

Course Teacher Evaluation: Biases and Best Practices

UUP Union Matters Event
March 11, 2020

UUP Working Group on CTEs
Presented by
Professor Tom Pasquarello
Political Science Department





Agenda for today's discussion

1. The origin of SUNY Cortland's policies regarding the use of CTEs
2. Review of recent studies and professional statements on CTEs
3. The status of CTEs at other institutions and in the courts
4. Starting a discussion about the future of CTEs at SUNY Cortland



The origins of our CTE system

- College policies regarding CTEs originate in a report from 1982:

FINAL REPORT
OF THE 1981-82 COMMITTEE FOR THE EVALUATION OF TEACHING EFFECTIVENESS AT
SUNY CORTLAND
TO
SUNY CORTLAND COLLEGE FACULTY SENATE

- The last 40 years have seen the publication of many scientific studies establishing deep and irrevocable problems with CTEs, yet this report from 1982 remains the basis of our CTE system today.



The origins of our CTE system

- The 1982 report acknowledged fundamental problems with CTEs:

[CTEs] can play a limited role **when used with caution** as one part of a teaching evaluation system.

Despite the **imprecision of CTE's as valid scientific indicators of teaching quality**, however, they can play a limited role **when used with caution** as one part of a teaching evaluation system... It is essential to keep in mind that **CTE's reflect student perceptions of teaching, not the quality of teaching** in any absolute sense. (pp. 5)

[Fear of a sort of Nielson-rating contest between faculty members] can be minimized, however, by a system which allows the faculty to control the evaluation process and by a college philosophy that places academic and professional standards above a reliance on statistical ratings of teaching. (pp 6)

... the systematic use of CTE's as a component must be designed to reassure faculty members that two common abuses with CTE's will be guarded against. The most common abuse is for evaluators to use CTE results as a highly reliable statistical indicator of the quality of one's teaching... Moreover, faculty members must be reassured that any data drawn from CTE's will be used as longitudinal data over a period of several years and several courses. (pp.g 6)



The origins of our CTE system

- An updated report on CTEs was produced in 2002, but largely copied the 1982 report without any substantial modifications.
- The 2002 report conflates CTEs with a holistic sense of evaluation, evident in the report title:

Course Teacher Evaluation At SUNY Cortland:

A guide to Teacher Evaluation at SUNY Cortland

- The most recent citation in the 2002 report dates to 1982.
 - 2 out of 19 citations are from the 1950's, 11 are from the 1970's, the remaining 7 are from the early 1980's.

How our understanding of CTEs has changed since 1982



The problem of bias: CTEs scores exhibit systemic bias with regard to race, gender, age, disability, charisma, and other factors, and often include comments that would constitute harassment in other contexts.

Basow, Codoes and Martin. “The effects of professors’ race and gender on student evaluations and performance.” *College Student Journal* **47**, 352 (2013).

Anderson and Smith. “Students’ preconceptions of professors: benefits and barriers according to ethnicity and gender.” *Hispanic Journal of Behavioral Sciences* **27**, 184 (2005).

Mitchell and Martin. “Gender bias in student evaluations.” *PS: Political Science & Politics* **51**, 649 (2018).

MacNell, Driscoll and Hunt. “What’s in a name: exposing gender bias in student ratings of teaching.” *Innovative Higher Education* **40**, 291 (2015).

... and many others.

How our understanding of CTEs has changed since 1982



Absence of significance: There is no significant evidence relating CTE scores to actual teaching effectiveness.

Boring, Ottoboni and Stark. “Student evaluations of teaching (mostly) do not measure teaching effectiveness.” *ScienceOpen Research* (2016).

Uttl, White and Wong-Gonzalez. “Meta-analysis of faculty’s teaching effectiveness: student evaluation of teaching ratings and student learning are not related.” *Studies in Educational Evaluation* **54**, 22 (2017).

The problem of negative correlations: studies have shown that student evaluations of professors in introductory math & science courses are negatively correlated with long-term success.

Braga, Paccagnella and Pellizzari. “Evaluating student evaluations of professors.” *Economics of Education Review* **41**, 71 (2014).

Carrell and West. “Does professor quality matter? Evidence from random assignment of students to professors.” *Journal of Political Economy* **118**, 409 (2010).

How our understanding of CTEs has changed since 1982



Unbiased, reliable, and valid student evaluations can still be unfair.

Esarey and Valdes. *Assessment & Evaluation in Higher Education*, 1 (2020).

The take-away:

Even ideal circumstances produce an “unacceptably high error rate.”

This study simulated randomized distributions of instructors using idealized CTE scoring (considered to be reliable, valid, and free of bias) using known correlations between actual teaching effectiveness and CTE score, and found that many excellent instructors still receive poor reviews that would lead administrators to incorrect decisions.

<https://www.insidehighered.com/news/2020/02/27/study-student-evaluations-teaching-are-deeply-flawed>



Summary of findings

- **CTEs have systemic problems** including statistical unreliability, negative correlations in some cases, and have issues of racial, gender, and other kinds of bias in student responses.
- **CTEs do not necessarily measure teaching effectiveness** and may have more to do with grade expectations, as established by many scientific studies.
- **Interpretation of CTEs is not simple:** high CTE scores can be a sign of good teaching or easy grading, and low CTE scores are not clearly an indication of bad teaching and may actually represent superior teaching.
- **Extreme caution is needed** in interpreting CTE results, and simple reasoning can lead to incorrect conclusions.



Position of professional organizations

From the American Sociological Association (ASA), September 2019:

(and endorsed by 19 other professional organizations)

1. Instruments are opportunities for student feedback about their course experience, not formal ratings of teaching effectiveness.
2. SETs should be one part of a holistic self-assessment of teaching effectiveness, not the sole measure of teaching effectiveness.
3. SETs should not be used to compare individual faculty members.
4. Quantitative scores should include distributions, sample sizes, response rates.
5. Evaluators (e.g., chairs, deans, personnel committees) should be trained in the interpretation of SETs.



The status of CTEs at other institutions

Ryerson University (Toronto, Ontario, Canada)

SETs are no longer “used to measure teaching effectiveness for promotion or tenure,” a result of resolution by the Ontario courts. (Farr, 2018)

Alternatives: Multi-modal teaching dossiers with peer evaluation to foster “an ongoing process of inquiry, experimentation, and reflection.”

University of Oregon (Eugene, Oregon)

Existing collective bargaining agreement updated to include a “commitment to evaluating professional, inclusive, engaged and research-led teaching practices” beyond student ratings.

Alternatives: Use of mid-term course surveys and end-of course “student experience evaluations;” new instructor self-reflection tool; new tool for peer review in departments; establishment of standards for effective teaching practices.



The status of CTEs at other institutions

University of Southern California (Los Angeles, California)

SETs will no longer be a part of decisions pertaining to tenure, promotion, or any other performance review, though SETs will still be given “to help faculty adjust their teaching practices.” (Flaherty, 2018)

Alternatives: Revising instrument; developing peer-review evaluation; sustained focus on incentivizing teacher professional development.

Stanford University (Palo Alto, California)

SETs still used as one measure of teaching effectiveness; changes to the instrument made allowing instructors to customize form with respect to their individual course with more qualitative/written feedback.

Alternatives: Along with revised SETs, mid-term feedback, reflective teaching portfolios, and peer evaluations are employed.



Status of CTEs at other SUNY schools

- **UUP state-level** is starting discussions with SUNY to review concerns regarding CTEs across all campuses.
- **Other UUP chapters** at multiple campuses have raised concerns about CTEs through Labor Management.
- **Empire State College** has been engaged in local negotiations between UUP and campus administration since 2016 and is currently in mediation.

The continued use of CTEs for personnel decisions requires further discussion



It is not clear that a majority of faculty have confidence that CTEs are being used appropriately in the current system of evaluation. Answers to the following questions may help guide a campus discussion.

- What institutional structures or policies are needed to ensure that CTEs are actually considered with caution?
- How do we correct/account for bias in CTE scores and how do we counter harassment in CTE comments?
- What recourse do we have for challenging inaccurate CTEs?
- How can CTEs be responsibly used to assess teaching effectiveness when they actually measure something else and are not statistically reliable?



Updating College Policies on CTEs

- **We must, at the least, update our policies** to reflect contemporary acknowledgement and understanding of issues of bias, statistical unreliability, and susceptibility to misinterpretation that create deep problems for responsible use of CTEs in personnel actions.
- **In many cases, CTEs play an excessive part in evaluations** in contrast to the “limited” role mandated by College policies, which is especially true for contingent faculty for whom CTEs can comprise the bulk of the evaluation.
- **Faculty have the authority** to change how CTEs are used in evaluations, as stated in College policies. This is an area that can be renegotiated through UUP.
- **SUNY Cortland can become a leader in the SUNY system** by developing a new guidelines for evaluation that are responsive to the science, ethically defensible, and in which faculty have confidence.



State-level policy for the College

- **The SUNY Board of Trustees (BOT) policies** in Article XII, regarding the Evaluation and Promotion of Academic and Professional Employees, reads:

In conducting evaluations pursuant to this Title, the chief administrative officer of the college concerned, or designee, **may consider**, but shall not be limited to consideration of, the following:

(b) Effectiveness in teaching - as demonstrated by such things as judgment of colleagues, development of teaching materials or new courses and student reaction, as determined from surveys, interviews and classroom observation.

- This means that BOT policies do not require the use of CTEs in performance evaluations.



SUNY Cortland policy for departments

- **College policy for personnel actions** was defined in a 1986 memo from former Provost Charles Warren:

The degree to which information about course teacher evaluations (CTE's) and course dossier evaluations (CDE's) is transmitted to administrative officials in the making of personnel decisions **is solely the judgment of each department.** It should be noted, however, that the Deans' Cabinet's current policy stipulates that no positive decision will be rendered on promotion, retention, tenure, or merit unless clear and persuasive evidence is offered of an instructor's teaching effectiveness.

- Therefore, the authority for determining how CTEs are used in personnel actions falls to the departments.
- Do we need new institutional limits on the weight given to CTEs given the known problems of bias and unreliability?



SUNY Cortland policy for members of personnel committees

- **Individual faculty are given substantial authority** in personnel actions, as defined in Section 220.07(E) of the SUNY Cortland Handbook:

While some types of evidence may be more important than others, it is the function of the recommender to judge the weight and quality of each item of evidence [in a portfolio].

- Therefore, if you believe that a piece of evidence, including an assessment tool, is flawed or biased, **it is the responsibility of individual members of personnel committees** to evaluate what weight it should be given.



Three things you can do

- **Ask your Faculty Senate representative** to bring this important issue to the floor for discussion with the administration.
- **Talk with your colleagues** about the studies and alternatives to CTEs.

<https://uuphost.org/cortland/>

- **Take the UUP survey on CTEs** - it takes less than 10 minutes!



<https://www.surveymonkey.com/r/UUPCortlandCTEs>



The ongoing UUP Survey on CTEs

1. Do you think CTEs are used appropriately or inappropriately for...

	Always or almost always used appropriately 1	Generally used appropriately 2	Sometimes used appropriately, sometimes used inappropriately 3	Generally used inappropriately 4	Always or almost always used inappropriately 5	Rarely or never used for this purpose	Don't know
Faculty development	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Faculty mentoring	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Renewal/ Reappointment decisions	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

2. Do you feel that policies regarding the use of CTEs in your department are clear for...

	Policies not at all clear 1	A little clear 2	Somewhat clear 3	Mostly clear 4	Policies extremely clear 5	Don't Know
Renewal/Reappointment	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Promotion	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Tenure	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



<https://www.surveymonkey.com/r/UUPCortlandCTEs>



SUNY Cortland Faculty Senators

- **School of Arts & Sciences:**

Li Jin

Anna Curtis

Brian Williams

Chris Badurek

Donna West

Ross Borden

Jenn McNamara

Herbert Haines

- **School of Education:**

Alexis Blavos

Vierne Placide

Alyson Dearie

Tadayuki Suzuki

Regina Grantham

Shufang Shi

- **School of Professional Studies:**

S. Shi

Vierne Placide

Jacob Hall

Regina Grantham

- **Library & Professional Staff:**

Rich Powell

Sonya Commins

Jeremy Zhe-Heimerman

Melissa Fox

Carol Costell-Corbin

April Myles