A QuantCrit Approach: Using Critical Race Theory as a Means To Evaluate If Rate My Professor Assessments Are Racially Biased

Chuck Alan Baker
Delaware County Community College, USA

ABSTRACT

Research on race is a paradigm in qualitative methodology. Researchers believe that when analyzing discrimination, numerical data may miss the subjective characteristics of bigotry. Since the early 1990s, research utilizing critical race theory in education employed a qualitative approach. Recent research using critical race theory includes a quantitative approach called QuantCrit. The online faculty evaluation site called Rate My Professor (RMP) is designed to allow students an opportunity to appraise faculty performance. Using evaluations of faculty in a Pennsylvania college from both RMP and IOTA360, this research examines the validity of RMP in analyzing minority faculty's teaching. As predicted by applying a QuantCrit approach, results support that RMP evaluations show a race bias.

Keywords: African American, Critical Race Theory, discrimination, faculty evaluations, minority, QuantCrit, racism

BACKGROUND

Issues of race, class, and gender are deliberated in colleges throughout the United States. Social science disciplines often acknowledge the vestiges of bigotry, and courses on racism, diversity, ethnicity, sexual orientation, and other manners of discrimination rooted in America's history are designed to enlighten students on the systemic and institutionalized manners of

discrimination that yet persist. Critical Race Theory (CRT) stipulates that race matters because of its systemic nature, that there are consistent reminders of it, and that an inherent white ownership of American cultural and legal doctrine requires that any progress in social equity must benefit non-minorities also (Ladson-Billings & Tate, 1995). Bell (1992) believed that the American structure relegates African Americans into a perpetual underclass that is enduring, and more recent CRT research places emphasis on school systems as a protagonist supporting inequality (Hiraldo, 2010; Ladson-Billings & Tate, 1995).

The schooling system doesn't simply instill manifest knowledge. One of its primary latent functions is to fortify traditional norms and normative behaviors. It is the schooling system that assists in the integration of citizens into society (Henslin, 2007). CRT takes the approach that minority integration into America reproduces disparity and that the institution of education is a leading protagonist of socialized inequality. CRT theorists do not claim that race relations have not changed. The central premise is that race matters and African Americans are the recipients of differential treatment in society (Bell, 1992; Hiraldo, 2010; Ladson-Billings, 1995). Much of the Civil Rights era was predicated upon the ideology that laws, integration, education, and assimilation dilutes racism. CRT theorists argue that laws, education, and integration may alter race relations but do not eradicate racial discrimination. Take, for instance, the Black Lives Matter movement. A 2017 Pew Center poll shows that 73% of postgrads, 63% of college graduates, 54% of those with some college education, and 47% of high school (or less) respondents support the Black Lives Matter movement (Neal, 2017). More educated people support the movement, and yet, social movements to offer liberty to minorities are nonetheless necessary. According to the National Science Foundation (2015), the largest percentage of doctorates conferred to African Americans is in the field of education, and as African Americans desire to effectuate change and establish egalitarianism, many well-educated minorities become college administrators (Hiraldo, 2010). Administrators, who are responsible for the fiscal health of an institution, place value upon student evaluations and may include them in decisions of faculty tenure and promotion. In contrast, faculty, who are responsible for educating students into their chosen professions, may see evaluations as likability of faculty appraisals similar to satisfaction surveys (Patton, 2010). Either the faculty is liked or not and those who aren't liked often offer rigorous courses. Given the importance that administrators place on student evaluations and their responsibilities of budget jurisdiction, online assessments provide evaluation data that seem

cost efficient. Online evaluations, however, have low response rates (Patton, 2010) and increase measurement error.

Research supports the distinction between online and face-to-face student evaluations. Nowell et al. (2010) showed that student evaluations of faculty who teach online are lower, on average, than faculty who teach in a traditional classroom orientation after controlling for factors that are not under the faculty's control (e.g., class size and time). In essence, when looking at characteristics that are under the faculty's control and are related to instructor capabilities, the difference between online evaluations and face-to-face class evaluations was significant. Since smaller sample sizes inflate measurement error, faculty often prefer evaluations of their classroom courses given the lower response rates of online students (Sax et al., 2003). CRT posits that race is another factor out of a faculty member's control that may influence student evaluations. Students often place salience upon unmonitored online evaluations that are easily accessible and gives them some information on faculty and the courses they deliver.

Rate My Professors (RMP; www.ratemyprofessors.com) is a free online website for users at any higher education institution to anonymously provide evaluations of their instructors. These evaluations are disconcerting for a number of reasons. First, RMP content does not include dialogue between faculty and administrators. In addition, anyone, a student or not, can complete RMP evaluations. Third, much of the ratings' content has very little to do with faculty's professional capabilities. Faculty ability, for instance, has nothing to do with physical attractiveness and yet, RMP has a hotness chili pepper to grade attractiveness. The website's novelty approach is very popular with people. Along with the superficial nature of assessing physical attractiveness, could other characteristics, unrelated to pedagogy, influence RMP evaluations? Since racism is salient in society and has motivated discriminatory practices in education and employment opportunities (Baker 2017; Omi & Winant, 1994), open forums like RMP could have a race bias in evaluations. If race matters, according to CRT, it would likely influence teacher ratings. While faculty may believe that the information stated on its site do not accurately depict the course delivered, students may use RMP evaluations to make decisions about which courses to take. From the CRT perspective by using a QuantCrit approach, this research examines if there is a race bias in faculty evaluations on RMP.

LITERATURE REVIEW

Critical Race Theory

Evolving in Bell's (1992) thesis titled "Faces at the Bottom of the Well: The Permanence of Racism," CRT posits that the structural inequalities embedded in America are ongoing due to the persistent apparatus supporting social inequity created through the legacy of discrimination. CRT has several fundamental creeds that allow for its application. First, society will often place salience on overt and vicious forms of racism and yet miss the subtle forms that perpetuate inequality. When the media coverage about the slayings of Trayvon Martin and Tamir Rice are discussed, they offer validity to concerns about the continued existence of racism. Yet, it is the covert, non-violent, and painful manners of racism that the discriminatory apparatus supports. Second, that racism is enduring. Although minorities have earned citizenship since the 19th century, discrimination is still pervasive. Third, race relations change when there are converging interests between minorities and whites. Minority social gains are couched in an ideology of social equity placed on a gradient leaning toward whites. Fourth, white privilege is a property right for whites—the right of home ownership, employment, and upward mobility in employment, and the lack of social restrictions that minorities face. Fifth is perceptions of equity and colorblindness under the law. When whites experience tangible losses (e.g., jobs) to minorities on Affirmative Action legislative justifications, they claim their 14th Amendment rights have been violated. In essence, the law designed to establish citizenship equality (e.g., U.S. Constitution) of minorities is used against Affirmative Action by whites.

The Education Pipeline

According to CRT, discrimination against people of color is an enduring element of American society. Groups desire to maximize their gains and minimize their losses (Hobfoll et al., 2002; Hobfoll & Dekel, 2007). Resource acquisition and control may converge around issues of race. Resource competition, resource scarcity, and resource control motivate intergroup conflict, and the group suitability of an individual is often based upon the other group members' acceptance of the original person's phenotype, which is often embedded in issues of race (Stephan et al., 2006; Stephan et al., 2009). Education is a resource that can be exchanged for economic gain in America. This is the intersection between education and race. At each phase in the education pipeline, from entering college and

getting adequate support to going into respective employment careers, racial diversity decreases (King, 2016). According to the National Center for Education Statistics (2013) only 19% of faculty in America are racial/ethnic minorities and many of these are adjunct faculty. In the 2011–2012 school year, 19% of the students who graduated high school were Latino. Yet, only 13% of first semester college freshmen were Latino. The African American percentages were flat for the transition from high school to college, but 57% of students who graduated high school were white and 62% who entered college were white. In essence, white college student populations were the only racial/ethnic group that increased in their percentage of students entering college while other groups show percentage declines (King, 2016).

The education pipeline provides training to students at key points in their college learning experience as they prepare to be educators. At the final ends of the pipeline, that of entering the workforce and maintaining employment, while educational institutions claim that they desire diversity in faculty, they often will erroneously state a lack of qualified applicants when search processes and hiring decisions are made. According to a report by the Association for the Study of Higher Education (ASHE), when minorities get employment in colleges, they are less likely to receive tenure and be promoted above the associate professor level (Museus, Ledesma, & Parker, 2015). This idea of sifting and sorting through the education pipeline supports the idea that only the most competent minorities survive this gauntlet process. Hiraldo (2010) discussed minority positioning in higher education reinforcing the property right interests of whites. Most African Americans who earn doctorate degrees, for example, do so in education administration. Most tenured faculty in colleges and universities are not minorities. The secure positions for professionals, in academia, are those of tenure. It is the contract of tenure for faculty that has been reaffirmed by the Supreme Court as a property interest in Perry v. Sinderman (Justia, 1972), which buttressed the current race inequality. CRT supports that these property interests serve to separate minority administrators from the minority students and reinforce the students' dependence on non-minority faculty. It is in this vein that the legacy of race relations perpetuates racialized oppression although the legal mechanism did not implement discrimination in a linear manner.

Race and Educational Standards

CRT offers critique of liberal doctrine. In education, as in most social institutions, the perception is that liberalism is unnecessary since social norms and laws are neutral. As institutions of education promote

diversity, it is important that diversity is not couched in color blind racial ideology. The premise that race is perceived through a color blind racial ideological lens denies the reality of established norms that promote inequality. Given the legacy of discrimination, minimizing race is similar to the apathy that promotes racism. To be clear, color blind racial ideology promotes conventionalism supporting a status quo which, in-turn, supports racial inequity in America (Poteat & Spanierman, 2012). CRT acknowledges that race matters in America. The durability of racial inequality is not only in the economic, political, and social realms of society, but it also penetrates educational norms and pedagogical instruction by educators. The white privilege doctrine turns the epistemological rationale for quality educators upon its head. Kuh, Nelson, and Umbach (2004) examined the role of race and gender in pedagogical delivery. Faculty of color and women are more likely to perform better on the effective use of several educational practices including (a) academically challenging curricula for students, (b) collaborative learning, (c) an emphasis on diversity experiences, (d) facilitating critical thought, and (e) fostering better educational experiences for students (Kuh et al., 2004). According to the Association of American Colleges and Universities (AACU, 2002), the activities that they stipulate as important in student–teacher engagement are (a) collaborative learning opportunities, (b) class assignments in which higher order intellectual tasks are required, (c) assignments that ask for diverse perspectives, (d) challenging activities, and (e) analysis versus memorization. The overlap between the AACU and Kuh et al. (2004) is substantial and yet, most faculty positions are held by white males. Along the lines of interest conversion in CRT, of those post-Civil Rights increases in minority faculty, most are of white women.

Student Evaluations

A primary difficulty with student evaluations is that they may be motivated by student gratification with the ease of the course, the students' grade in a course, or a host of factors for which faculty are not responsibe, rather than reflective of mastering the material (Aleamoni, 1999; Kulik, 2001; Lawrence, 2018; Neath, 1996). Take, for example, faculty office hours. Fusani (1994) showed that one in four students never contacted an instructor outside of the classroom. Griffin et al.'s (2014) research reported that 66% of students never used office hours in their course. Although students do not attend office hours, they often evaluate faculty with criteria that includes the student's ability to understand course content and access to the professor. Other factors that are not under the faculty's control are class

size, course level, tutoring, and whether the course is an elective or not (Griffin et al., 2014). Scriven (1995) also cited assignments outside the classroom, course attendance, and textbook costs as associated with student appraisals, although these factors may have little to do with faculty competence. In this manner evaluations prove as a means to evaluate faculty on student's perceptions of their grade and how much they like the faculty (Germain & Scandura, 2005). Student perceptions of the grade they should earn correlate with the evaluations of instructors (Snyder & Claire, 1976). Thus, if there are ample opportunities to receive the grade desired, the instructor's rating is more favorable (Aronson & Linder, 1965).

Research by Bavishi, Madera, and Hebl (2010) further supports that students evaluate professors on issues beyond their control, such as on race and gender. Bavishi et al. had a sample of students who had just begun their college experience rate the capability of a faculty by examining the faculty's resume. The researchers manipulated perceptions of the race of the faculty by changing their organizational affiliations and disguised gender through changing the name on the resume. The researchers found that resumes appearing to be from white males got the most favorable resume rating, even when students were examining the same resume and the only changes were organizations and gender name. When the resume had an African American organization affiliation designation, it was given the least favorable rating (Bavishi, Madera, & Hebl, 2010).

RMP and IOTA360 Solutions Evaluation Tools

RMP has over 19 million ratings about professors and colleges (www.ratemyprofessors.com). As a point of advertisement, the website discusses these ratings as a manner of "joining the fun." Items included in the rating of a professor include the amount of homework and reading students must fulfill, as well as if faculty are caring, respectful, and inspirational. RMP's design offers no means to filter out people who have not taken a course with the professor who is being evaluated. In essence, people who do not know the professor may offer a rating to inflate or deflate the faculty's rating.

IOTA360 is another website that offers an evaluation instrument so that students may assess faculty competence (iota360.com). Among other questions, IOTA360 faculty evaluations request student feedback on syllabus clarity, course design, and instructor expectations. These criteria are part of the faculty's job responsibilities. At a college in Pennsylvania students are given access to IOTA360 evaluation surveys of faculty several weeks before the semester ends. During these few weeks, students are

prompted several times to complete the evaluation. In addition, the evaluations are secure and only the students in the course, who have their own personal code, can access the evaluation survey.

QuantCrit Approach

CRT describes the structural factors embedded in institutions that reinforce racial inequalities. Historically, much of the research has been qualitative and over the last 15 years has focused upon education (Ladson-Billings & Tate, 1995). Quantitative methods have been less than desirable given the manners in which such data is used to mislead, perpetuate inequality, and dilute attempts at social justice. In 2001, for example, the American Council on Education reported that 32% of all doctorates were conferred upon Asians. A close examination of the data, however, shows that 86% of those doctorates earned by Asians were to non-U.S. citizens from Asian nations (Teranishi, 2007). In the present, attempts are being made to revive quantitative approaches to examine race issues (Gillborn et al. as cited in Garcia et al., 2017). Gillborn et al. authored several defining principles for a QuantCrit approach. These principles include that (a) racism is multifaceted and not readily quantifiable, (b) data sources may be biased and often benefit white race interests, (c) race categories are not fixed or innate and therefore, first-hand knowledge of minority experiences is important in such analysis, and (d) statistical assessment is not the entirety of race analysis but may be part of the encounter for race equity. In this manner, Gillborn et al. stated the valuable but limited role that quantitative approaches play in examining racism. Zuberi (2001) discussed the use of data during the eugenics movement of the 20th century and in 2001, posed the question "How do we deracialize the social conditions that produce racialized inequalities?" In "The Death of White Sociology: Essays on Race and Culture" Ladner (1973) stated that personal values and social location should be filtered out of quantitative approaches in research. Garcia et al. (2017) made clear that any quantitative analysis on racial justice must take into account the intersection of powerful social, political, economic, and historical factors. The social implications of institutionalized education is that it is a middle class pedagogy and in addition, historically, schools have taught an ethnocentric white doctrine. The QuantCrit approach taken in this research acknowledges the intersectionality of these factors and their embeddedness in the institution of education. This research uses a QuantCrit approach to deconstruct how social dynamisms through online media evaluation tools may reinforce perceptions of race inequality.

RESEARCH METHOD

RMP is a free website that gives people the opportunity to evaluate the performance of college professors. In RMP, people may appraise whether they would "take the professor again," the "level of difficulty" of the course(s), the "hotness" of the professor, and the "overall quality" of the course(s). In addition to these four criteria, statements may be made in which the student describes the professor through a tag and, in addition, makes open-ended statements about the professor. Several of the possible tags may include *tough grader, participation matters, beware of pop quizzes, respected, lots of homework, test heavy, inspirational, get ready to read, hilarious, caring, graded by a few things,* and *amazing lectures*. Furthermore, RMP scores are on a Likert scale of 1–5 for the overall quality of a professor's course and the overall difficulty of the professor's course. The overall quality of a professor's course is rated by how well the professor teaches, how helpful he/she is during and after class, and the professor's approachability. The level of difficulty is a lone question based upon a scaled score (1–5).

This research utilizes the aggregate overall quality and difficulty RMP scores of minority and non-minority faculty in a college located in Pennsylvania. The RMP non-minority teaching quality evaluations scores range from a low 2.8 to 4.8. The mean score is 4.16. The non-minority level of course difficulty scores range from 1.0 to 3.8. The mean course difficulty evaluation score for non-minority faculty on RMP is 3.10. The range of scores for RMP evaluations of minority faculty's teaching quality is a low of 1.9 and a high of 5.0. The mean teaching quality score for minority faculty was 3.71. The range for the minority faculty course difficulty score is a low of 1.9 to a high of 4.0 with a mean of 3.10.

IOTA360 Solutions offers standardized evaluation products that colleges utilize to examine their faculty. Some of the questions used for the evaluation of faculty are "exams and quizzes and assignments reflect important aspects of the course," "how challenging is this course when compared to other courses," "this course stimulated my critical and analytical thinking," "assigned readings/materials are valuable in learning this course content," "the pace at which we are covering the material is about right," and "this course is meeting the objectives outlined by its competencies." In addition, students may answer questions that evaluate their desire to learn material. These questions include "effort put into this course when compared to other courses," "I am motivated to learn in this course," and "on average how many hours do you spend on this course." To

maintain equivalence with the RMP quality of course factors the data recorded by IOTA360 student evaluations were "instructional material adequate," "the pace of the course is about right," "the course stimulates analytical and critical thought," "the instructor is well prepared for class," "the instructor responds effectively to student questions," "instructor creates an environment in which students equally," "instructor knows when students do not understand," "instructor willing to assist outside of class," and "students can freely approach professor." The IOTA360 student evaluation responses are on a Likert scale from 1 to 4. The responses range from *Strongly Disagree* (1), *Disagree* (2), *Agree* (3), and *Strongly Agree* (4). These survey questions address the RMP factors for the overall quality of the professor's class. For white faculty, the range of IOTA360 student evaluation scores was 2.8 at the low boundary and at the upper boundary it was 4.0. The mean white teaching score from was 3.71. For minorities, the mean student evaluation score was 3.42. The minority scores range from 2.1 at the lower boundary to 4.0 at the upper boundary.

Participants

During the summer of 2018, a purposive sample of 48 instructors in a college in Pennsylvania was used to examine if there is a race difference in student evaluations. Race designation was gained from the college's records office. Through voluntary reply, race designation is recorded by the Department of Human Resources during the hiring process. Similar to Census Bureau designations, whites were non-Hispanic white. There are 153 full-time faculty at the college. Of the 153 full-time faculty, 125 are white. To get an appropriate sample, faculty who are not white were included in the minority strata, and these other race groups include African Americans, Asians, Native Americans, Hispanics, and people who are biracial. There are 28 are minority faculty. The minority faculty classifications are 13 African Americans, 10 Asians, four Latinos, and one faculty with two stated race designations. The data from the sample was separated along minority ($n = 21$) and non-minority ($n = 27$) strata for the total ($N = 48$) faculty. The samples were matched with IOTA360 data. These were the race designations used in the IOTA and RMP assessment. These data allow to examine the following hypothesis:

Since RMP is an uncontrolled online evaluation tool, students and non-students may access it. Students who participate may be motivated by a highly positive or negative affect of the professor. Thus, a large proportion of students may avoid participating. To examine the validity of RMP as an evaluation tool, this research compares RMP evaluation results to IOTA360

student evaluation results. IOTA360 gives students access to evaluate their faculty and they are persistently asked to participate. IOTA360 has been shown to depict professor capabilities.

Hypothesis 1: RMP results will differ from IOTA360 results in student evaluation of faculty quality.

Twenty-first century racism is often a subjective experience for minorities with objective results. The application of CRT has traditionally taken a qualitative approach due to the subjective nature of racism. The QuantCrit design lends itself, however, to analyzing statistical differences in academic credentials when comparing minorities and non-minorities. As stated previously, minorities have a more demanding journey through the education pipeline. Perceptions about the pipeline can be examined qualitatively, but degree completion is a quantitative assessment. By the time minorities are hired, there should be a leveling effect when comparing the credentials between the groups. Most institutions of higher learning require faculty to have a master's or doctoral degree to be hired for a faculty position. Given that they are credentialed and have already been through much of the education pipeline, for those who are employed as faculty, race should not influence differences in student ratings. For research purposes, the operationalization of education is that faculty with a master's degree were given a value of one and those with a doctorate were given the value of two.

Hypothesis 2: Minority and non-minority faculty will have similar levels of educational capital and thus, perceived similarities of intellectual competence in their chosen fields.

Since students may perceive that they are purchasing a product instead of an opportunity to learn, they may seek out courses that are not rigorous. Yet, competent faculty may offer demanding classes to facilitate student acumen and mastery of subject matter. After all, people do not want an incompetent surgeon to perform surgery upon them. Thus, competent faculty are motivated to create thorough courses to maintain integrity and ensure that graduates have a commensurate skill set, but students may desire listless courses and the likelihood of social promotion affects enrollment. Thus, it is predicted that the lower the faculty rating by people in RMP, the higher the course difficulty rating.

Hypothesis 3: High RMP teaching quality scores will associate with low RMP course difficulty scores.

CRT supports that society may be willing to confront overt manners of racism. It is the covert and yet structurally entrenched forms of discrimination that perpetuate inequality. Given the salience of race and the aspiration to appear impartial in evaluating minority faculty, people will desire to justify low ratings of minority faculty. Since evaluations are anonymous, based upon CRT they will be likely to have latent discrimination embedded in their application and in the aggregate, minorities will be scored as less competent. Therefore, people who evaluate faculty on RMP will appraise minority faculty courses as more difficult and minority faculty as less capable of teaching than non-minority faculty.

Hypothesis 4: Minority faculty will be more likely to get higher course difficulty scores and lower teaching quality scores in RMP when compared to non-minority faculty.

RESULTS

Tables 1–6 in the analysis uses a comparison of independent means t-test approach. The formula for a *t* test of independence presents as follows:

$$\frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{(N_1 - 1)\sigma_1^2 + (N_2 - 1)\sigma_2^2}{N_1 + N_2 - 2} * [\frac{1}{N_1} + \frac{1}{N_2}]}}$$

Where $\overline{X}_1$ is the mean of group one (e.g., RMP), $\overline{X}_2$ is the mean of group two (e.g., IOTA360), σ_1^2 is the variance of group one scores, σ_2^2 is the variance of group two scores, N_1 is the number of respondents in the first group, and N_2 is the number of respondents in the second group.

Tables 8–9 use an analysis of variance (ANOVA) approach. Table 8 allows for a robust comparison of IOTA360 and RMP course difficulty and teaching competence scores. Table 9 removes the white RMP scores for statistical assessment. The ANOVA formula presents as follows.

$$\text{SUM OF SQUARES BETWEEN} \quad \frac{\sum^K X_i^2 - C.T.}{\sum n_1}$$

SUM OF SQUARES WITHIN $\quad\quad \sum_{ij} \sum (X_{ij} - \overline{X}_i)^2{}_{ij}$

<u>SSB (k-1)</u>

SSW (n-k) = Computed Ratio

Where k is the number of categories, and $\sum X^2{}_i$ is the sum of the squared scores. C.T. is the correction term. $(X_{ij} - \overline{X}_i)^2{}_{ij}$ is the variance within sum of squared scores. The computed *F* score ratio is the sum of squares between divided by (k − 1) as the numerator of the computed ratio, and the sum of squares within is the denominator of the computed ratio.

Table 1 below shows the *t* test of independent means results for quality of teaching at a college in Pennsylvania as measured on RMP and IOTA360. The mean RMP teaching quality score is 3.96. The mean for the IOTA360 evaluations is 3.60. The *t* test results show statistical significance at $p < .05$ with $t = 2.5748$ at 94 degrees of freedom (*df*). RMP scores are statistically higher than IOTA360 evaluation scores.

Table 1. Rate My Professor versus IOTA360 student evaluations (teaching quality).

	N	*M*	*SD*	*SE*
Rate My Professor	48	3.96	0.807	0.117
IOTA360	48	3.60	0.554	0.080

Note. t test $= 2.5748*$, $df = 94$, two-tailed test, $*p < .05$.

Hypothesis 1 states that the RMP results of faculty quality will differ from those of IOTA360. Table 1 shows that this statistical difference does exist. The next hypothesis focuses upon the similarities between the levels of education when comparing non-minority faculty with minority faculty. Faculty education level was designated as 1 (for master's degrees) or 2 (for doctoral degrees). Table 2 shows the results from this analysis.

Table 2. Non-minority and minority education acquired.

	n	*M*	*SD*	*SE*
Non-minority education	27	1.26	0.447	0.086
Minority education	21	1.48	0.518	0.112

Note. t test $= -1.5663$, $df = 46$, two-tailed test

The *t* test for Table 2 above shows that there is no statistically significant difference between non-minority and minority faculty's level of education. The minority faculty have a slightly higher average level of education at 1.48, but the mean is not substantially different than the white mean level of education at 1.26. The standard deviation of non-minorities was 0.447 and the standard deviation for minorities was 0.518. The *t* test showed not to be statistically significant and supports that the level of education between the faculty race groups is similar.

Table 3. Rate My Professor teaching quality and course difficulty.

	N	M	SD	SE
Quality	48	3.96	0.807	0.117
Difficulty	48	2.86	0.690	0.100

Note. t test = 7.1915***, *df* = 94, two-tailed test; *** $p < .001$.

Table 3 shows the results when comparing RMP teaching quality scores with RMP course difficulty scores. Notice that the RMP teaching quality mean is higher than the RMP course difficulty mean. The *t* test results show that the difference between the course difficulty and quality of teaching results in RMP are substantially statistically significant at $\alpha <$ .001. The *t* test is 7.1915 at 94 *df*. The results from the analysis of Hypothesis 3 are corroborated in a statistically meaningful way. Higher difficulty scores are associated with lower teaching quality scores.

Table 4 shows the correlations and coefficients of determination results examining the relationships of RMP teaching and difficulty quality when comparing minorities and non-minorities. The statistics show that for teaching quality scores for white faculty, the 4.16 mean is much higher than the white course difficulty mean of 2.69. The correlation coefficient is −0.5531. There is a fairly strong relationship between difficulty and teaching quality scores for white faculty. In addition, as perceptions of course difficulty diminish, teaching quality increases. Minorities have a similar relationship. Yet, the mean teaching quality score for minorities is only 3.71, while the perceptions of course difficulty are higher than whites at a 3.10 mean value. The correlation for minorities is −0.5093. The coefficient of determination states that perceptions of course difficulty explain about 31% of the quality of teaching for whites and 26% of the teaching quality for minorities. The difference between the white faculty and minority faculty coefficients of determination is 5% better in predicting quality of teaching scores when using course difficulty scores for whites than for minorities.

Table 4. Non-minority and minority mean, correlation, and coefficient of determination (Rate My Professor [RMP] and difficulty).

	n	M	R	R^2
Non-minority RMP	27	4.16	−0.5531	0.3059
Non-minority difficulty	27	2.69		
Minority RMP	21	3.71	−0.5093	0.2595
Minority difficulty	21	3.10		

Hypothesis 4 states that minority faculty will be more likely to get higher difficulty scores and lower quality scores in RMP when compared with non-minority faculty. Table 4 bears this out. Table 5 shows that the teaching quality scores are statistically different at $t = 1.969$ at 46 *df* at $p < .05$.

Table 5. Non-minority and minority Rate My Professor results for teaching quality.

	n	M	SD	SE
Non-minority	27	4.16	0.738	0.142
Minority	21	3.71	0.839	0.183

Note. t test = 1.9685*, *df* = 46, one-tailed test; *$p <$.05.

According to Table 6, the t-test result is significant at -2.034. Thus, minorities are more likely than whites to have their courses rated as 'difficult' in the RMP evaluations. This supports hypothesis 4 that minorities will have lower teaching quality scores and higher course difficulty scores when compared to non-minorities.

Table 6. Non-minority and minority Rate My Professor results for difficulty of course.

	n	M	SD	SE
Non-minority	27	2.70	0.708	0.136
Minority	21	3.10	0.627	0.137

Note. t test = −2.0342*, *df* = 46, one-tailed test; *$p < .05$.

Examining IOTA360 student evaluations in Table 7 shows that the *t* test was not statistically significant at 1.853, *df* = 46, $p < .05$ for a two-tailed test. Therefore, IOTA360 results show that, statistically, non-minority and minority faculty are similar.

Table 7. Non-minority and minority IOTA360 results for teaching quality.

	n	*M*	*SD*	*SE*
Non-minority	27	3.71	0.484	0.093
Minority	21	3.42	0.613	0.134

Note. t test = 1.8534, *df* = 46, two-tailed test

The IOTA360 results show less difference in teaching quality when comparing minority and non-minority faculty than do those from RMP. Thus, students who can be confirmed to be class participants do not evaluate faculty differently, in a statistical manner, based upon race. In contrast, people who go on RMP to evaluate white professors evaluate them much more positively than they do minorities and much more favorably than IOTA360 results (see Table 5 for RMP results and see Table 7 for IOTA360 results). On a scale from 1to 5, the RMP mean evaluation teaching quality score is 4.16 for white faculty. It is 3.71 for minority faculty. Thus, it seems as if the RMP evaluations have a racial lag on minority evaluation results when contrasted with results for white faculty.

This perceived lag on minority RMP scores can be viewed through an ANOVA computation of white and minority RMP and IOTA360 evaluation scores. Table 8 shows the ANOVA results from this assessment.

Table 8. Analysis of variance for non-minority and minority IOTA360 and Rate My Professor evaluations.

	SSE	*df*	MSE	*f*
Between	6.8026	3	2.2675	5.0349**
Within	41.4332	92	0.4504	
Total	48.2358	95		

Note. SSE = sum of squares error; MSE = means square error. **$p < .01$.

The ANOVA results in Table 8 show that the difference between white and minority RMP and IOTA360 scores is highly statistically significant. The *f* score is 5.0349 and statistically significant at $\alpha < .01$. The mean for white RMP teaching quality scores ($\bar{X} = 4.16$) is much higher than the mean for white IOTA360 ($\bar{X} = 3.71$), minority RMP ($\bar{X} = 3.71$), and minority IOTA360 scores ($\bar{X} = 3.42$). Table 9 is another ANOVA, but it omits the white RMP scores to verify that these white RMP scores are causing the statistical difference.

Table 9. Analysis of variance of non-minority and minority IOTA360 evaluations and minority Rate My Professor results

	SSE	*df*	MSE	*f*
Between	1.3995	2	0.6997	1.69353
Within	27.2702	66	0.4132	
Total	28.6697	68		

Note. SSE = sum of squares error; MSE = means square error.

Table 9 shows that when comparing the IOTA360 results for white faculty with the RMP results of minority faculty, and IOTA360 evaluation scores for minority faculty, there is no statistical difference. The *F* score is 1.69353 and is not statistically significant at $\alpha = .05$. Therefore, it is the white RMP scores, as shown in Table 8, which cause this difference. RMP teaching quality scores show atypical favoritism toward non-minority professors.

DISCUSSION

Since the latter 20th century, faculty have perceived that student evaluations are more akin to course satisfaction surveys than evaluation of faculty competence. Students who receive low grades tend to project their academic results upon the professors' pedagogical capabilities through giving faculty poor evaluations (Aleamoni, 1999; Kulik, 2001; Lawrence, 2018; Neath, 1996). Given the legacy of discrimination and students' desire for the path of least resistance, RMP is a response to student angst that may more likely be hostile to minority faculty. In 1992, Derrick Bell (1992) discussed the permanence of racism, and CRT supports that minority faculty are more likely to be negatively evaluated. This research addresses four hypotheses to examine if RMP shows a racial bias. The first hypothesis was "RMP results will differ from IOTA360 student evaluations of faculty." The statistical results show this to be the case. The RMP scores were statistically higher than IOTA360 results in Table 1. The next issue that white professors likely to get better results on RMP evaluations and IOTA360 evaluations than minorities. To address any difference of racial bias, initially, it is important to see if minorities and non-minorities have similar training experiences. To examine this issue Hypothesis 2 states "Minority and non-minority faculty will have similar levels of educational capital and thus, perceived similarities of intellectual competence in their chosen fields." The results in Table 2 show that there is no statistically significant difference between

minority and white faculty in education capital. Next, this research examines if "high RMP teaching quality scores are associated with low RMP course difficulty scores." Table 3 shows a large statistically significant difference between course difficulty and teaching quality scores. High teaching quality scores are related to low course difficulty scores in RMP. The fourth hypothesis states, "Minority faculty will be more likely to get higher course difficulty scores and lower teaching quality scores in RMP when compared with non-minority faculty." Table 4 shows that the correlation coefficients are negative and moderately strong for both minorities and whites. The coefficients of determination show that for whites, 31% of the variability in RMP teaching quality scores are explained by knowing how people rate the difficulty of the course and 26% of the teaching quality scores can be explained by course difficulty scores for minorities. Table 5 further corroborates the difference between evaluations of teaching quality in minorities and non-minority faculty, and Table 6 does so for course difficulty race differences. Since the IOTA360 evaluations are offered to all of the students in a course, the response rate is much higher and therefore, a more accurate depiction of faculty performance than is RMP. Table 7 shows that there are no statistically significant differences between the assigned race groups. Yet, there seems to be a white privilege status in RMP teaching quality scores at the Pennsylvania college (see Tables 8 and 9).

CONCLUSION

This research uses a QuantCrit approach to examine if RMP evaluations are biased against minorities. Using a quantitative CRT approach, the results support that minority faculty are given lower teaching quality scores and higher difficulty of course scores than are non-minorities. Given the inequities embedded in the schooling system, this research analyzes if race plays a role in faculty evaluations at a college in Pennsylvania. The student rationale for the difference between white and minority results could be that minority faculty are less competent in their ability to serve students. Yet, this position is not supported. Initially, the literature supports that minorities must go through a much more rigorous process to acquire an advanced degree (King, 2016), have more difficulty getting tenure, and struggle to gain upward mobility after being hired (ASHE, 2015). In addition to the education pipeline being more rigorous for minorities, the IOTA360 results (see Table 7) show that minorities and whites have similar levels of education capital (see Table 2) and similar capabilities according to IOTA360 teaching quality student evaluations (see Tables 7 and 9). Thus,

the most rationale reason for these statistical differences is white privilege and systemic racism. In addition, the difficulty scores in RMP were significant. The persistence in low difficulty scores for white faculty suggests a bias that is not due only to the course design. Another explanation for the lower RMP competence scores for minorities is the lack of controls to deny non-students the ability to evaluate a professor. It is impossible to know who is rating the faculty and therefore, the RMP evaluations lack scientific validity.

IMPLICATIONS

Although offering important findings, the QuantCrit approach had limitations. First, the sample was small due to data limitations. Of the 50 faculty selected through purposeful sampling, two were eliminated due to a lack of RMP results. Thus, the two strata with 27 white faculty and 21 minority faculty barely approaches research standards. Secondly, the small sample size limits the external validity of these results beyond the institution in Pennsylvania. Future research could focus upon increasing the sample size of this research design. Third, the relationship between IOTA360 evaluation index questions and RMP evaluation index questions were not exactly the same and there is a possibility that participants in the evaluation tools had differing perceptions of their meanings. Although this is unlikely and they seem very similar, nevertheless there is such a possibility. Fourth, IOTA360 does not offer a course difficulty evaluation that could mimic RMP's and therefore, IOTA360 course difficulty ratings were excluded. Fifth, since race is a subjective construct, racial inequality is multidimensional and not easily measured numerically. This research minimized that factors other than race could explain the higher teaching competence scores for whites. The two factors embedded in the design that could explain competence differences are level of education and types of occupation. Teaching competence was shown to be comparable (see Table 2) and these evaluations are on tenured faculty teaching in similar fields. From a CRT theoretical purview, these factors lend support that the statistical differences in this research is likely due to the social privilege status that white faculty experience.

REFERENCES

Aleamoni, L. M. (1999). Student rating myths versus research facts from 1924 to 1998. *Journal of Personnel Evaluation in Education*, *13*, 153–166.

American Association of University Professors. (2018). What do faculty do? Washington D.C.: Author. Retrieved from https://www.aaup.org/ issues/faculty-work-workload/what-do-faculty-do

Aronson, E., & Linder, D. (1965). Gain and loss of esteem as determinants of interpersonal attractiveness. *Journal of Experimental Social Psychology*, *1*, 156–171.

Association of American Colleges and Universities. (2002). *Greater expectations: A new vision for learning as a nation goes to college*. Washington, D.C.: Author.

Baker, C. (2017). *Shades of intolerance: How capitalism and terrorism shape discrimination*. San Diego, CA: Cognella Academic Publishing.

Basow, S.A., Codos, S., & Martin J. L. (June 1, 2013). The effects of professor's race and gender on student evaluations and performance. *College Student Journal*, *47*(2), 352–363.

Bavishi, A., Madera, J. M., & Hobl, M. R. (2010). The effect of professor ethnicity and gender on student evaluations: Judged before met. *Journal of Diversity in Higher Education*, *3*, 245–256.

Bell, D. (1992). *Faces at the bottom of the well: The permanence of racism*. East Rutherford, NJ: Basic Books.

Fusani, D.S. (1994). Extra-class communication: Frequency, immediacy, self-disclosure, and satisfaction in student-faculty interaction outside the classroom. *Journal of Applied Communication Research, 22*(3), 232–255.

Garcia, N. M., Lopez, N., & Velez, V. N. (2017). QuantCrit: Rectifying quantitative methods through CRT. *Race, Ethnicity, & Education*, *21*(2), 149–157.

Germain, M. L., & Scandura, T. A. (Mar. 1, 2005). Grade inflation and student individual differences as systematic boas in faculty evaluations. *Journal of Instructional Psychology*, *32*(1) 58–67.

Griffin, W., Cohen, S. D., Berndtson, R., Burson, K. K., Camper, K. M., Chen, Y., & Smith, M. A. (Jul 16, 1994). Starting the conversation: An exploratory study of factors that influence student office hour use. *College Teaching, 62* (3), 94–99.

Henslin, J. (2007). *Sociology: A down to earth approach*. New York: Pearson.

Hiraldo, P. (2010). The role of CRT in higher education. *The Vermont Connection*, *31*, 53–59.

Hobfoll, S., & Dekel, R. (2007). The impact of resource loss on Holocaust survivors facing war on terrorism in Israel. *Aging Mental Health, 11*(2), 159–167.

Hobfoll, S., Jackson, A., & Hobfoll, I. (2002). The impact of communal-mastery versus self-mastery on emotional outcomes during stressful conditions: A prospective study of Native American women. *American Journal of Community Psychology, 50*(6), 853–871.

King, J. B., Jr., McIntosh, A., & Bell-Ellwanger, J. (2016). *The state of racial diversity in the educator workforce*. Washington, D.C.: Policy and Program Studies Service, Office of Planning, Evaluation and Policy Development, U.S. Department of Education.

Kuh, G.D., Nelson-Laird, T. F., & Umbach, P. D. (2004). Aligning faculty activities—Student behavior: Realizing the promise of greater expectations. *Liberal Education, 90*(4), 24-31.

Kulik, J. A. (2001). Student ratings: Validity, utility, and controversy. *New Directions for Institutional Research*, *109*, 9–25.

Ladner, J.A. (1973). *The death of white sociology*. Baltimore, MD.: Black Classic Press.

Ladson-Billing, G., & Tate, W. (1995). Toward a CRT of education. *Teachers College Record*, *97*(1), 47–68.

Lawrence, J. W. (May/June, 2018). Student evaluations of teaching are not valid. SET scores are poor measures of teaching effectiveness and likely undermine educational standards. *Academe*, *108*(4), 16–18.

Museus, S. D., Ledesma, M. C., & Parker, T. L. (2015). Systemic racism in higher education. *ASHE Higher Education Report*, *42*(1), 49–71.

National Center for Educational Statistics. (2013). Retrieved from https://nces.ed.gov/

National Science Foundation, National Center for Science and Engineering Statistics. (2016). Which fields attract students? In *Doctorate Recipients from U.S. Universities 2015*. Special Report NSF 17-306. Arlington, VA: Author. Retrieved from https://www.nsf.gov/statistics/2017/nsf17306/report/which-fields-attract-students/minorities.cfm

Neal, S. (2017). Views of racism as a major problem increase sharply, especially among Democrats. *Pew Research Center*. Retrieved from http://www.pewresearch.org/fact-tank/2017/08/29/views-of-racism-as-a-major-problem-increase-sharply-especially-among-democrats/

Neath, I. (1996). How to improve your teaching evaluations without improving your teaching. *Psychological Reports*, *78*, 1363–1372.

Nowell, C., Lewis, R., & Handley, B. (July, 2010). Assessing faculty performance using student evaluations of teaching in an uncontrolled setting. *Assessment & Evaluations in Higher Education*, *35*(4), 463–475.

Omi, M., & Winant, H. (1994). *Racial formation in the United States: From the 1960s to the 1990s*. Abington, UK: Routledge.

Patton S. (May 19, 2015). Student evaluations: Feared loathed and not going anywhere. *Chronicle Vitae*. Retrieved from https://chroniclevitae.com/news/1011-student-evaluations-feared-loathed-and-not-going-anywhere

Perry v. Sindermann. 408 U.S. 593 (1972). Retrieved from https://supreme.justia.com/cases/federal/us/408/593/

Poteat, P. V., & Spanierman, L. B. (2012). Modern racism attitudes among White students: The role of dominance and authoritarianism and the mediating effects of racial color-blindness. *The Journal of Social Psychology*, *152*(6), 758–774.

Sax, L. J., Gilmartin, S. K, & Bryant, A. N. (August, 2003). Assessing response rates and nonresponse bias in web and paper surveys. *Research in Higher Education*, *44*(4), 409–432.

Scriven, M. (1995). *Student ratings offer useful input to teacher evaluations*. Eric Digest/AE. Retrieved from https://eric.ed.gov/?id=ED398240

Snyder, C. R., & Claire, M. (1976). Effects of expected and obtained grades on teacher evaluation and attribution performance. *Journal of Educational Psychology*, *68*, 75–82.

Stephan, W., Renfro, L., Duran, A., & Clason, D. (2006). The role of threat in attitudes toward affirmative action and its beneficiaries. *Journal of Applied Social Psychology, 36*(1), 41–74.

Stephan, W. G., Ybarra, O., & Morrison, K. R. (2009). Intergroup threat theory. In T. D. Nelson (Ed.), *Handbook of prejudice, stereotyping and discrimination* (pp. 43–59). New York: Psychology Press.

Teranishi, R. T. (2007). Race, ethnicity, and higher education policy: The use of critical quantitative research. *New Directions for Institutional Research, 2007*(133), 47–59.

Vescio, T. K., & Biernat, M. (1999). When stereotype-based expectancies impair performance: The effect of prejudice, race, and target quality on judgments and perceiver performance. *European Journal of Social Psychology*, *29*, 961–969.

Zuberi, T. (2001). *Thicker than blood: How racial statistics lie*. Minneapolis: University of Minnesota Press.

DR. CHUCK A. BAKER is a tenured professor of sociology at Delaware County Community College. He teaches diversity courses and acknowledges the importance of differing perceptions among people. He has worked to foster intellectual diversity and pushes classroom dialogue into areas beyond the traditional topics of discrimination, promoting conversation independent of emotional talking points. Email: CBaker2@dccc.edu