

## Minutes for the Minority Affairs Meeting Agenda

Tuesday January 21, 2014

10:00 a.m. – 11:00 a.m.

Auraria Library Friends Conference Room

**Attendees:** Ji Chen (Business); Philip Joseph (English); Denise Pan (Library); Stephanie Santorico (CLAS); Craig Sisneros (Business); Brenda Allen (Diversity and Inclusion); Lando Archibeque (Library); Charles Musiba (CLAS)

### Meeting Agenda

1. Review minutes from December meeting

Minutes approved. No changes from circulated draft.

2. Planning for focus group to investigate the potential impact of FCQs on the retention of underrepresented racial-ethnic faculty
  - a. FCQ research – preliminary data from IR (Craig and Stephanie)

Analyses were conducted for FCQ course and instructor ratings. Analyses were conducted twice with respect to classification of minorities: minorities and underrepresented minorities. In general, minority faculty score less, female faculty on average score less, female minority faculty have a small increase in FCQ score over male minority faculty.

Suggestions from Brenda: Would be interesting to distinguish between different disciplines, e.g., social sciences, humanities and STEM. Also, how does the FCQ question regarding “Instructor treatment of ethnic minorities and women” relate to FCQ score? Is there any access to comment data?

- b. Review dates and details
  - i. March 6 at 11-12:15 -- Confirmed with Dr. Allen
  - ii. Catering -- \$500 budget funded by the Office of Diversity and Inclusion
  - iii. Room reservation – Auraria Library Enhanced Learning Center
  - iv. Participant gift – Auraria Library flash drive with Literature Review by Lando Archibeque

Meeting will collect qualitative data. Thanks to the Office of Diversity and Inclusion for providing funding the catering service!

Meeting will be held in the library so there is no facility use cost. Lando will be doing a literature review to be distributed at the meeting.

c. Marketing/Promotion/Planning Discussion

i. Event goals and target audience

Philip suggested opening up the meeting to others interested in the use of FCQs in the evaluation of teaching.

Goals:

Document the problem. Relate to the merit process and the effects of weighing FCQs in the context of merit. Move beyond documentation to make recommendations. (Philip)

How can we utilize the qualitative data from FCQs? How do recommendations relate to our committee's charge? Note that benefitting the larger population of faculty also benefits minority faculty. (Brenda)

An APS exists regarding multiple means of teaching. Can this be utilized as a point of discussion when moving beyond documentation of issues? (Brenda)

What are the quantitative and qualitative measures that are problematic when it comes to evaluation of minority faculty? Regarding peer evaluation, most faculty are not trained in pedagogy. (Craig)

Philip: The English Dept. is breaks down teaching as follows. FCQ 's can be applied to the first category (Student Learning and Growth), but there are other forms of evidence for that category, as well:

- Student Learning and Growth, 40%
- Teaching Materials, 20%
- Instructor Development, 20%
- Curriculum Innovation & Development, 20%

ii. Event name: Reconsidering FCQs (proposed title)

iii. Venues to promote event:

meetings - FA (First Tuesday noon-1:30); AD/ASG (First Tuesday 8:30-10:00 AM); others?  
emails - Today at CU Denver; FA; AD/ASG; others?

iv. Action items and due dates to assign to MAC members: Create flyer, online registration form, handouts (lit review and other items?), moderator speaking notes, other items?

Invite: Target audience as inclusive as possible to identify problems and biases in the FCQs with respect to minority faculty.

Format: Present data. Provide discussion questions and use small focus groups.

Philip will work on wording of information to be disseminated.

Craig will take the lead on summarizing institutional data.

Philip suggested getting a list of minority faculty to target an invitation specifically to them.

- d. Agenda/Itinerary 11:00-12:15 (proposed, based on schedule from Inclusion Dialogues)
  - i. 11:00 arrival and self-serve lunch
  - ii. 11:15 begin program, presentation of findings and discussion
  - iii. 11:20 begin discussion re: defining the problems/issues.
  - iv. 11:40 discussion re: identifying solutions/strategies
  - v. 12:00 Wrap-up: Call to action
3. Other updates or announcements?
  - a. Dr. Allen invited a MAC member to represent CU Denver on a Steering Committee for the Metropolitan State University's annual Higher Education Diversity Summit.
    - i. Denise volunteered to participate and represent CU Denver and MAC.
  - b. Craig and Stephanie will be meeting with Institutional Research on Friday January 24, 2014.
  - c. March 6 at 11-12:15, "Reconsidering FCQs" event.

Brenda: Thanks to Philip for serving on system level commitment to excellence committee. They are looking at the process of evaluation for these proposals, e.g., reviewing in diversity office first and then prioritizing the same number from different campuses.

April 17 – System level Diversity Summit (will be on AMC): heighten awareness of faculty and staff with respect to diversity and inclusion. What kinds of topics would we like to see as part of the breakout sessions? Do we have suggested speakers or facilitators? Will have information coming from institutional research, human resources, and library.

**Next Meeting: February 11, 2014 from 10am-11am in Auraria Library Friends Conference Room.**

**Minutes for the Minority Affairs Meeting Agenda**  
**Tuesday March 11, 2014**  
**10:00 a.m. – 11:00 a.m.**  
**Auraria Library Friends Conference Room**

**Attendees:** Lando Archibeque (Library); Farah Ibrahim (SEHD); Philip Joseph (English); Charles Musiba (CLAS); Denise Pan (Library); Stephanie Santorico (CLAS)

**Meeting Agenda**

1. Review minutes from February meeting – [Approved](#)
2. “Reconsidering FCQs” Event Planning
  - a. Videotaping session – [funded by Office of Diversity and Inclusion](#). We will inform participants of the recording and that the recording will only be made available to the PTR committee given the time conflict. Note: we will ask media services to arrange for a microphone that would not be on video camera. In addition, we will provide note cards for participants to provide questions and/or suggestions that would be read by a MAC Committee member.
  - b. Review registered attendee list - [19 people registered as of today](#)
  - c. Event schedule –
    - i. [11:30 -- Arrival and boxed lunches.](#)
    - ii. [11:45 -- MAC presentation of OIRE findings and literature](#)
    - iii. [12:00 -- Audience discussion: Defining the problems/issues.](#)
      - [Do assumptions about gender, race, sexual orientation, national origin and language shape FCQ student responses?](#)
      - [What is the potential impact of FCQs on the underrepresented racial-ethnic faculty of CU Denver?](#)
      - [Are FCQs an accurate measurement of effective teaching?](#)
      - [Are there other limitations/hurdles in the tenure and promotion process for underrepresented racial-ethnic faculty?](#)
    - iv. [12:20 -- Audience discussion: Identifying solutions/strategies](#)
      - [What are the alternative methods of evaluating teaching?](#)
      - [How can we support faculty in the tenure and promotion process?](#)
      - [How can we educate students on providing constructive feedback?](#)
    - v. [12:45 – Inclusion Dialogue Wrap-up: Call to action](#)  
[Let people know about our Fall event focusing on how we can change evaluation of teaching so it is not so heavily focused on FCQs \(on a department and an individual level\).](#)
    - vi. [1:00 – Conclusion](#)
  - d. Room setup/shutdown: We will use the room as currently set up. Stephanie will bring notecards and a box for anonymous questions and/or suggestions. Denise will create name tags and copies of handout. Bryan (CU Denver Media Production), Denise, Stephanie will be at room at 10:30. Others invited to assist as well.

3. MAC Operating Procedures (Denise sent via email on Feb 21) – Changes made by Faculty Assembly were reviewed and approved.
4. Other updates or announcements?
  - a. EMAC (Philip) – The committee has looked at a New York Times Opinion Editorial: [http://opinionator.blogs.nytimes.com/2014/02/05/improving-economic-diversity-at-the-better-colleges/?\\_php=true&\\_type=blogs&\\_r=0](http://opinionator.blogs.nytimes.com/2014/02/05/improving-economic-diversity-at-the-better-colleges/?_php=true&_type=blogs&_r=0). Of particular interest, Franklin & Marshall College has shifted money away from merit based scholarships to needs based scholarship, demonstrating that this is working to both increase the depth of student learning and diversity. EMAC is bringing this to the university chief diversity officers to consider.
  - b. Higher Education Diversity Summit (Denise) – All day event on Thursday April 10: <http://aurariaevents.wix.com/heds>  
Theme: Civil rights @ 50: Advancing Equal Access to All  
Keynote speaker: Lily Ledbetter

**Reconsidering FCQs Event: March 14, 11:30am-1pm**

**Next Meeting: April 8, 2014 from 10am-11am in Auraria Library Friends Conference Room.**

**Minutes for the Minority Affairs Meeting Agenda**  
**Tuesday April 8, 2014**  
**10:00 a.m. – 11:00 a.m.**  
**Auraria Library Friends Conference Room**

**Attendees:** Farah Ibrahim (SEHD); Philip Joseph (English); Denise Pan (Library); Stephanie Santorico (CLAS); Craig Sisneros (Business)

**Agenda:**

1. Review minutes from March meeting (attached)

Approved.

2. "Reconsidering FCQs" Recap event
  - a. Videotaping session
  - b. Review attendee list – 19 registered attendees (includes MAC members)
  - c. Feedback
  - d. Review draft of Summary Report from Craig

Craig is working on a summary which we will discuss at the May meeting. Perhaps we could consider turning this into a best practices document. Other ideas: provide information to evaluation committees so that they are aware of the biases that do exist in FCQ responses as well as the implicit biases in students, departments and faculty. Provide training for faculty on how to conduct evaluations. Often training is focused on what you shouldn't do rather than on what you should do. Also, teach faculty and students how to provide useful critiques.

3. MAC Executive Committee Elections
  - a. Chair or two Co-Chairs
  - b. Secretary

Stephanie Santorico is on sabbatical next year but will suggest a substitute.

Denise Pan will not be the library representative; though will still serve on the committee. The new representative for the library will be Vera Gao.

Denise Pan and Craig Sisneros will continue on as co-chairs.

Charles Musiba and Philip Joseph will not be continuing on the committee.

We'll leave the secretary position open for election in the Fall 2014.

4. Begin planning for next year

Tabled for next meeting.

5. Other updates or announcements?

a. EMAC (Philip) –

Mark Knowles, chair of EMAC, has sent an email concerning the article below to Brenda Allen and other chief diversity officers in the university system. EPCC is trying to move forward on a discussion that proposes to shift some scholarships away from merit based to needs based.

This pertains to the New York Times Opinion Editorial:

[http://opinionator.blogs.nytimes.com/2014/02/05/improving-economic-diversity-at-the-better-colleges/?\\_php=true&\\_type=blogs&\\_r=0](http://opinionator.blogs.nytimes.com/2014/02/05/improving-economic-diversity-at-the-better-colleges/?_php=true&_type=blogs&_r=0). Here, Franklin & Marshall College has shifted money away from merit based scholarships to needs based scholarship, demonstrating that this is working to both increase the depth of student learning and diversity.

b. Higher Education Diversity Summit (Denise)

Theme: Civil Rights @ 50: Advancing Equal Access for All

Keynote: Lily Ledbetter

Thursday April 10 <http://aurariaevents.wix.com/heds>

Registration is around 400 people. There is a reception at 3:30 which would be a nice option to attend.

**Next Meeting: Tuesday May 13, 2014 from 10am-11am in Auraria Library Friends Conference Room.**

**Minority Affairs Meeting Minutes**  
**Tuesday May 13, 2014 at 10:00 – 11:00 a.m.**  
**Auraria Library Friends Conference Room**

**Attendees:** Farah Ibrahim, Philip Joseph, Denise Pan, Craig Sisneros, and Rebecca Hunt -- representing Educational Policy University System (EPUS) and Faculty Assembly Education Planning and Policy Committee (EPPC).

**Minutes:**

1. Review minutes from April meeting
  - a. No changes minutes approved.
2. Review draft of Reconsidering FCQs Summary Report from Craig
  - a. Denise updated committee that the final cost for videotaping and editing the faculty forum by CU Media Productions was \$500.
  - b. Craig shared his draft, but still needs to write a summary and conclusion by Friday May 16.
  - c. He will send drafts to MAC members and Rebecca.
3. Begin planning for next year
  - a. Review past minutes/discussion ideas, see January 2014 minutes c. i. Event goals and target audience.
  - b. Next steps ideas.
    - i. The Faculty Forum identified the issue that junior instructors are intimidated and do not experiment with teaching methodology.
    - ii. Use the Reconsidering FCQs Summary Report as a starting place to develop a document to endorse EPUS multiple means of evaluation.
    - iii. Assign MAC members to talk their Associate Deans about the school/college's approach and what is happening at the primary unit evaluation. We will need to establish a set of questions to ask.
    - iv. Develop a recommendation from MAC and share with EPPC and Faculty Assembly to endorse.
    - v. Then take the recommendation to Deans/Chairs of Schools and Colleges.
4. Other updates or announcements?
  - a. EMAC – No volunteers. Denise will send Joanne an email.
  - b. Excellence in Equity – The committee will be wrapping up in June. Craig will have a report during fall semester.
  - c. Higher Education Diversity Summit (HEDS) – The conference was a success. The HEDS committee has begun to plan the 2015 event. MSU Denver and Auraria Library are exploring the possibility of building a website for the event, and expanding tri-institutional collaboration beyond the event to an Inclusive Excellence Site. Denise is scheduling meetings with the Chief Diversity Officers from CU Denver and CCD to discuss further.
  - d. We will not meet in June and July.
  - e. In early August, Denise will send out an email to welcome new members and a Doodle poll to schedule standing meetings.



**Minority Affairs Meeting Minutes**  
**Friday September 19, 2014**  
**10:00 a.m. – 11:00 a.m.**  
**Auraria Library Friends Conference Room**

**Attendees:** Denise Pan (Auraria Library); Vera Gao (Auraria Library); Ji Chen (Business); Faye C. Caronan (CLAS); Omar Swartz (CLAS); Carlos Reali (CLAS).

**Minutes**

1. Welcome and introductions
2. Election of MAC Secretary
  - a. No volunteers. Tabled for next meeting. Potentially will need to draw lots to select Secretary.
3. Reconsidering FCQs Faculty Forum
  - a. Review event report written by Craig.
  - b. Suggestion to create an executive summary version.
4. Begin planning for 2014-15
  - a. Members agreed that they want to continue discussing FCQs and potential impact on underrepresented faculty. Some concerns discussed included:
    - i. How to make the evaluation process meaningful for everyone
    - ii. How to analyze and interpret FCQ comments/scores
    - iii. Multiple means of evaluation for annual/tenure reviews
    - iv. Issues with completion rates with online FCQs
  - b. Possible ideas discussed included:
    - i. Developing presentation on the issues and suggestions on how to address concerns for Department Chairs and/or teaching faculty.
    - ii. Possibly including real examples or fictions scenarios.
    - iii. Partner with Center for Faculty Development
    - iv. Potential venues could be CLAS Chairs meeting, Academic Support Group and Associate Deans (AD/ASG), Faculty Assembly (FA) meetings and Executive Committee.
5. Other updates or announcements
  - a. Faculty Council EMAC – Omar
    - i. Committee interested in reviewing results and developing findings from the CU System Climate Survey
  - b. Diversity and Inclusion Initiatives Fund by CLAS – Omar
    - i. Accepting proposals to enhance diversity across the college and foster a culture of inclusion
  - c. Interdisciplinary programming grant opportunities from CLAS - Omar
  - d. CU System Diversity Summit October 24, 2014 – Denise  
<https://www.cu.edu/office-academic-affairs/2014DiversitySummit>
  - e. Higher Education Diversity Summit April 10, 2015 <http://heds.auraria.edu> - Denise
6. Action items
  - a. Denise - forward copy of revised FCQ to MAC list
  - b. Denise – develop draft of Exec Summary version of report
  - c. Chen, Faye, Carlos, and other teaching faculty – develop more ideas on how to develop presentation on issues and solutions/strategies

**Next meeting: Friday Oct 3, 2014 at 10 am in the Auraria Library Friends Conference Room**

October 2014

Agenda

1. Review of minutes from May and September 2014 (attached)
  - A. Approved
2. Nominations and election of Secretary
  - A. Faye- nominated and approved
3. Meeting days/times
  - A. 2nd Wednesday of month? No.
  - B. Proposing 2nd Friday. 11/14 and 12/12
3. Review Executive Summary and Report for Reconsidering FCQs (attached)
  - A. Denise recommended some changes to the report and presented executive summary.
  - B. Approved
4. Continue planning discussion
  - A. Developing a best practices workshop on how fcq data is used and how to teach students to provide useful recommendations.
  - B. General question of how to do evaluations
  - C. Working Workshop name- Multiple Means of Evaluation? : Identifying Best Practices (in classrooms, for administrators, for rtp evaluation)
  - D. Apply for Diversity and Excellence Grant- 10/31 deadline: Faye and Denise
  - E. Present at Laura Goodwin's FA mtg / FA Exec mtg / Chair mtg
5. Other updates or announcement
  - A. Read through emac minutes
  - B. HEDS April 10, 2015: Who's Sitting Next to You?
  - C. Diversity event calendar site
  - D. Getting in touch with student organizations

December 2014  
Agenda

Present: Vera Gao, Carlos, Denise, Faye, Chen Ji, Farrah Ibrahim

1. Review minutes from October 2014
  - A. Approved
2. Recap of FCQ meeting from Faye
  - a. Want to start new system, independent from Boulder. Start on budget to see if feasible. If it is feasible, then questions remain the same for a couple of years for continuity before changing questions to tailor for schools/depts.
  - b. Farrah- said that in different meeting, there is no money to migrate to new system independent from Boulder. Advocates that individuals should add own questions.
3. Planning workshop: Multiple Means of Evaluation? : Identifying Best Practices (in classrooms, for administrators, during RTP evaluation)
  - a. Review draft of proposal for the President's Diversity Fund
    1. Based on findings from last forum, workshop model for administrators, student participation, and faculty RTP.
    2. Farrah suggested skill-soft training for faculty who evaluate peers in RTP- hold off for a future recommendation to add to or workshop findings.
    3. Add recording into budget- no. Better to protect anonymity.
    4. Draft approved.
  - b. Begin discussion of planning logistics – proposed dates and other details.
    1. April 2<sup>nd</sup> 11-12:15pm or 12:30-1:45pm or 3<sup>rd</sup>? Noon-1pm
    2. Draft email with doodle poll to find which option would be the best day and time. Farrah write email draft and Denise will set up doodle poll for committee to send out to contacts. Draft will be sent out for review Monday, 1/12. Feedback due by Wednesday 1/14 and then final draft to be sent out first week of classes.
4. Meeting schedule for Spring term: Jan 23 10-11am, Feb. 24 10-11am, Mar 10 10-11am, Mar 31 10-11am
5. Denise will set up a dropbox folder for group for collaboration and minutes.
6. Continue discussion from October meeting: promoting collaboration between faculty and students – Could **MAC** help faculty engage with students by serving as a communication conduit between students groups and interested faculty? If so, how?
  - a. First step to reach out to student groups

# Reconsidering FCQs

## LUNCH AND LEARN FOR CU DENVER FACULTY

MARCH 14 2014 | 11:30am - 1:00pm | AURARIA LIBRARY ENHANCED LEARNING CENTER

Event funded by the Office of Diversity and Inclusion

Please RSVP: <http://tinyurl.com/ReconsideringFCQs>

Join the **CU Denver Minority Affairs Committee (MAC)** for lunch to learn about the quantitative data from **CU Denver Office of Institutional Research & Effectiveness (OIRE)** and to provide insights and direction on finding alternative methods of evaluating teaching. **Reconsidering FCQs Lunch and Learn for CU Denver Faculty** is a forum to discuss the role of FCQs in promotion and tenure, and their limitations as a measurement of effective teaching.

### DISCUSSION...

With excerpts from the literature and OIRE research as a starting point, we hope to initiate a discussion on the potential impact of FCQs on the retention of CU Denver faculty from underrepresented racial-ethnic faculty groups.

#### **Potential Inquiries For Discussion:**

- Do assumptions about gender, race, sexual orientation, national origin and language shape FCQ student responses?
- What is the potential impact of FCQs on the underrepresented racial-ethnic faculty of CU Denver?
- Are FCQs an accurate measurement of effective teaching?
- What are the alternative methods of evaluating teaching?

### RESEARCH...

MAC is charged with evaluating and recommending policies affecting minority faculty with particular regards to hiring, support, and retention.

Our first step was to ask the OIRE to explore FCQ data to discern whether or not minority faculty on our campus are subjected to the biases noted in national research. OIRE used the seven most recent fiscal years' of FCQ data collected for the Denver Campus and included only those course sections with at least 10 enrolled students and at least 10 FCQ responses. This resulted in resulting in 18,205 distinct course section observations (cross-listed courses counted as one observation).

Controlling for factors (through regression analyses) that could influence FCQ scores (e.g., course level, instructional method, course enrollment, term, student age, student cumulative GPA, course GPA, instructor rank, and instructor gender), findings indicated that the average overall course and faculty FCQ scores for minority faculty were lower, on average, by approximately 0.15 to 0.25 points (on a six-point FCQ rating scale).

MAC believes, depending on the Primary Unit Criteria, this could have an effect on periodic evaluation and/or tenure and promotion at the margin.

### LITERATURE REVIEW...

#### **Student Ratings**

A review of the recent (2005 – present) literature on student ratings of faculty members shows that there are significant problems in using end-of-course questionnaires to evaluate the teaching performance of faculty members, and particularly minority faculty members. Students come in with preconceptions of professor warmth, capability, and political bias, and these are reflected in faculty course questionnaires (*Anderson et al., 2005*).

#### **Student Evaluation Studies**

A recent study (Reid, 2010) of student evaluations of faculty at 25 highly ranked liberal arts colleges found that students evaluated racial minority faculty more negatively than white faculty. Similarly, end-of-course questionnaires of students in a College of Education at a Research 1 institution were analyzed (*B. Smith, 2007*); student ratings for white faculty members were "very good" while the ratings of black faculty were "good".

One study (*G. Smith et al., 2005*) examined the impact of gender, ethnicity (Latino/a or Anglo), and teaching style (lenient or strict) on student perceptions of social science instructors; among other findings this study concluded that Latino/a faculty members with a strict teaching style received the least favorable marks of all faculty groups. While there are not very many recent empirical studies pertaining to end-of-course questionnaires and minority faculty members, the studies confirm that there are definite and significant bias issues with this type of student evaluation.

#### **To Read More:**

<http://library.auraria.edu/content/bias-student-evaluations-minority-faculty>



Office of Diversity and Inclusion

UNIVERSITY OF COLORADO  
DENVER | ANSCHUTZ MEDICAL CAMPUS

# What's in a Name: Exposing Gender Bias in Student Ratings of Teaching

Lillian MacNell · Adam Driscoll · Andrea N. Hunt

© Springer Science+Business Media New York 2014

**Abstract** Student ratings of teaching play a significant role in career outcomes for higher education instructors. Although instructor gender has been shown to play an important role in influencing student ratings, the extent and nature of that role remains contested. While difficult to separate gender from teaching practices in person, it is possible to disguise an instructor's gender identity online. In our experiment, assistant instructors in an online class each operated under two different gender identities. Students rated the male identity significantly higher than the female identity, regardless of the instructor's actual gender, demonstrating gender bias. Given the vital role that student ratings play in academic career trajectories, this finding warrants considerable attention.

**Keywords** gender inequality · gender bias · student ratings of teaching · student evaluations of instruction

---

**Lillian MacNell** is a doctoral candidate in Sociology at North Carolina State University. She received her Master's degree in Sociology at the University of Central Florida. Her research and teaching interests include food access, food justice, and the environment.

**Adam Driscoll** is Assistant Professor of Sociology at the University of Wisconsin-La Crosse. He received his Master's degree in Sociology at East Carolina University and his Ph.D. in Sociology at North Carolina State University. His research and teaching focus upon the environmental impacts of industrial agriculture and effective online pedagogy.

**Andrea N. Hunt** has a Ph.D. in Sociology from North Carolina State University and is currently Assistant Professor in Sociology and Family Studies at the University of North Alabama. Her research interests include gender, race and ethnicity, mentoring in undergraduate research, engaging teaching practices, and the role of academic advising in student retention.

L. MacNell (✉)

Department of Sociology and Anthropology, 334 1911 Building, Campus Box 8107,  
Raleigh, North Carolina 27695, USA  
e-mail: loconne@ncsu.edu

A. Driscoll

University of Wisconsin-La Crosse, La Crosse, WI, USA  
e-mail: adriscoll@uw.lax.edu

A. N. Hunt

University of North Alabama, Florence, AL, USA  
e-mail: ahunt3@una.edu

Student ratings of teaching are often used as an indicator of the quality of an instructor's teaching and play an important role in tenure and promotion decisions (Abrami, d'Apollonia, & Rosenfield, 2007; Benton & Cashin, 2014). Gender bias in these ratings constitutes an important form of inequality facing women in academia that is often unaccounted for in such decisions. Students perceive, evaluate, and treat female instructors quite differently than they do male instructors (Basow, 1995; Centra & Gaubatz, 2000; Feldman, 1992; Young, Rush, & Shaw, 2009). While a general consensus exists that gender plays a vital role in how students perceive and interact with their instructors, there is conflicting evidence as to whether or not this translates into a bias in student ratings due to variations in several mediating factors such as teaching styles and subject material.

Prior studies of student ratings of instruction have been limited in their ability to test for the existence of gender bias because it is difficult to separate the gender of an instructor from their teaching practices in a face-to-face classroom. In online courses, however, students usually base the categorization of their instructor's gender on the instructor's name and, if provided, photograph. It is possible for students to believe that their instructor is actually a man, based solely on a name or photograph, when in reality she is a woman, or vice versa. Therefore, the online environment affords researchers a unique opportunity to assign one instructor two different gender identities in order to understand whether or not differences in student ratings are a result of differences in teaching or simply based on unequal student expectations for male and female instructors. Such experimentation allows researchers to control for potentially confounding factors and therefore attribute observed differences solely to the variable of interest—in this case, the perceived gender of the instructor (Morgan & Winship, 2007).

This study analyzed differences in student ratings of their instructors<sup>1</sup> from an online course, independent of actual gender. The course professor randomly assigned students to one of six discussion groups, two of which the professor taught directly. The other four were taught by one of two assistant instructors—one male and one female. Each instructor was responsible for grading the work of students in their group and interacting with those students on course discussion boards. Each assistant instructor taught one of their groups under their own identity and the second group under the other assistant instructor's identity. Thus, of the two groups who believed they had the female assistant instructor, one actually had the male. Similarly, of the two groups who believed they had the male assistant instructor, one actually had the female (see Table 1). At the end of the course, the professor asked students to rate their instructor through the use of an online survey. This design created a controlled experiment that allowed us to isolate the effects of the gender identity of the assistant instructors, independent of their actual gender. If gender bias was present, then the students from the two groups who believed they had a female assistant instructor should have given their instructor significantly lower evaluations than the two groups who believed they had a male assistant instructor.

## Student Ratings of Teaching

Though far from perfect, student ratings of teaching provide valuable feedback about an instructor's teaching effectiveness (Svinicki & McKeachie, 2010). They may be reliably interpreted as both a direct measure of student satisfaction with instruction and as an indirect

<sup>1</sup> To clarify the language we use throughout the paper, we refer to all three persons responsible for grading and directly interacting with students as "instructors." The course "professor" was the person responsible for course design and content preparation, while the two "assistant instructors" worked under the professor's direction to manage and teach their respective discussion groups.

**Table 1** Experimental Design.

Discussion Group	Instructor's Perceived Gender	Instructor's Actual Gender
Group A ( <i>n</i> =8)	Female	Female
Group B ( <i>n</i> =12)	Female	Male
Group C ( <i>n</i> =12)	Male	Female
Group D ( <i>n</i> =11)	Male	Male

measure of student learning (Marsh, 2007; Murray, 2007). They also play an important role in the selection of teaching award winners, institutional reviews of programs, and student course selection (Benton & Cashin, 2014). More importantly to the careers of educators, these ratings are “used by faculty committees and administrators to make decisions about merit increases, promotion, and tenure” (Davis, 2009, p. 534). In particular, quantitative evaluations of instructors’ overall teaching effectiveness are frequently emphasized in personnel decisions (Centra & Gaubatz, 2000). Given the widespread reliance on student ratings of teaching and their effect on career advancement, any potential bias in those ratings is a matter of great consequence.

### Gender Bias in Academia

Sociological studies of gender and gender inequality are careful to distinguish between sex (a biological identity) and gender (a socially constructed category built around cultural expectations of male- and female-appropriate behavior). Gender is part of an ongoing performance based on producing a configuration of behaviors that are seen by others as normative. West and Zimmerman (1987) suggested that people engage in gendered behaviors not only to live up to normative standards, but also to minimize the risk of accountability or gender assessment from others. Thus, gender is a process that is accomplished at the interactional level and reinforced through the organization of social institutions such as academia (Lorber, 1994). Gender then contributes to a hierarchal system of power relations that is embedded within the interactional and institutional levels of society and shapes gendered expectations and experiences in the workplace (Risman, 2004).

An examination of gender bias in student ratings of teaching must be framed within the broader context of the pervasive devaluation of women, relative to men, that occurs in professional settings in the United States (Monroe, Ozyurt, Wrigley, & Alexander, 2008). In general, Western culture accords men an automatic credibility or competence that it does not extend to women (Johnson, 2006). Stereotypes that women are less logical, less confident, and occupy lower positions still pervade our organizational structures (Acker, 1990). Conversely, men are automatically assumed to have legitimate authority, while women must prove their expertise to earn the same level of respect. This disparity has been well documented in the field of academia, where men tend to be regarded as “professors” and women as “teachers” (Miller & Chamberlin, 2000) and women face a disparate amount of gender-based obstacles, relative to men (Morris, 2011).

In experiments where researchers gave students identical articles to evaluate—half of which bore a man’s name and half of which bore a woman’s—the students rated the research they thought had been done by men more highly (Goldberg, 1968; Paludi & Strayer, 1985). In a similar study, college students evaluated two hypothetical applicants for a faculty position and tended to judge the male candidate as more qualified despite the fact that both applicants had identical credentials (Burns-Glover & Veith, 1995). Additionally, a study of student



evaluations of instructors' educational attainment revealed that students misattribute male instructors' education upward and female instructors' education downward (Miller & Chamberlin, 2000). Overall, women in academia tend to be regarded as less capable and less accomplished than men, regardless of their actual achievements and abilities.

### Gender Role Expectations

Students often expect their male and female professors to behave in different ways or to respectively exhibit certain "masculine" and "feminine" traits. Commonly held masculine, or "effectiveness," traits include professionalism and objectivity; feminine, or "interpersonal," traits include warmth and accessibility. Students hold their instructors accountable to these gendered behaviors and are critical of instructors who violate these expectations (Bachen, McLoughlin, & Garcia, 1999; Chamberlin & Hickey, 2001; Dalmia, Giedeman, Klein, & Levenburg, 2005; Sprague & Massoni, 2005). Consequently, instructors who adhere to gendered expectations are viewed more favorably by their students (Andersen & Miller, 1997; Bennet, 1982). When female instructors exhibit strong interpersonal traits, they are viewed comparably to their male counterparts. When female instructors fail to meet these gendered expectations, however, they are sanctioned, while male instructors who do not exhibit strong interpersonal traits are not (Basow & Montgomery, 2005; Basow, Phelan, & Capotosto, 2006). At the same time, students are less tolerant of female instructors whom they perceive as lacking professionalism and objectivity than they are of male instructors who lack the same qualities (Bennet, 1982). In general, "students' perceptions and evaluations of female faculty are tied more closely to their gender expectations than for male faculty" (Bachen et al., 1999, p. 196).

These different standards can place female instructors in a difficult "double-bind," where gendered expectations (that women be nurturing and supportive) conflict with the professional expectations of a higher-education instructor (that they be authoritative and knowledgeable) (Sandler, 1991; Statham, Richardson, & Cook, 1991). On the one hand, students expect female instructors to embody gendered interpersonal traits by being more accessible and personable. However, these same traits can cause students to view female instructors as less competent or effective. On the other hand, female instructors who are authoritative and knowledgeable are violating students' gendered expectations, which can also result in student disapproval. Therefore, female instructors are expected to be more open and accessible to students *as well as* to maintain a high degree of professionalism and objectivity. Female instructors who fail to meet these higher expectations are viewed as less effective teachers than men (Basow, 1995).

Male instructors, however, are rated more highly when they exhibit interpersonal characteristics in addition to the expected effectiveness characteristics (Andersen & Miller, 1997). In other words, female instructors who fail to exhibit an ideal mix of traits are rated lower for not meeting expectations, while male instructors are not held to such a standard. Consequently, gendered expectations represent a greater burden for female than male instructors (Sandler, 1991; Sprague & Massoni, 2005). An important manifestation of that disparity is bias in student ratings of instructors, where female instructors may receive lower ratings than males, not because of differences in teaching but for failing to meet gendered expectations.

### Methodological Concerns with Previous Studies of Gender Bias

Studies of gender bias in student ratings of instruction have presented complicated and sometimes contradictory results. Sometimes men received significantly higher ratings (Basow & Silberg, 1987; Sidanius & Crane, 1989), sometimes women (Bachen et al., 1999; Rowden & Carlson, 1996), and sometimes neither (Centra & Gaubatz, 2000; Feldman, 1993). The



variety of results in these studies suggests that gender does play a role in students' ratings of their instructors, but that it is a complex and multifaceted one (Basow et al., 2006).

One reason why prior research on gender bias in student ratings of teaching has provided such inconclusive results may lie in the research design of these previous studies. A large portion of research on student ratings of teaching directly utilized those ratings for their data (e.g. Basow, 1995; Bennett, 1982; Centra, 2007; Centra & Gaubatz, 2000; Marsh, 2001). This strategy allows for the analysis of a large amount of data, but it does not control for differences in actual teaching and therefore may fail to capture gender bias in student ratings. Studies that compare student ratings of instructors explore whether or not there are differences—not whether or not those differences are the result of gender bias (Feldman, 1993). For example, a study of ratings may find that a female instructor received significantly lower scores than a male peer, but it could not assess whether that indicates a true difference in teaching quality. Perhaps she was not perceived as warm and engaging; failing to meet the gendered expectations of the students, she may have been rated more poorly than her male peer despite being an equally effective instructor. Similarly, the lack of a gender disparity in student ratings of instruction could actually obscure a gender bias if at a particular institution the female faculty members were, on average, stronger instructors than the males, yet were being penalized by the students due to bias (Feldman, 1993).

Additionally, a number of situational elements may serve to sway student ratings of male versus female instructors as male and female professors tend to occupy somewhat different teaching situations. Men are overrepresented in the higher ranks of academic positions as well as in STEM fields. They are also more likely to teach upper-level courses whereas women are more likely to teach introductory courses (Simeone, 1987; Statham et al., 1991). Women are also more likely than men to be employed in full-time non-tenure track positions as well as in part-time positions (Curtis, 2011). These factors are highly relevant because instructor rank, academic area, and class level of the course have all been found to directly impact student ratings of instruction (Feldman, 1993; Liu, 2012). All of these factors serve to complicate the relationship between instructor gender and student ratings of instruction and obfuscate the conclusions that can be drawn from direct studies of such ratings. Studies of actual student ratings of instruction may tell us more about women's position in academia than about actual gender bias in student ratings. In contrast, experimental studies allow the researcher to control for both the quality and character of the teaching as well as the academic position of the instructor; ensuring that any differences registered in student ratings indicate, as much as possible, a bias rather than an actual difference in teaching (Feldman, 1993).

## Research Question and Related Hypotheses

The fundamental question examined in this study is whether or not students rate their instructors differently on the basis of what they perceive those instructors' gender to be. We expected that there would be no difference between the ratings for the actual male and female instructors in the course as every attempt was made to minimize any differences in interaction and teaching. However, we expected that student ratings of instructors would reflect the different expectations for male and female instructors discussed above. Instructors whom students perceived to be male would be afforded an automatic credibility on their competence and professionalism. Furthermore, they would not be penalized for any perceived deficiency in their interpersonal skills. Therefore, we expected that students would rate the instructors they *believed* to be male more highly than ones they believed to be female, regardless of the instructors' actual gender.

## The Study and Methodology

This study examined gender bias in student ratings of teaching by falsifying the gender of assistant instructors in an online course and asking students to evaluate them along a number of instructional criteria. By using a 2-by-2 experimental design (see Table 1), we were able to compare student evaluations of a perceived gender while holding the instructor's actual gender (and any associated differences in teaching style) constant. Any observed differences in how students rated one perceived gender versus the other must have therefore derived from bias on the students' part, given that the exact same two instructors (one of each gender) were being evaluated in both cases.

### Subjects

Data were collected from an online introductory-level anthropology/sociology course offered during a five-week summer session at a large (20,000+), public, 4-year university in North Carolina. The University's institutional review board had approved this study (IRB# 2640). The course fulfilled one of the university's general education requirements, and the students represented a range of majors and grade levels. The majority of the participants were traditional college-aged students with a median age of 21 years. The instructors taught the course entirely through a learning management system and students' only contact with their instructors was either through e-mail or comments posted on the learning management system. The professor delivered course content through assigned readings and written PowerPoint slideshow lectures. The course was broken up into nine different content sections. For each section, students were required to read the assigned material and make a series of posts on a structured discussion board. The course had 72 students who were randomly divided into six discussion groups for the entirety of the course. All discussion board activity took place within the assigned discussion group. Each discussion group had one instructor responsible for moderating the discussion boards and grading all assignments for that group. The course professor took two groups and divided the remaining four between the two assistant instructors, each taking one group under their own identity and a second under their fellow assistant instructor's identity (see Table 1). All instructors were aware of the study being conducted and cooperated fully.

The section discussion boards were the primary source of interaction between students and the course instructors and, as such, represented 30% of the students' final grades. The discussion boards were also an important part of student learning because they were the main arena in which students could analyze and voice questions about course concepts and material. The instructor assigned to each discussion group maintained an active presence on each discussion board, offering comments and posing questions. The instructor also graded students' posts and provided detailed feedback on where students had lost points. The two assistant instructors for the four discussion groups employed a wide range of strategies so as to maintain consistency in teaching style and grading. The two assistant instructors composed personal introduction posts that indicated similar biographical information and background credentials. They posted on the discussion boards and graded assignments at the same time of day three days each week to ensure that no group received significantly faster or slower feedback than others. The professor provided detailed grading rubrics for the discussion boards, and the instructors coordinated their grading to ensure that these rubrics were applied to students' work equitably.<sup>2</sup>

<sup>2</sup> A one-way ANOVA test confirmed that there was no significant variation among all six groups' discussion board grades and overall grades for the course.

Toward the end of the course the professor sent students reminder e-mails requesting that they complete an online evaluation of their instructor. These evaluations were explained as serving the purpose of providing the professor with feedback about the instructors' performance. The survey asked students to rate their instructor on various factors such as accessibility, effectiveness, and overall quality. Over 90% of the class completed the evaluation. For the purpose of this study, we only analyzed data from the discussion groups assigned to the assistant instructors, leaving us with 43 subjects.

### Instrument

The instructor evaluation consisted of 15 closed-ended questions that ask students to rate their instructors on a variety of measures using a five-point Likert scale (1 = Strongly disagree, 2 = Disagree, 3 = Neither Agree nor Disagree, 4 = Agree, 5 = Strongly agree). The survey had six questions designed to measure effectiveness traits (e.g. professionalism, knowledge, and objectivity) and six questions designed to measure interpersonal traits (e.g. respect, enthusiasm, and warmth). In addition, there were two questions designed to measure communication skills and one question that asked students to evaluate the instructor's overall quality as a teacher. We also asked students to indicate which discussion group they were in and to provide basic demographic and academic background information including gender, age, year in school, and number of credit hours currently being taken. All students fully completed the evaluation, leaving us with no missing data.

We performed all analyses with the 13<sup>th</sup> version of the Stata statistical analysis program. We used exploratory factor analysis to test how well the separate questions reflected a common underlying dimension. Principal component factor analysis revealed that 12 of our items characterized a single factor for which the individual factor loadings ranged from .7370 to .9489; sufficiently high to justify merging them into a single index (Hair, Anderson, Tatham, & Black, 1998). This indicates that those 12 questions on our survey were all measuring the same latent variable, which we interpret to be a general evaluation of the instructor's teaching. A reliability test yielded a Cronbach's alpha above .950 for the 12 questions. In order to confirm the factor structure, we used structural equation modeling to test a single latent variable indicated by our 12 separate questions. Our model was a strong fit to the data ( $N=43$ ,  $\chi^2(47)=59.18$  (not significant), RMSEA =0.078, CFI =0.980, SRMR =0.043) with all loadings significant at the  $p < 0.001$  level. Therefore, we extracted a factor score, *student ratings index*, which weighed each question by how strongly it loaded onto the single factor, providing us with a single representation of how well each student evaluated their instructor's teaching.

### Analysis

To test for the existence of gender bias in student ratings of teaching, we made two types of comparisons. First we compared across the *actual* gender of the assistant instructor, combining the two groups that had the female assistant instructor (one of which thought they had a male) into one category and doing the same with the two groups that had the male assistant instructor. Second, we compared across the *perceived* gender of the assistant instructor, combining the two groups that thought they had a female assistant instructor (one of which was actually a male) into one category and doing the same with the two groups that thought they had a male assistant instructor. We made both comparisons for the 12 individual questions, as well as the *student ratings index*. We used Welch's *t*-tests (an adaptation of the Student's *t*-test that does not assume equal variance) to establish the statistical significance of each difference. We also ran two general linear multivariate analyses of variance (MANOVAs) on the set of 12 variables

to test the effects of instructor gender (perceived and actual) on all of the questions considered as a group. A MANOVA allows a researcher to test a set of correlated dependent variables and conduct a single, overall comparison between the groups formed by categorical independent variables (Garson, 2012). This *F*-test of all means addresses the potential for false positive findings as the result of multiple comparisons.<sup>3</sup>

## Results

### Student Ratings of Perceived and Actual Gender

By comparing differences across the *actual* gender of the assistant instructor with those observed across the *perceived* gender of the instructor it is possible to observe whether or not students rated their instructors differently depending on the gender of the instructor. The results of this comparison are found in Table 2.

Our MANOVAs indicate that there is a significant difference in how students rated the perceived male and female instructors ( $p < 0.05$ ), but not the actual male and female instructors. When looking at the individual questions as well as the *student ratings index*, there are no significant differences between the ratings of the actual male and female instructor (the first and second columns in Table 2). Students in the two groups that had the female assistant instructor (one of which thought they had a male) did not rate their instructor any differently than did the students in the two groups that had the male assistant instructor. The left two columns of Fig. 1 provide a graphic representation of this comparison for the *student ratings index*. The overlapping error bars ( $\pm$  one standard error) indicate the lack of a significant difference between how students rated the actual male and female assistant instructors.

When comparing between the perceived gender identities of the instructors (the fourth and fifth columns in Table 2), we found that the male identity received significantly higher scores on professionalism, promptness, fairness, respectfulness, enthusiasm, giving praise, and the *student ratings index*.<sup>4</sup> Looking at the *R*-squares, all seven of these comparisons yielded a medium sized effect. It is worth noting, particularly given the small sample size, that the male instructor identity also received higher scores on the other six questions, though not to a statistically significant degree. Students in the two groups that perceived their assistant instructor to be male rated their instructor significantly higher than did the students in the two groups that perceived their assistant instructor to be female, regardless of the actual gender of the assistant instructor. This comparison is represented graphically by the right two columns of Fig. 1, where a clear difference can be observed.

<sup>3</sup> We acknowledge that the application of parametric analytical techniques (ANOVA, MANOVA, and *t*-tests) to ordinal data (the Likert scale responses) remains controversial among social scientists and statisticians. (See Knapp (1990) for a relatively balanced review of the debate.) We side with the arguments of Gaito (1980) and Armstrong (1981) and argue that it is appropriate to do so in our case as the concept being measured is interval, even if the data labels are not. This practice is common within higher education research. (e.g. Centra & Gaubatz [2000] Young, Rush, & Shaw [2009]; Basow [1995]; and Knol et al. [2013])

<sup>4</sup> While we acknowledge that a significance level of .05 is conventional in social science and higher education research, we side with Skipper, Guenther, and Nass (1967), Labovitz (1968), and Lai (1973) in pointing out the arbitrary nature of conventional significance levels. Considering our study design, we have used a significance level of .10 for some tests where: 1) the results support the hypothesis and we are consequently more willing to reject the null hypothesis of no difference; 2) our hypothesis is strongly supported theoretically and by empirical results in other studies that use lower significance levels; 3) our small *n* may be obscuring large differences; and 4) the gravity of an increased risk of Type I error is diminished in light of the benefit of decreasing the risk of a Type II error (Labovitz, 1968; Lai, 1973).

**Table 2** Comparison of means of student ratings of teaching across the actual gender of the assistant instructor and the perceived gender of the assistant instructor

Question	Actual Female	Actual Male	Difference	Perceived Female	Perceived Male	Difference
Caring	4.00 (1.257)	3.87 (0.868)	0.13 (0.004)	3.65 (1.226)	4.17 (0.834)	-0.52 (0.071)
Consistent	3.80 (1.322)	3.70 (1.020)	0.10 (0.002)	3.50 (1.357)	3.96 (0.928)	-0.47 (0.045)
Enthusiastic	4.05 (1.191)	3.78 (0.850)	0.27 (0.019)	3.60 (1.314)	4.17 (0.576)	-0.57† (0.112)
Fair	4.05 (1.050)	3.78 (0.951)	0.27 (0.018)	3.50 (1.192)	4.26 (0.619)	-0.76* (0.188)
Feedback	4.10 (1.252)	3.83 (1.029)	0.27 (0.015)	3.70 (1.380)	4.17 (0.834)	-0.47 (0.054)
Helpful	3.65 (1.309)	3.83 (0.834)	-0.18 (0.008)	3.50 (1.192)	3.96 (0.928)	-0.46 (0.049)
Knowledgeable	4.20 (1.056)	4.09 (0.949)	0.11 (0.003)	3.95 (1.191)	4.30 (0.765)	-0.35 (0.038)
Praise	4.35 (0.988)	4.09 (0.900)	0.26 (0.020)	3.85 (1.089)	4.52 (0.665)	-0.67* (0.153)
Professional	4.30 (1.218)	4.35 (0.935)	-0.05 (0.000)	4.00 (1.414)	4.61 (0.499)	-0.61† (0.124)
Prompt	4.10 (1.252)	3.87 (0.919)	0.23 (0.013)	3.55 (1.356)	4.35 (0.573)	-0.80* (0.191)
Respectful	4.30 (1.218)	4.35 (0.935)	-0.05 (0.001)	4.00 (1.414)	4.61 (0.499)	-0.61† (0.124)
Responsive	4.00 (1.124)	3.57 (0.843)	0.43 (0.052)	3.65 (1.137)	3.87 (0.869)	-0.22 (0.013)
Student Rating Index	0.09 (1.165)	-0.08 (0.850)	0.17 (0.008)	-0.33 (1.267)	0.284 (0.584)	-0.61† (0.128)
N	20	23		20	23	

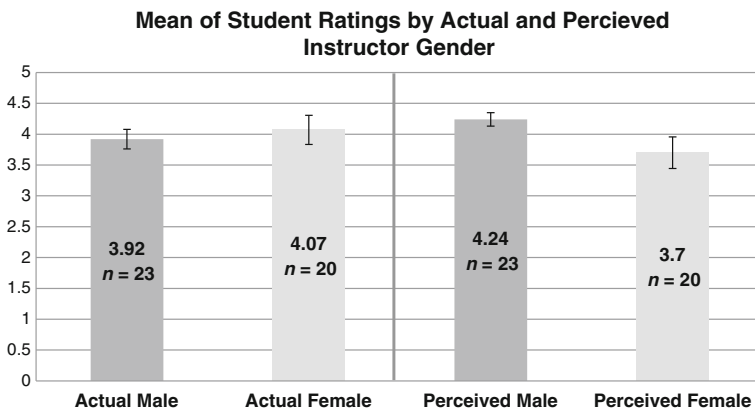
Note: Each cell contains the mean student response for the question with the standard deviations in parentheses. The cells in the Difference columns contain the difference between the means with the *r*-squared in italics and parentheses. Welch's *t*-tests were used to establish the significance of the observed differences.

† *p* < =0.10.

\* *p* < =0.05.

### Discussion

With the design of this experiment, we are able to attribute any differences between how students rated the two perceived genders to gender bias as the students actually evaluated the same two instructors in each case. Our findings support the existence of gender bias in that



**Figure 1** Comparison of the mean of student ratings across actual instructor gender (left two columns) and perceived instructor gender (right two columns). The difference between the right two columns is significant to the *p* < =0.10 level.

students rated the instructors they perceived to be female lower than those they perceived to be male, regardless of teaching quality or actual gender of the instructor. The perceived female instructor received significantly lower ratings on six of the 12 metrics on the survey, as well as on the *student ratings index*.

The difference between how students rated the two perceived genders stands in stark contrast to the fact that neither the actual male nor actual female instructor received significantly higher ratings than the other. Both instructors performed equally well from the students' perspective. However, in both cases the *same* instructor received different ratings depending solely on their perceived gender. In other words, when the actual male instructor was perceived to be female, he received significantly lower ratings than when he was perceived to be a male. For example, when the actual male and female instructors posted grades after two days *as a male*, this was considered by students to be a 4.35 out of 5 level of promptness, but when the same two instructors posted grades at the same time *as a female*, it was considered to be a 3.55 out of 5 level of promptness. In each case, the same instructor, grading under two different identities, was given lower ratings half the time with the only difference being the perceived gender of the instructor. Similarly, students rated the perceived female instructors an average of 0.75 points lower on the question regarding fairness, despite both instructors utilizing the same grading rubrics and there being no significant differences in the average grades of any of the groups. These findings support the argument that male instructors are often afforded an automatic credibility in terms of their professionalism, expertise, and effectiveness as instructors. Despite the fact that the students were equally satisfied with the promptness and fairness of the *actual* instructors, the instructor that students perceived to be male was considered to be more effective.

Similarly, both actual instructors demonstrated the same level of interpersonal interaction in their attempts to create a sense of immediacy in the online classroom. Yet the perceived male instructor received higher ratings on all six interpersonal measures, three of them significantly. We contend that female instructors are *expected* to exhibit such traits and therefore are not rewarded when they do so, while male instructors are perceived as going above and beyond expectations when they exhibit these traits. In other words, students have higher interpersonal standards for their female instructors (Sandler, 1991). Our findings support the existence of this bias. In the online environment, it is more difficult to create immediacy through verbal communication, and nonverbal communication and body language are eliminated entirely (O'Sullivan, Hunt, & Lippert, 2004). Students sanctioned the perceived female instructor for failing to demonstrate strong interpersonal traits, yet did not do the same for the perceived male instructor. Both instructors were working within the same confines of online, text-based communications, but students only penalized the instructor they perceived to be female for this shortcoming.

Although this experiment was conducted in the online environment, we believe that the findings apply more broadly to all student ratings of teaching. Rather than testing for gender bias in the online environment, we used this environment as a natural laboratory to test for the existence of gender bias in student ratings as a whole. We argue that the demonstrated bias exists in the general student population and will manifest itself in both online and face-to-face classrooms. The combination of higher expectations and lower automatic credibility translates into very real differences in student ratings of female versus male instructors. Though it is easier to affect interpersonal characteristics in a face-to-face environment, the fact remains that some

professors are *expected* to do so while others are given a ratings boost for those same behaviors.

Because student ratings of teaching are considered an important measure of teaching proficiency, the existence of gender bias in those scores needs to be better understood and acknowledged within the institutional framework of our higher-education system. These results provide strong evidence that gender bias exists in student ratings of their instructors, but more work is needed. First and foremost, these results need to be replicated in other similar online classes. A single case study cannot establish a broad pattern. However, it does suggest the existence of one and provides incentive for further exploration. Additional studies of this type could lend weight to these findings and better establish the existence of this bias throughout academia. Additionally, courses in other subject areas with a variety of both male and female instructors should follow a similar model to corroborate these findings.

## Conclusions

Our findings show that the bias we saw here is *not* a result of gendered behavior on the part of the instructors, but of actual bias on the part of the students. Regardless of actual gender or performance, students rated the perceived female instructor significantly more harshly than the perceived male instructor, which suggests that a female instructor would have to work harder than a male to receive comparable ratings. If female professors and instructors are continually receiving lower evaluations from their students for no other reason than that they are women, then this particular form of inequality needs to be taken into consideration as women apply for academic jobs and come up for promotion and review.

These findings represent an important contribution to existing debates over the validity of student ratings of teaching. (See Benton & Cashin, 2014; Perry & Smart, 2007; and Theall, Abrami, & Mets, 2001 for reviews.) These debates have highlighted a number of weaknesses and shortcomings of student ratings of teaching as a reflection of the quality of instruction being rated (Greenwald, 1997; Johnson, 2003; Svanum & Aigner, 2011). They have also shown that there is substantial room for updating and improving how student ratings of teaching are collected, interpreted, and utilized (Hampton & Reiser, 2004; Subramanya, 2014). However, for better or worse, they remain one of the primary tools used to evaluate educators' teaching for the purposes of promotion and tenure decisions (Davis, 2009; Svinicki & McKeachie, 2010). This study demonstrates that gender bias is an important deficiency of student ratings of teaching. Therefore, the continued use of student ratings of teaching as a primary means of assessing the quality of an instructor's teaching systematically disadvantages women in academia. As this limitation is one of numerous problems associated with the emphasis on quantitative student ratings of teaching, this work adds to the growing call for re-evaluation and modification of the current system of evaluating the quality of instruction in higher education (Hampton & Reiser, 2004; Morrison & Johnson, 2013).

It is also worth noting that this experiment is only scratching the surface of what is possible with gender studies in the online environment. The online environment presents a unique opportunity to experiment directly with gender identity. Analyzing the difference in online behavior of individuals when they perceive that they are interacting with a male or female could provide a wealth of data on how gender is constructed and treated. We hope that this experiment serves as a model for future work that will enhance our ability to test for gender bias in order to further our understanding of its basis, means of perpetuation, and potential avenues of amelioration.



## References

- Abrami, P. C., d'Apollonia, S., & Rosenfield, S. (2007). The dimensionality of student ratings of instruction: What we know and what we do not. In R. P. Perry & J. C. Smart (Eds.), *The scholarship of teaching and learning in higher education: An evidence-based perspective* (pp. 385–445). Dordrecht, The Netherlands: Springer.
- Acker, J. (1990). Hierarchies, job, and bodies: A theory of gendered organizations. *Gender and Society*, 4, 81–95.
- Andersen, K., & Miller, E. D. (1997). Gender and student evaluations of teaching. *Ps-Political Science and Politics*, 30, 216–219.
- Armstrong, G. D. (1981). Parametric statistics and ordinal data: A pervasive misconception. *Nursing Research*, 30, 60–62.
- Bachen, C. M., McLoughlin, M. M., & Garcia, S. S. (1999). Assessing the role of gender in college students' evaluations of faculty. *Communication Education*, 48, 193–210.
- Basow, S. A. (1995). Student evaluations of college professors: When gender matters. *Journal of Educational Psychology*, 87, 656–665.
- Basow, S. A., & Montgomery, S. (2005). Student ratings and professor self-rating of college teaching: Effects of gender and divisional affiliation. *Journal of Personnel Evaluation in Education*, 18, 91–106.
- Basow, S. A., Phelan, J. E., & Capotosto, L. (2006). Gender patterns in college students' choices of their best and worst professors. *Psychology of Women Quarterly*, 30, 25–35.
- Basow, S. A., & Silberg, N. T. (1987). Student evaluations of college professors: Are female and male professors rated differently? *Journal of Educational Psychology*, 79, 308–314.
- Bennett, S. K. (1982). Student perceptions of and expectations for male and female instructors: Evidence relating to the question of gender bias in teaching evaluation. *Journal of Educational Psychology*, 74, 170–179.
- Benton, S. L., & Cashin, W. E. (2014). Student ratings of instruction in college and university courses. In M. B. Paulsen (Ed.), *Higher education: Handbook of theory and research* (pp. 279–326). Dordrecht, The Netherlands: Springer.
- Burns-Glover, A. L., & Veith, D. J. (1995). Revisiting gender and teaching evaluations: Sex still makes a difference. *Journal of Social Behavior and Personality*, 10, 69–80.
- Centra, J. A. (2007). *Differences in responses to the student instructional report: Is it bias?* Princeton, NJ: Educational Testing Service.
- Centra, J. A., & Gaubatz, N. B. (2000). Is there gender bias in student evaluations of teaching? *Journal of Higher Education*, 71, 17–33.
- Chamberlin, M. S., & Hickey, J. S. (2001). Student evaluations of faculty performance: The role of gender expectations in differential evaluations. *Educational Research Quarterly*, 25, 3–14.
- Curtis, J. W. (2011). *Persistent inequity: Gender and academic employment*. Report from the American Association of University Professors. Retrieved from [http://www.aaup.org/NR/rdonlyres/08E023AB-E6D8-4DBD-99A0-24E5EB73A760/0/persistent\\_inequity.pdf](http://www.aaup.org/NR/rdonlyres/08E023AB-E6D8-4DBD-99A0-24E5EB73A760/0/persistent_inequity.pdf)
- Dalmia, S., Giedeman, D. C., Klein, H. A., & Levenburg, N. M. (2005). Women in academia: An analysis of their expectations, performance and pay. *Forum on Public Policy*, 1, 160–177.
- Davis, B. G. (2009). *Tools for teaching* (2nd ed.). San Francisco, CA: Jossey-Bass.
- Feldman, K. A. (1992). College students' views of male and female college teachers: Evidence from the social laboratory and experiments – Part 1. *Research in Higher Education*, 33, 317–375.
- Feldman, K. A. (1993). College students' views of male and female college teachers: Evidence from the social laboratory and experiments – Part 2. *Research in Higher Education*, 34, 151–211.
- Gaito, J. (1980). Measurement scales and statistics: Resurgence of an old misconception. *Psychological Bulletin*, 87, 564–567.
- Garson, G. D. (2012). *General linear models: Multivariate GLM & MANOVA/MANCOVA*. Asheboro, NC: Statistical Associates.
- Goldberg, P. (1968). Are women prejudiced against women? *Trans-action*, 5, 28–30.
- Greenwald, A. G. (1997). Validity concerns and usefulness of student ratings of instruction. *American Psychologist*, 52, 1182–1186.
- Hair, J. F., Jr., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate data analysis with readings* (5th ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Hampton, S. E., & Reiser, R. A. (2004). Effects of a theory-based feedback and consultation process on instruction and learning in college classrooms. *Research in Higher Education*, 45, 497–527.
- Johnson, V. E. (2003). *Grade inflation: A crisis in college education*. New York, NY: Springer.
- Johnson, A. (2006). *Power, privilege, and difference*. Boston, MA: McGraw-Hill.
- Knapp, T. R. (1990). Treating ordinal scales as interval scales: An attempt to resolve the controversy. *Nursing Research*, 39, 121–123.



- Knol, M. H., Veld, R., Vorst, H. C. M., van Driel, J. H., & Mellenbergh, G. J. (2013). Experimental effects of student evaluations coupled with collaborative consultation on college professors' instructional skills. *Research in Higher Education, 54*, 825–850.
- Labovitz, S. (1968). Criteria for selecting a significance level: A note on the sacredness of .05. *The American Sociologist, 3*, 220–222.
- Lai, M.K. (1973). *The case against tests of statistical significance*. Report from the Teacher Education Division Publication Series. Retrieved from <http://files.eric.ed.gov/fulltext/ED093926.pdf>
- Liu, O. L. (2012). Student evaluation of instruction: In the new paradigm of distance education. *Research in Higher Education, 53*, 471–486.
- Lorber, J. (1994). *Paradoxes of gender*. New Haven, CT: Yale University Press.
- Marsh, H. W. (2001). Distinguishing between good (useful) and bad workloads on students' evaluations of teaching. *American Educational Research Journal, 38*, 183–212.
- Marsh, H. W. (2007). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases and usefulness. In R. P. Perry & J. C. Smart (Eds.), *The scholarship of teaching and learning in higher education: An evidence-based perspective* (pp. 319–383). Dordrecht, The Netherlands: Springer.
- Miller, J., & Chamberlin, M. (2000). Women are teachers, men are professors: A study of student perceptions. *Teaching Sociology, 28*, 283–298.
- Monroe, K., Ozyurt, S., Wrigley, T., & Alexander, A. (2008). Gender equality in academia: Bad news from the trenches, and some possible solutions. *Perspectives on Politics, 6*, 215–233.
- Morgan, S. L., & Winship, C. (2007). *Counterfactuals and causal inference: Methods and principles for social research*. Cambridge, MA: Cambridge University Press.
- Morris, L. V. (2011). Women in higher education: Access, success, and the future. *Innovative Higher Education, 36*, 145–147.
- Morrison, K., & Johnson, T. (2013). Editorial. *Educational Research and Evaluation, 19*, 579–584.
- Murray, H. G. (2007). Low-inference teaching behaviors and college teaching effectiveness: Recent developments and controversies. In R. P. Perry & J. C. Smart (Eds.), *The scholarship of teaching and learning in higher education: An evidence-based perspective* (pp. 145–183). Dordrecht, The Netherlands: Springer.
- O'Sullivan, P. D., Hunt, S. K., & Lippert, L. R. (2004). Mediated immediacy: A language of affiliation in a technological age. *Journal of Language and Social Psychology, 23*, 464–490.
- Paludi, M. A., & Strayer, L. A. (1985). What's in an author's name? Differential evaluations of performance as a function of author's name. *Sex Roles, 12*, 353–361.
- Perry, R. P., & Smart, J. C. (Eds.). (2007). *The scholarship of teaching and learning in higher education: An evidence-based perspective*. Dordrecht, The Netherlands: Springer.
- Risman, B. J. (2004). Gender as a social structure: Theory wrestling with activism. *Gender & Society, 18*, 429–450.
- Rowden, G. V., & Carlson, R. E. (1996). Gender issues and students' perceptions of instructors' immediacy and evaluation of teaching and course. *Psychological Reports, 78*, 835–839.
- Sandler, B. R. (1991). Women faculty at work in the classroom, or, why it still hurts to be a woman in labor. *Communication Education, 40*, 6–15.
- Sidanius, J., & Crane, M. (1989). Job evaluation and gender: The case of university faculty. *Journal of Applied Social Psychology, 19*, 174–197.
- Simeone, A. (1987). *Academic women: Working toward equality*. South Hadley, MA: Bergin and Garvey.
- Skipper, J. K., Guenther, A. C., & Nass, G. (1967). The sacredness of .05: A note concerning the uses of statistical levels of significance in social science. *The American Sociologist, 1*, 16–18.
- Sprague, J., & Massoni, K. (2005). Student evaluations and gendered expectations: What we can't count can hurt us. *Sex Roles, 53*, 779–793.
- Statham, A., Richardson, L., & Cook, J. A. (1991). *Gender and university teaching: A negotiated difference*. Albany, NY: State University of New York Press.
- Subramanya, S. R. (2014). Toward a more effective and useful end-of-course evaluation scheme. *Journal of Research in Innovative Teaching, 7*, 143–157.
- Svanum, S., & Aigner, C. (2011). The influences of course effort, mastery and performance goals, grade expectancies, and earned course grades on student ratings of course satisfaction. *British Journal of Educational Psychology, 81*, 667–679.
- Svinicki, M., & McKeachie, W. J. (2010). *McKeachie's teaching tips: Strategies, research, and theory for college and university teachers* (13th ed.). Belmont, CA: Wadsworth.
- Theall, M., Abrami, P. C., & Mets, L. A. (Eds.). (2001). *The student ratings debate: Are they valid? How can we best use them?* San Francisco, CA: Jossey-Bass.
- West, C., & Zimmerman, D. H. (1987). Doing gender. *Gender & Society, 1*, 125–151.
- Young, S., Rush, L., & Shaw, D. (2009). Evaluating gender bias in ratings of university instructors' teaching effectiveness. *International Journal of Scholarship of Teaching and Learning, 3*, 1–14.

## ***Reconsidering FCQs (Faculty Course Questionnaires)***

**University of Colorado Denver  
Faculty Assembly Minority Affairs Committee (MAC)**

---

### **Background**

MAC is charged with making recommendations on the hiring, support, and retention of a diverse faculty at CU Denver. In 2012-13, MAC held a series of Diversity Dialogues to gather impressions on the environment at CU Denver with respect to inclusiveness. These Dialogues were conducted in three sessions with students, faculty/staff, and administrators, respectively.

In 2013-14, MAC followed this series with a goal of tackling the issue of retention more directly. The first step was to partner with the Office of Institutional Research and Effectiveness (OIRE) to gather and analyze FCQ data to determine if bias exists with respect to faculty members from underrepresented ethnic/racial groups. OIRE gathered and analyzed the data and reported findings to MAC.

---

### **Event Summary**

Seventeen people in total attended the session on March 14, 2014 from 11:00-12:15 in the Auraria Library from a wide variety of departments including Auraria Library, Sociology, Business School, Women's and Gender Studies, English, Anthropology, Political Science, Integrative Biology, CLAS, OIRE, CU Online/OIT. Rank and titles ranged from Lecture to full Professors, and also included a Research Analyst, Senior Business Professional, Associate Dean, and Vice Chancellor.

The forum was moderated by a MAC co-chair and was an open question/answer session. The campus RTP committee had a conflicting meeting and was unable to send a representative, but asked that the session be taped instead.

First, audience members were given a brief overview of some research in the area. For example, minority faculty are rated lower on student evaluations, there are preconceptions of warmth capability and political bias reflected in questionnaires<sup>1</sup>, Latino faculty are rated lower in general<sup>2</sup>, and strict Latino faculty are rated even lower<sup>3</sup>.

Next, the findings from OIRE were then summarized. In short, controlling for as many factors known to affect FCQ's (course level, instructional method, enrollment, term, percentage female, average student age, average student GPA, average course GPA, instructor rank, instructor gender), minority instructors were scored approximately 0.25 points lower on both the instructor

---

<sup>1</sup> Anderson, K. J., & Smith, G. (2005). Students' preconceptions of professors: Benefits and barriers according to ethnicity and gender. *Hispanic Journal of Behavioral Sciences*, 27(2), 184-201.

<sup>2</sup> Reid, L. D. (2010). The role of perceived race and gender in the evaluation of college teaching on RateMyProfessors.com. *Journal of Diversity in Higher Education*, 3(3), 137-152.

<sup>3</sup> Smith, G., & Anderson, K. J. (2005). Students' ratings of professors: The teaching style contingency for Latino/a professors. *Journal of Latinos & Education*, 4(2), 115-136.

and course questions on FCQ's, on average. This effect was moderated but not eliminated for the interaction of female and minority faculty members.

The moderator invited participants to join in the discussion by asking the following questions:

- Do assumptions about gender, race, sexual orientation, national origin and language shape FCQ student responses?
- What is the potential impact of FCQs on the underrepresented racial-ethnic faculty of CU Denver?
- Are there other limitations/hurdles in the tenure and promotion process for underrepresented racial-ethnic faculty?
- What are the alternative methods of evaluating teaching?
- How can we support faculty in the tenure and promotion process?
- How can we educate students on providing constructive feedback?
- How can MAC help you?

Detailed notes on the responses can be found in the full report.

---

## **Summary & Conclusion**

---

As the population of students on the CU Denver campus continues to diversify, it is apparent that conversations like these need to continue and turn into actionable items to create a welcoming and inclusive environment for both faculty/staff as well as students. Several takeaways from this conversation make it apparent that we are not where we need to be, yet. Many faculty members voiced concerns about “teaching to the FCQ’s” by altering their normal methods in an attempt to please everyone as well as shying away from topics germane to the course to not offend anyone. The general feeling among the audience implied that this environment with so much emphasis on the FCQ is restrictive and possibly even unhealthy for the instructors as well as the students. General sentiment was that faculty members must take it upon themselves to prove excellence via multiple means of evaluation. There was also some discussion about guidance on this being located somewhere in the Regents’ documents. However, it is not apparent what resources are available to faculty in this regard or where to go for help. Coupled with this extra burden for finding and documenting teaching excellence are two factors that can harm minority faculty even more in the tenure and promotion process. First, on the research side some of the topics that are of interest to minority faculty may be considered “fringe” research and thus more difficult to achieve publication in the premier outlets. Second, on the service side many minority faculty feel a sense of “cultural taxation” by being asked/expected to serve in multiple service capacities to help these committees feel diverse. This can result in a self-fulfilling negative effect on tenure and promotion as faculty are burdened with more work, marginalized in research, and restrained in teaching. Creating an environment that welcomes diverse faculty as well as helps them to be successful should be of the highest priority. The next level of conversation should be shifted toward the actionable items in this regard.

**Reconsidering FCQ's, an event hosted by the  
Minority Affairs Committee (MAC) of  
University of Colorado Denver Faculty Assembly**

**Background**

MAC is charged with making recommendations on the hiring, support, and retention of a diverse faculty at CU Denver. In 2012-13, MAC held a series of Diversity Dialogues to gather impressions on the environment at CU Denver with respect to inclusiveness. These Dialogues were conducted in three sessions with students, faculty/staff, and administrators, respectively.

In 2013-14, MAC followed this series with a goal of tackling the issue of retention more directly. The first step was to partner with the Office of Institutional Research and Effectiveness (OIRE) to gather and analyze FCQ data to determine if bias exists with respect to faculty members from underrepresented ethnic/racial groups. OIRE gathered and analyzed the data and reported findings to MAC.

**Attendees**

Seventeen people in total attended the session on March 14, 2014 from 11:00-12:15 in the Auraria Library from a wide variety of departments including Auraria Library, Sociology, Business School, Women's and Gender Studies, English, Anthropology, Political Science, Integrative Biology, CLAS, OIRE, CU Online/OIT. Attendees rank/titles:

- Lecturer = 2
- Sr. Instructor = 1
- Asst. Professor = 5
- Assoc. Professor = 3
- Professor = 1
- Assoc. Dean = 1
- Vice Chancellor = 1
- Director = 1
- Sr. Business Professional = 1
- Research Analyst = 1

The forum was moderated by a MAC co-chair and was an open question/answer session. The campus RTP committee had a conflicting meeting and was unable to send a representative, but asked that the session be taped instead.

## Questions and Responses

First, audience members were given a brief overview of some research in the area as well as findings for CU Denver. For example, minority faculty are rated lower on student evaluations<sup>1</sup>, there are preconceptions of warmth capability and political bias reflected in questionnaires<sup>2</sup>, Latino faculty are rated lower in general, and strict Latino faculty are rated even lower<sup>3</sup>.

Next, the findings from OIRE were then summarized. In short, controlling for as many factors known to affect FCQ's (course level, instructional method, enrollment, term, percentage female, average student age, average student GPA, average course GPA, instructor rank, instructor gender), minority instructors were scored approximately 0.25 points lower on both the instructor and course questions on FCQ's, on average. This effect was moderated but not eliminated for the interaction of female and minority faculty members.

The moderator invited participants to join in the discussion by asking the following questions:

### **Do assumptions about gender, race, sexual orientation, national origin and language shape FCQ student responses?**

- Associate Professor: Would also be interesting to know how you change your teaching based on knowledge that assumptions about gender, race, sexual orientation, national origin and language shape FCQ student responses?
- Yes (comment from lecturer): It can be uncomfortable, and I struggle with feelings pertaining to job security. I also pay attention to how information on the dominant group is presented, being careful not to appear biased.
- Yes (Associate Professor): I often need to reorient my teaching toward the subset of class that will react negatively; this affects the other students that are there to learn.
- Vice Chancellor: Students can look for “agendas” of minority faculty. During PTR, it is necessary to ensure that the faculty under review can assist with going through the comments to understand the context of the FCQ scores.
- Strict Latino professor – try to be more communicative.
- Latina instructor – careful about how she is teaching information about dominant groups. She does not want to be accused of tearing down dominant groups. Impacts the books/readings she selects. Does not feel like she has job security.
- Teaches human evolution – student comment “I did not come from monkeys, but maybe you did.” Yes, students perceptions do have an impact.
- Forum is an important opportunity to share experience.
- If a faculty member is teaching a topic that is already challenging to the students and the instructor is from racial-ethnic group, this combination causes more tensions.
- Stereotype threat matters.

---

<sup>1</sup> Anderson, K. J., & Smith, G. (2005). Students' preconceptions of professors: Benefits and barriers according to ethnicity and gender. *Hispanic Journal of Behavioral Sciences*, 27(2), 184-201.

<sup>2</sup> Reid, L. D. (2010). The role of perceived race and gender in the evaluation of college teaching on RateMyProfessors.com. *Journal of Diversity in Higher Education*, 3(3), 137-152.

<sup>3</sup> Smith, G., & Anderson, K. J. (2005). Students' ratings of professors: The teaching style contingency for Latino/a professors. *Journal of Latinos & Education*, 4(2), 115-136.

## **What is the potential impact of FCQs on the underrepresented racial-ethnic faculty of CU Denver?**

- Associate Professor: I have heard from junior faculty that they don't think they could be here another year. There are other environments that may be more welcoming.
- Assistant Professor: In tenure considerations, it can be hard to prove excellence in teaching. This is even more difficult for non-majority faculty given the biases leading to lower than average FCQ scores.
- Lecturer: TAs and lecturers may need to especially "teach" towards the FCQs since future jobs are affected by it – to the point of buying donuts, pizza, etc. near FCQ time.
- Associate Professor: It is important to move emphasis away from FCQs to other indicators of effective teaching.
- Associate Professor: There are other factors looked at rather than FCQs; however, it is hard to resist the temptation to put emphasis on something that is a numeric score.
- Current FCQs build on what students already know. This can be tough since it is hard to build off of students with particularly poor backgrounds.
- Vice Chancellor: There is often useful information in the FCQs that can be used to improve a faculty member's teaching; however, if the faculty member believes they are biased, that may limit the use of FCQs towards improvement.
- Vice Chancellor: FCQs can have more of an effect on lecturers since they don't have multiple means of teaching evaluation to the extent that tenured faculty members do through the PTR processes.
- Associate Professor: I have found it helpful to look back at comments over the last 5 years for patterns, especially in the written comments. I use these to demonstrate how they have changed the way I teach.
- Have heard from minority faculty who do not intend to stay after 1 year because not feeling welcome or supported.
- Advice received by racial-ethnic faculty. Excellence in teaching and research. Teaching is too hard, so focus on research.
- MAC wants to move the conversations away from student satisfaction, and towards student learning.
- Tenure committees are looking at more than FCQs. How is teaching improving?
- Students may not remember everything that occurred during the class. Only administered at one time. How reliable are the scores? How students are evaluating is like a roller coaster.
- Less flexibility for instructors and TAs for evaluations. They rely heavily on FCQs. How can we be supportive of all faculty?
- Advice to faculty. In dossier, acknowledge that you are "rigorous" in the narrative. But also describe how you are "fair." Be reflective. If receiving the same comments, then it's not an outlier.
- At comp review, 1 member of the primary unit committee looked through all my FCQs and found 1 negative comment and included it in the letter. Received 1 negative vote.

### **Are there other limitations/hurdles in the tenure and promotion process for underrepresented racial-ethnic faculty?**

- Vice Chancellor: Often different expectations exist for non-majority faculty members, often related to service, which then eats away at time for research and teaching. This is particularly true when the faculty member is being chosen to “represent” their group (cultural taxation).
- Lecturer: On positive side, universities are looking for minority faculty. Perhaps that is hopeful.
- Recruiting practices are often good for increasing the number of entering minority faculty members; however, afterwards institutions often do little for retention and adopt a “you are on your own” attitude.
- Students are often unaware of the biases they bring to the process as are administrators and other faculty members.
- Vice Chancellor: Often minority faculty use different approaches in their scholarship compared to majority faculty or main stream research. They often publish in other outlets. During evaluation this can be a detriment. Education is needed to explain these differences during evaluation and demonstrate contribution.

### **What are the alternative methods of evaluating teaching?**

- Find a way to connect teaching, service and research activities to society and societal problems.
- Associate Professor: I teach a summer field school, often without pay. These types of classes aren’t given much credit despite the immense contribution to learning.
- Vice Chancellor to MAC: We should be thoughtful in terms of not only what we as individuals can do to address these biases but also think about systemic change, raising the consciousness of those making decisions in administration. Consider policies and practices along these lines. What are other institutions doing? What systematic changes can be suggested?
- Vice Chancellor: I suggest faculty members conduct a thoughtful mid-semester evaluation to look at feedback and make changes as needed. What is going well? What is not going well?
- Responding to the needs of society. Beyond me and I. What is the impact? Service learning.
- We need systemic change. Looking at policies. Need to raise consciousness of these issues. If we really want to have under-represented faculty, we need to have look at everything.
- CU Online has midterm survey that mirrors FCQs.
- Ask mid-term questions in the middle of class.
- Students are hesitant to hand write their comments. Afraid that the instructor will recognize their handwriting. Suggest using google docs, which can be completely anonymous.

### **How can we support faculty in the tenure and promotion process?**

- When I first arrived, there was a mentoring program run by Brenda Allen. It made a huge difference to me, e.g., helped with feelings of isolation. I now try to provide this to others, e.g., sit in their classes to provide feedback.
- Regarding peer feedback on teaching: Should be done more than just at the last semester before it is needed. When receiving peer feedback, have an initial meeting (explain teaching philosophy etc.) and then again post meeting. It is helpful to provide a form or suggested questions for evaluation. These things ensure the peer evaluation will be more likely useful.
- Associate Dean: Faculty should be proactive in the creation of the dossier. Be familiar with departmental criteria.
- Creating community
- Having faculty willing to sit in on my teaching – peer evaluation. Not consistent evaluators. Better to go through Center for Faculty Development. Using a template. Should happen at different times. Problem some departments have rotating process – Jr faculty evaluating a Sr. faculty. Power dynamics matters.
- RTP looks at who are giving the peer-evaluations. Independent evaluators are more meaningful.
- Can modify tenure criteria through department actions.
- FCQs have less value these days on FTP cases. Looking for outliers. Multiple layers of decision makers.
- Chairs also need more training on mentoring Jr faculty. They need to know how to document evidence for their dossier.

### **How can we educate students on providing constructive feedback?**

- Encourage students to think about providing comments that would be helpful for future offerings. Ask them to be reflective and take the time needed to provide helpful critique.
- CU Online: They do mid semester evaluation with questions geared towards final FCQ questions. This allows faculty a chance to make necessary changes.
- Have evaluations done using google docs so comments are typed and clearly anonymous.

### **How can MAC help you?**

- Seek funding to do focus groups and find out how people are feeling. Make the university aware of the climate.
- Provide information to evaluation committees so that they are aware of the biases that do exist in FCQ responses as well as the implicit biases in students, departments and faculty.
- Provide training for faculty on how to conduct evaluations. Often training is focused on what you shouldn't do rather than on what you should do. Also, teach faculty and students how to provide useful critiques.
- Perhaps consider creating and providing a best practices document.
- Help educate the problem with unintentional bias from student, faculty colleagues, and administrators.
- Brenda is doing a diversity workshop for search committees. Should we do something similar for faculty?



## Summary and Conclusion

As the population of students on the CU Denver campus continues to diversify, it is apparent that conversations like these need to continue and turn into actionable items to create a welcoming and inclusive environment for both faculty/staff as well as students. Several takeaways from this conversation make it apparent that we are not where we need to be, yet. Many faculty members voiced concerns about “teaching to the FCQ’s” by altering their normal methods in an attempt to please everyone as well as shying away from topics germane to the course to not offend anyone. The general feeling among the audience implied that this environment with so much emphasis on the FCQ is restrictive and possibly even unhealthy for the instructors as well as the students. General sentiment was that faculty members must take it upon themselves to prove excellence via multiple means of evaluation. There was also some discussion about guidance on this being located somewhere in the Regents’ documents. However, it is not apparent what resources are available to faculty in this regard or where to go for help. Coupled with this extra burden for finding and documenting teaching excellence are two factors that can harm minority faculty even more in the tenure and promotion process. First, on the research side some of the topics that are of interest to minority faculty may be considered “fringe” research and thus more difficult to achieve publication in the premier outlets. Second, on the service side many minority faculty feel a sense of “cultural taxation” by being asked/expected to serve in multiple service capacities to help these committees feel diverse. This can result in a self-fulfilling negative effect on tenure and promotion as faculty are burdened with more work, marginalized in research, and restrained in teaching. Creating an environment that welcomes diverse faculty as well as helps them to be successful should be of the highest priority. The next level of conversation should be shifted toward the actionable items in this regard.

# MARCH 14, 2014

11:30am - 1:00pm

AURARIA LIBRARY

ENHANCED LEARNING CENTER

PLEASE  
**SAVE  
THE  
DATE**

## Reconsidering FCQs FOR CU DENVER FACULTY

Reconsidering FCQs is a forum to discuss  
the role of FCQ's in promotion and tenure,  
and their limitations as a measurement  
of effective teaching.

Research has indicated that student responses to their instructors can often be shaped by assumptions about gender, race, sexual orientation, national origin, and language.

With this research in mind, we hope to gather a diverse group of faculty to share experiences, and to discuss the potential impact of FCQs on the retention of CU Denver faculty from underrepresented racial-ethnic faculty groups.

Join the CU Denver Minority Affairs Committee for lunch to learn about the quantitative data from CU Denver Office of Institutional Research & Effectiveness and to provide insights and direction on finding alternative methods of evaluating teaching.

Please RSVP for lunch by clicking on this link  
<http://tinyurl.com/ReconsideringFCQs>

\*This event is funded by the Office of Diversity and Inclusion.