Notes on Student Evaluations of Teaching (SET)
Last edited 8 September 2020.

Experimental and quasi-experimental results

- weak or negative association with objective measures of learning (Carrell & West, 2010; Braga et al., 2014; Boring et al., 2016)
- substantial bias from gender
  - gender bias can make female instructors rate worse than objectively less effective male instructors (Boring et al., 2016)
  - bias affects ratings of “objective” items like promptness (MacNell et al., 2015; Boring et al., 2016)
  - varies by discipline &c (Boring et al., 2016; Mengl et al., 2018)
- bias from ethnicity & gender (Chisadza et al. 2019)
- strong association with grade expectations, but not necessarily with learning (Boring et al., 2016)
- grades—not learning—“rewarded” with high SET (Cho et al., 2015; Carrell & West, 2010; Braga et al., 2014)
- providing cookies during class increases ratings of instructors and of course materials (Hessler et al., 2018)
- the number of points on the rating scale affects gender differences (Rivera & Tilcsik, 2019)

Laboratory studies

- bias in favor of young male instructors (Arbuckle & Williams, 2003)
- ratings predicted by responses to 30 seconds of silent video (Ambady & Rosenthal, 2003)
- race and gender matter (Basow et al., 2013)

Meta-analyses

- weak or negative association with objective measures of learning (Uttl et al., 2016)
- association between author conflicts of interest and conclusions that SET are valid (Uttl et al., 2019)

Observational studies and $n = 1$ experiments

- strong association with student enjoyment (Stark, unpublished)
- data unreliable: substantial fraction of students give demonstrably–apparently deliberately–false answers to objective questions (Stanfel, 1995)
- gendered language in evaluations (Schmidt, 2015, *inter alia*)
• bias against older instructors and female instructors (Bianchini et al., 2013; Wagner et al., 2016)
• bias against non-native English speakers, (Subtirelu, 2015, inter alia), URM (Wagner et al., 2016)
• bias in favor of physically attractive instructors (Wolbring & Riordan, 2016; Feeley, 2002; Hamermesh & Parker, 2004)
• biases from physical condition of room, time of day, mathematical level, class size, . . . (Bedard & Kuhn, 2005, inter alia)
• “halo effect”: students conflate enthusiasm, attractiveness, & other things with effectiveness; enthusiasm not associated with learning (Williams & Ceci, 1997; Feeley, 2002; inter alia)
• negative association with learning (Stroebe, 2016)
• association with attractiveness, esp. for female instructors (Babin et al. 2020)

Surveys
• students deliberately falsify ratings (Clayson and Haley, 2011)
• students and faculty use the same adjectives differently (Lauer, 2012)
• comments incommensurable across disciplines (Stark & Freishtat, 2014)
• bias against quantitative classes (Uttl et al., 2013)

Statistical abuses (Stark & Freishtat, 2014)
• averages of categorical data are meaningless/misleading
• response rate matters
• report distributions, not means
• SET are an incomplete census, not a random sample. Nonresponse bias not ignorable.

Litigation/Arbitration
• U. Florida
• Ryerson U. https://ocufa.on.ca/blog-posts/significant-arbitration-decision-on-use-of-student-questionnaires-for-teaching-evaluation/
• U. Toronto
• UNLV
• American Sociological Association seeks to crowd fund a class action https://www.change.org/p/american-sociological-association-end-the-use-of-biased-student-evaluations-of-teaching-in-employment-decisions-e3ff4761-0d64-4b51-9fce-f160d743e690/sign

Policy changes at other universities
• USC https://academicsenate.usc.edu/teaching-evaluations-update/
• U. Oregon https://provost.uoregon.edu/revising-uos-teaching-evaluations
SELECTED REFERENCES

- Basow, S., S. Codos, and J. Martin, 2013. The Effects of Professors’ Race and Gender on Student Evaluations and Performance, *College Student Journal, 47*(2), 352-363.
- Cho, W., W. Baek, and J. Cho, 2015. Why do good performing stu-


• Subtirelu, N.C., 2015. “She does have an accent but…”: Race and language ideology in students’ evaluations of mathematics instructors on RateMyProfessors.com, *Language in Society, 44*, 35-62. DOI 10.1017/S0047404514000736

• Uttl, B., C.A. White, and A. Morin, 2013. The Numbers Tell it All: Students Don’t Like Numbers!, *PLoS ONE, 8*(12): e83443, DOI 10.1371/journal.pone.0083443


