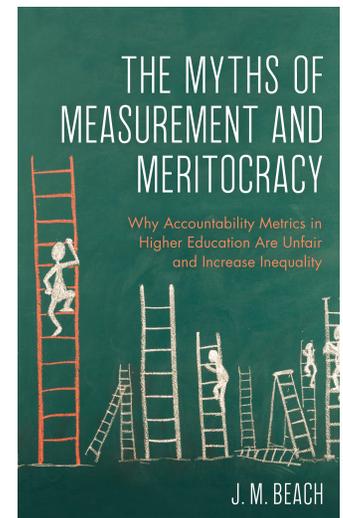


Getting Beyond the Flawed Metrics of Student Evaluation Surveys

Josh M. Beach
21st Century Literacy
Austin, TX

Transforming Institutions Conference, June 2021



Do Students Learn in College?

- The Spellings Commission on the Future of Higher Education:
- “lack of useful data and accountability” metrics by colleges and universities, which meant that there was “no solid evidence” for student learning or other outcomes (4, 13).
- No one actually knew which colleges or universities “do a better job than others of not only graduating students but of teaching them what they need to learn” (vii).

Do Institutions of Higher Education Cause Student Learning?

- Donald Levine, a former dean of the University of Chicago:
- “The scandal of higher education in our time is that so little attention gets paid, in institutions that claim to provide an education, to what it is that college educators claim to be providing”

(qtd in Delbanco 2007).

Do Institutions of Higher Education Cause Student Learning?

- Up until the 21st century, Scholars found “little consistent evidence to suggest that either postsecondary education in general or the type of institution attended in particular had a differential effect on knowledge acquisition for different kinds of students”

(Pascarella and Terenzini 2005, 134; Keeling and Hersh 2012, 27).

Student Evaluation Surveys

- Common Sense Assumptions
 - valid accountability instruments
 - accurately measure the means and ends of education.
- Based on Traditional Assumptions of Schooling
 - Teachers *produce* student learning
 - Most students learn in school
 - Students easily recognize their learning
 - Teachers are responsible for student learning and customer satisfaction.
- Surveys are high-stakes accountability instruments for adjunct instructors

Common Sense Assumptions FALSE

- Surveys are improperly designed, calibrated, administered, and interpreted
 - Surveys do not measure either teaching or learning
 - Teachers do not control student learning
 - Students can't easily recognize their own learning
 - Many students don't want to learn
 - Surveys are biased and unfair to teachers

Early Student Evaluation Surveys

- Hermann Henry Remmers, 1928
- Relationship between grades and students rating of professor
- Traditional “teacher-effectiveness theory”
 - Teachers control student learning
 - Students want to learn
 - Students learn what teacher tells them to learn
 - Grades reflect learning
 - High grades = high learning

Early Student Evaluation Surveys: Bias?

- Remmers, Martin, & Elliot (1949)
- Different kinds of students rated same teacher differently
 - Advanced students rated teacher high
 - Weaker students rated teacher low
- When teacher adjusted difficulty down
 - Advanced students rated teacher low
 - Weaker students rated teacher high

New Psychology & Attribution Theory

- Mind NOT a passive receiver of stimulus
 - Mind subjectively interprets
- Behavior not controlled by stimulus
 - Mind creatively acts upon stimulus
- Attribution theory
 - How people construct their subjective world and personal identity through interpretation
- Humans interpret world with a “self-serving bias” (Pinker, 2011, p. 490)

Insights from the New Psychology

- Students construct “good student” identity to “look smart” instead of learn or become educated (Dweck 2002, p. 24)
- Students act to maintain positive self-image and high self-esteem
- People have “unrealistically positive views” of themselves (Edmondson, 2012, p. 155)
- Dissonance theory/Confirmation bias: People focus on information that confirms positive self-image and prior beliefs. They distort, discount, ignore negative information (Bruner, 1983, p. 114)
- People defensively use “anti-learning” techniques to protect themselves from embarrassment or threat (Argyris, 1993, p. 15)
- Self-handicapping: Fearful people use self-limiting behavior to maintain status quo and not learn or do new things (Pfeffer 2010, 13)
- People lie to themselves and others: protect their self-identity & self-esteem (Ariely, 2012, pp. 158, 165; Tavis & Aronson, 2015).

Insights from Sociology

- Students are not “raw material” that teachers can control, forcing learning (Waller, 1932; Milner, 2006, p. 18)
- Students have “a mind of their own,” and they often “resist being transformed” by teachers and schools (Milner, 2006, p. 18)
- Students forced to go to school by parents, school officials, police, & labor market
- Students resist learning, subvert the official pretext of instruction and academic success (Steinberg, 1996)
- School as a social environment to be with friends (Steinberg, 1996)

The Grade Satisfaction Theory of Schooling: *A Quid Pro Quo* “Compromise”

(Sizer, 1992; Labaree, 1997; Labaree, 2010)

- Teachers ritualize schooling, water down curriculum, inflate grades
- Teachers please administrators, parents, and students
- Students resist learning, but feel subjectively successful
- Low standards for less work, little or no learning
- Easy assignment for high grades
- Students attend, comply, complete assignments, & reward teacher with high survey ratings
- Teachers conform to school policies, meet accountability criteria, So administrators give freedom, praise, raises, & promotions

What Is the Result of Schooling?

- Students can only “plug and chug” useless, abstract information, “memorizing formulae, sticking numbers in the right equation or the right vocabulary into a paper... understanding little. When the class is over, they quickly forget much of what they have ‘learned.’

(Bain, 2004, p. 24)

- “Overwhelming body of educational research” has shown that “even those students who have been well trained and who exhibit all the overt signs of success – faithful attendance at good schools, high grades and high test scores, accolades from their teachers – typically do not display an adequate understanding of the materials and concepts” they were supposed to learn in school.

(Gardner, [1995] 2011, p. 3)

Student Evaluation Surveys

Do Not Measure Teaching or Learning

- Rodin and Rodin (1972)
 - Correlation of $-.746$ to $-.754$ between surveys and grades
 - Students who learned the most with highest grades gave TAs lowest ratings
- Holmes (1972),
 - Students with unexpected low grade → lower ratings for teacher on 10/19 items
- Johnson (2003)
 - Grades explained 36% evaluation rating; prior student interest 30%
 - 66 percent of teachers rating not related to teaching or curriculum
- Uttl (2017)
 - No correlation between survey scores and student learning
 - Students do not learn more from teachers with high ratings
 - At best, surveys explain 1% of variance in student learning

Student Evaluation Surveys Measure Student Grade Expectations

“Students’ responses to the survey were significantly affected by the grades that the students either expected to receive or already had received. For most items, the influence that students’ grades had on their responses to the survey ranged from about one-fourth to one-half of the importance of the consensus rating...grades do, in fact, represent a serious bias to student evaluations of teaching...It seems that students measure their success in a course, and implicitly the quality of instruction, by the grade they receive or expect to receive. Poor grades are thus associated with poor teaching; students who receive low grades tend to denigrate instruction when they complete teacher-course evaluation forms” (Johnson, 2003, p. 100)

Insights from Statistics

Evaluation Surveys Invalid Measurements

- Evaluation numbers are statistically meaningless (Stark & Freishtat, 2014, pp. 3, 5)
- Polarized student opinions are averaged: “Such averages and comparisons make no sense, as a matter of statistics” (Stark & Freishtat, 2014, pp. 5-6).
- Many nonsensical answers:
 - Rating teacher highly + rating the teacher poorly does NOT mathematically equal the instructor’s teaching was mediocre.
 - 2.78 for “the instructor was available during office hours.” Either the instructor held office hours or not. Score violates Law of the Excluded Middle: “No middle ground between true and false” (Sigmund 2017, 211).

Insights from Statistics

Evaluation Surveys Invalid Measurements

- Ratings averaged down into a single index number to rate and compare individuals
- Nobel Prize economist Paul A. Samuelson mathematically demonstrated, all index numbers are flawed in practice because “there will always be some bias which will make even the most ideal index number subject to some ambiguity” (Breit and Ransom 1998, 114)
- Statistician and management professor W. Edwards: Natural variability in all data; impossible to precisely compare individuals using any statistical model.
- Only three statistically meaningful categories:
 - People outside the control limits “on the bad side”
 - People outside the control limits “on the good side”
 - Most people who fit within the control limits (Walton 1986, 92; Aguayo 1990)
- No way to accurately disaggregate the majority of people in the middle within the control limits.
- Everyone in this category should receive the same rating with the same rewards (Walton 1986, 92)

Evaluation Survey Discriminatory & Unfair

- Based on subjective emotions & beliefs of students, beyond instructor's control
- Teachers with high standards & high student learning get low ratings
- Ratings are correlated with prejudices
 - race
 - nationality
 - gender
 - attractiveness
- Ratings “hypersensitive”
 - “influenced by the tiniest and most irrelevant things, from the comfort of the seats to the time of day the course is offered” (Nichols 2017, 97; Pascarella and Terenzini 2005, 115)

Can Students Evaluate Teacher or Learning?

The Fox Effect

- People use “likability” heuristic when evaluating others
 - Attractiveness, gender, race, personality, personal beliefs
- Judgments subconsciously formed within seconds
- “Rapid social judgments” have “no validity” in terms of competence
(Tversky 2019, 46)
- Just because you don’t like someone, doesn’t mean you can’t learn from them. Successful people learn from all experiences, even from people who are different or who you don’t like/respect (Pfeffer 2010, 8-9)

Students Want “Edu-tainment,” NOT Learning

- Waller (1932): a good teacher is a good actor (p. 233)
- Students want “edu-tainment,” rather than education (Pfeffer 2015, 25, 29)
- Naftulin, Ware, and Donnelly (1973); Ware & Williams (1975, 1976, 1977): “Fox Effect”
- Professor taught undergrads, grads, & working professionals
 - 97 % was interested in his subject.
 - 91 % used adequate examples.
 - 87 % aroused their interest in the topic.
 - 81 % teaching was interesting
 - 70 % was well organized

Students Want “Entertainer,” NOT Teacher

- Professor was a paid actor pretending teach
- Lectures false & nonsense
- Entertaining = high teacher ratings
- Quality of lecture/curriculum NOT correlated with teacher ratings, but highly correlated with exam scores
- Teacher evaluations NOT correlated with student learning
- Many studies have replicated these findings

Are Preferences Unconscious?

- Ambady & Rosenthal (1993)
- Control group of students: only three ten-second video bits, no sound
- Second group of students: took whole semester class
- Teacher evaluation ratings essentially the same
- Study suggests students unconsciously formed opinions in seconds based on physical appearance & non-verbal communication
- Study Replicated: political science & organizational leadership

Leadership = Acting ?

- Most people expect “leaders,” which includes teachers, to act out a stereotypical script of leadership.
- People expect leaders “to act like a leader, to act in a way that inspires confidence and garners support”
(Pfeffer 2015, 98-99)
- Ohio State University undergraduate: “A good prof in a lecture course is an entertainer...a bad prof is a prison guard” (qtd. in Sperber 2000, 87)

Johnson (2003)

The use of student evaluations of teaching for administrative purposes and as measures of overall teaching effectiveness has been an unqualified failure. Not only has their use for these purposes had the unintended consequence of altering the dynamics of student-instructor interactions in ways that are not yet fully understood, but current teacher-course evaluation forms are, at best, only modestly correlated with actual student achievement (p. 151)

Stop Using Student Surveys as High Stakes Teacher Evaluations

- 18 scholarly associations
- Student evaluation surveys should not be relied upon “as a measure of teaching quality” (Supiano 2019, para. 2)

Conclusion:
Make Surveys Low-Stakes, Data Tool in Classroom

Stark and Freishtat (2014): "It's totally valuable to ask [students] about their experience, but it's not synonymous with good teaching."

Make Student Surveys Low-Stakes Use as Data Tool in Classroom

- Teachers should design student surveys and administer them often in classroom
- Surveys should explore student background, motivation, & effort
- Surveys should be aligned with curricular objectives
- Data should be used for purposes of teaching and learning
- Do NOT use surveys for employment, rewards, or punishment
- Administrators should not see survey data

But....

Do students really want an education and learning?

- Even if it involves hard work, stress, failure, negative emotions, lower self-esteem, & self-doubt?

Or, do students want less work, high self-esteem, more time with their friends, easy to obtain credentials, social prestige, high-paying jobs?

But....

Do instructors really want (or know how) to teach?

- Even if it involves more work, more complaints, and failure?

Or, do instructors want to minimize classroom time, keep instruction ritualized and less work, so they can focus on their research?

But...

Do college administrators really want instructors to teach?

- Even if it involves higher costs, more complaints from students, increased failure, and more complaints from tax payers, the media, and politicians?

Or, do administrators want low costs, organizational efficiency, less work, social prestige, and ritualized traditions

Question?

- Is it possible for schools to actually educate students, rather than school them?

J. M. Beach

CAN WE MEASURE WHAT MATTERS MOST?

Why Educational Accountability Metrics
Lower Student Learning
and Demoralize Teachers



Find this presentation

& more information

at

www.jmbeach.com

