had a direct impact on establishing a task force at the University Senate.
[S3] Testimonial from University of Rochester stating that the research had a direct impact on the use of teaching evaluations within the University and will be used to establish a committee to consider this area.
[S4] University of Alberta. How should teaching evaluations at the University of Alberta be used?
[S4a] Maastricht correspondence about gender bias in student evaluations
[S4b] Testimonial from Hertie School of Governance Berlin, stating that the research has triggered an evaluation phase at the school.
[S4c] The McGill Tribune. Evaluating gendered bias in course evaluations
[S4d] Toulouse Economist. Toulouse Economist's Missing Women
[S5] The use of the research at Birmingham can be demonstrated by screenshots of the relevant webpages
[S5a] A testimonial evidencing the use of the research in PGCHEP teacher training.
[S6] Email from University of Colorado Announcements that confirms the use of the research in training there.
[S7] Testimonials from junior teachers at Tilburg University, University of Birmingham, Maastricht University and the University of Sheffield as well as a report to the senate at Tilburg University.
[S8] Report by the Norwegian Academy of Young Researchers, submitted to the Norwegian government, which cites the research in support of their recommendations.
[S9] German Economic Association Newsletter to members and invitation to speak at panel about Gender Bias in Teaching Evaluations. Invitation to speak also at European Women in Mathematics.
[S10] Collated media coverage in the form of screenshots or copies of the articles covering the research.