Original Article:

STUDENT EVALUATIONS OF INSTRUCTOR WARMTH AND COMPETENCE: COURSE DIFFICULTY COUNTS MORE THAN CHARACTER

Ivan Carbajal, B.A.
University of North Texas, USA

Jamie S. Hughes, Ph.D.
University of Texas of the Permian Basin, USA

Abstract

Subjective rating sites, such as ratemyprofessors.com, depend on descriptive characteristics, stereotype expectations, and grade expectations. Here, we used the stereotype content model and grading leniency hypothesis to examine student decisions to enroll in a class. In Study 1, participants judged a male or female professor described as cold or warm. In Study 2, participants judged a male or female professor whose course was described as difficult or easy. Supporting the leniency hypothesis student ratings of professors were highly influenced by course difficulty. Students were more likely to enroll in easy courses.

AUTHOR NOTE: Please address all correspondence to Dr. Jamie S. Hughes, The University of Texas of the Permian Basin, Psychology Department, 4901 E. University Blvd., Odessa, TX 79762, U.S.A. Email: hughes_j@utpb.edu

INTRODUCTION

Student evaluations are important for any teacher or professor. Evaluations can determine promotions or reprimands, lesson plan adjustments, and self-reflections. Research conducted on student evaluations typically examines characteristics of a successful professor, such as their reputation or personal characteristics. Previous research indicates that first impressions can be more important in evaluations than a professor's reputation (Buchert, Laws, Apperson, & Bregman, 2008). A crucial part in evaluating professors is rating the type of characteristics they display. Research typically describes warm characteristics in a professor by how approachable, available, and competent they are. Characteristics that are related to warmth, such as approachability and availability, are stereotypically categorized as feminine characteristics (Kelley, 1950; Widmeyer & Loy, 1988). Cold characteristics on the other hand, such as being assertive, dominating, and critical are seen as more masculine.

Solomon Asch's (1946) work on first impressions was the first to examine cold and warm characteristics and their effects on impressions. In Asch's (1946) study, he found warmth and coldness were essential in establishing a unified impression of a person. Upon encountering a person for the first time (or learning about a person's traits as in Asch's study) one gathers information about them and forms an impression, or an overall mental picture of what a person's characteristics are. In his work, Asch (1946) found first impressions could be altered by describing a person as either warm or cold.

The stereotype content model, developed by Fiske, Cuddy, and Glick (2006), shows the dynamic relationship between warmth and competence, and the importance of these two personality dimensions when forming impressions. Warmth is defined through characteristics such as friendliness, trustworthiness, and morality. Competence is shown through intelligence, efficacy, and skill. Warmth and competence are important when perceiving others because of its usefulness in survival. As an evolutionary adaptation, first impressions served as a tool that helped our ancestors distinguish friends from foes. Although we do not have the same concerns as our ancestors now, these two dimensions help perceivers distinguish between individuals they like and respect and those they think will harm them. The most noticeable dimension is warmth. People are more sensitive to warmth than they are to competence. Warmth is easily discerned by others, especially by women, due to traditional female gender roles that value warmth over competence (Abele, 2003; Fiske et al., 2006). Fiske et al., (2006) suggested there is an axis wherein people can be seen on a continuum from warm to cold and a continuum from competent to incompetent. The axis makes it possible for people to be categorized based on these characteristics. For example, people can be described as highly competent but low on warmth, highly warm and competent, or low in competence but high on warmth. Fiske et al.'s (2006) work involves examining judgments of those with different status positions or positions of power and their perceived warmth and competence. Social group members are often stereotyped in terms of these dimensions. Upon seeing a homeless person there is a

tendency to perceive the person as low on warmth and competence. In contrast, housewives are perceived to be high in warmth and relatively competent. This work implies that expectations about warmth and competence depend on a person's social position or role in society.

Warmth and competence stereotypes may be combined to predict behavior (Fiske et al., 2006). For example, when meeting a professor for the first time students may evaluate his or her competence and warmth and use these impressions to predict what will happen in the classroom (e.g., warm or cold, competent or incompetent). These expectations may bias or color evaluation of the professor. When people lack warmth and competence they are not viewed favorably. Further, gender stereotypes color our expectations of warmth and competence. For example, work by Bennett (1982) showed that female professors were negatively evaluated when they failed to meet the stereotypical gender appropriate characteristic of being warm. Male professors on the other hand, are denigrated for incompetence but not for lack of warmth. If a male is cold his behavior is seen as gender appropriate (i.e., he may be seen as appropriately authoritarian. However, if he is warm he is not punished for it in evaluations. On the contrary, men who exhibit warmth are evaluated positively (Bennet, 1982).

Women in typically male-dominated roles, such as academia, also experience what is known as sex role spillover, in which women are expected to be more nurturing than men in the same work roles (Gutek & Morasch, 1982). Gutek and Morasch (1982) showed that people are typically identified through their gender first in professions. For example, a female professor is seen as a female first, then as a professor, meaning being a female has its own expectations (e.g., such as being friendly, nurturing, and warm) apart from expectations of being a professor (i.e., being competent). Not only are female professionals supposed to be competent in their field, they are also expected to adhere to stereotypical gender roles. Further, men are more comfortable around women who fulfill a nurturing role (Gutek & Morasch, 1982). Men are also expected to adhere to their respective gender roles, and are judged in terms of competence rather than warmth. That is, men are expected to be aggressive and dominating in the workplace, but unlike their feminine counterparts, it is not necessary for males to be warm.

Gender differences are found in research on students evaluations of their professors (Bennett, 1982; Patrick, 2011; Remedios & Lieberman, 2008; Vaillancourt, 2013). Consistent with the stereotype, some research suggests that men are seen as more effective and competent than women overall (Sidanius & Crane, 2006).While both female and male professors are seen as intellectual, women are rated higher on interpersonal aspects of instruction (i.e., availability and approachability) than men. Further, if a female professor lacks warmth she is perceived more negatively by her students than male professors (Bennett, 1982). Centra and Gaubatz (2000) found that female students viewed female professors as better organized, better at communication, more interactive, and better providers of feedback than males. There is a bias in favor of female professors by female

students because of their teaching style. This may be due to the finding that female professors engage in more discussions, while male professors typically lecture (Centra & Gaubatz, 2000). Allowing discussion and interacting with the students may reflect warmth and increase the perception that the professor is approachable and caring. However, Bennett (1982) found that students demanded higher standards of formal preparation and organization from female professors than they did for male professors, and this may be because of a need to prove competence on the female instructor's part.

Research on warmth shows it correlates positively with student rating of teacher performance (Bennett, 1982). Typically a student's rating of an instructor as warm means being perceived as more supportive and interested in others. Warmth also increases the likelihood of engagement in class discussions from students (Buchert et al., 2008; Kelley, 1950). Warm individuals are well liked because warmth carries more weight in affective and behavioral reactions, and warm individuals are better remembered (Fiske et al., 2006). The work by Best and Addison (2000) showed that warmth can be primed using comments about the professor that suggest warm traits, such as an interest in learning names or via stating aloud they care about their students. Students formed a general impression of the professor with these descriptions, just as Asch was able to in his study with the words cold and warm.

Kindred and Mohammed (2005) explored teacher ratings on ratemyprofessors.com and found that students used the ratings as guides. Student who used ratemyprofessors.com found student opinions and information about teaching quality beneficial as it aided in their course selection. Students indicated a strong preference for selecting a course based on written comments over numerical ratings. Further, students in the study believed competence, knowledge, clarity, and helpfulness were more important attributes for a professor compared to appearance, race, or gender. Finally, Kindred and Mohammed (2005) found a strong positive relationship between easiness and quality ratings, meaning that as easiness ratings increased so did the professor's overall rating.

A controversial topic in the field is whether grades influence student evaluation (Greenwald & Gilmore, 1997; cf. Marsh & Roche, 1997). Jewell and McPherson's (2012) research suggests that female professors are more likely to inflate grades and therefore receive higher student evaluations. However, according to Vaillancourt (2013), grade inflation increases ratings for both genders. Vaillancourt tested the hypothesis that female professors who gave negative feedback would be rated lower than male professors who did the same. Surprisingly, gender had no effect, and grades had a substantial effect on ratings.

The work on student evaluations is important to instructors in higher education because faculty promotion and tenure decisions are influenced by student ratings. Here, we were interested in how the ease of a class and gender (features over which instructors have little control over), and warmth and competence (characteristic over which instructors have some control) affect positive perceptions of instructors. Although many faculty would prefer to be judged based on the effectiveness of their courses in helping students learn,

faculty promotion decisions rest on student's inferences about warmth, competence and/or their biases about gender, and their subjective impressions of course difficulty. As we merge into a digital age, in which students evaluate instructors anonymously online, and that content becomes freely available, it is important to evaluate the factors that influence positive and negative evaluations. Instructor ratings may be used to evaluate instructors for promotion and tenure decision, and to make decisions about job continuation and scheduling. As the politician and others increasingly look to institutions of higher education to be more and more business oriented, instructor evaluations may be relied on more to facilitate and validate employment decisions. However, as with any decision making tool it is important to understand exactly what the data imply. Here, we examined how factors such as gender, warmth and competence, and course difficulty influences university student's ratings of their instructors. Furthermore, it is important to understand how students are using technology and instructor evaluations to make decisions about their education. Are students interested in identifying competent professors or those who will give them an easy A?

We identified two gaps in the literature. First, it is not clear whether a professor's personality characteristics or grading leniency influence student evaluations. It is also not clear how gender stereotypes influence impressions. Based on the stereotype content model, student evaluations of female professors compared to male professors should be particularly negative when a professor lacks warmth. Further, student evaluations of male professors should be negative when they lack competence. In contrast, the grading leniency hypothesis suggests that perceptions of the ease or difficulty of a class should influence evaluations to a greater extent than gender. To test these competing hypotheses, we conducted two studies. Student participants were asked to judge professors based on descriptions akin to comments on the website *www.ratemyprofessors.com*. Participants in Study 1 read characteristics describing a cold or warm, male or female professor. In Study 2, male and female professors' courses were described as difficult or easy. In both studies, we asked students to rate the likelihood that they would take a class with the professor described. These course likelihood ratings were used to gauge student evaluations.

* STUDY 1 *

In Study 1, we tested the hypothesis that gender stereotypes would affect perceptions of professors and increase the likelihood that students would prefer taking a class with an unfriendly male instructor compared to an unfriendly female instructor. Participants were asked to judge a male or female instructor described as warm or cold. We measured perceptions of competence and warmth and asked students to rate the likelihood that they would take a class with each of the target professors.

METHOD

Participants

After removing data from participants who were not students or did not complete the study ($N = 36$), 104 students (80 women) remained. We recruited participants for the online study via the university psychology subject pool ($N = 45$) and from Psychological Research on the Net ($N = 59$), a crowd sourcing site for researchers. Students from the university subject pool were slightly older on average ($M = 27.47$; due to a number of non-traditional students) than those from the crowd sourcing site ($M = 22.47$), but no other differences between samples emerged. Participants were 25 years of age on average ($SD = 8.83$), and the majority were Caucasian or White (47%), Hispanic or Latino (29%), African-American or Black (9%), mixed race (9%), or Asian, Indian or Middle Eastern (6%).

Materials and Procedures

Participants were directed to the Survey Monkey site to complete a study about the kinds of information students use to evaluate professors. They were asked to read eight statements and try to form an impression of a professor. The statements represented comments left by other students on sites like "ratemyprofessors.com." Participants were randomly assigned to read a description of a professor who was cold or warm, and male or female in a 2(Gender: male vs. female) x 2(Characteristics: warm vs. cold) between-participant design. Gender was manipulated via the professors names and the pronouns used in the vignette. In the cold characteristic conditions participants read student feedback such as, "You can tell he/she enjoys history a lot with his/her non-stop lectures. During office hours he/she is usually on his/her computer or reading, doesn't seem to really like to have students around, seems socially awkward." In the warm characteristic conditions, participants read student feedback such as, "Great professor, you can tell he/she cares about his/her students and is always more than happy to help and answer questions in class."

Following presentation of the feedback, students were asked to rate several items related to the construct of warmth (i.e., they rated how friendly, warm, and caring the professor was), and the construct of competence (i.e., how intelligent, competent, and smart the professor was). Finally, they were asked, "How likely would it be that you would take a class with this professor?" Items were rated on 7 point scale ranging from 1 (*Not Very*) to 7 (*Very*). Finally, participants were asked whether they would like to share any other comments about the professor and completed demographic questions.

RESULTS

Competence, intelligence, and smart ratings were combined and averaged to create a competence scale (α = .82). Further, friendly, warm, and caring rating were combined to create a warmth scale (α = .98). Both scales were found to be internally consistent. A factorial ANOVA with Gender and Characteristics was used to test the hypotheses regarding the effect of instructor gender and warm or cold characteristics on perceptions of competence, warmth, and likelihood to take a class. The competence scale was examined first. The main effect of Characteristics, $F(1, 100) = 7.78$, $p = .006$, $\eta^2_p = .07$, indicated that those who were described as warm were seen as more competent ($M = 6.38$, $SD = 0.89$) than those described as cold ($M = 5.85$, $SD = 1.11$). The main effect of Gender approached significance, $F(1, 100) = 7.78$, $p = .06$, $\eta^2_p = .03$. Contrary to the stereotype content model, female professors were rated as more competent ($M = 6.29$, $SD = 0.92$) than male professors ($M = 5.92$, $SD = 1.14$). The interaction effect did not reach significance, $F(1, 100) = 0.31$, $p = .58$.

Inferences about warmth followed a similar trend (see Table 1). Unsurprisingly, the main effect of Characteristics indicated that those who were described as warm were seen as more warm ($M = 6.49$, $SD = 1.03$) than those described as cold ($M = 2.70$, $SD = 1.67$), $F(1, 100) = 187.08$, $p < .001$, $\eta^2_p = .65$. There was no effect of Gender, $F(1, 100) = 0.23$, $p = .63$, or an interaction effect, $F(1, 100) = 0.01$, $p = .91$.

Table 1. **Descriptive Statistics by Condition for Study 1**

Gender & Characteristics	Study 1 Competence	Warmth	Likelihood to Enroll in Class
Female			
Warm	6.50 (0.81)	6.57 (0.98)	6.20 (1.35)
Cold	6.06 (1.00)	2.79 (1.61)	3.78 (1.79)
Male			
Warm	6.24 (1.00)	6.40 (1.11)	6.25 (1.42)
Cold	5.58 (1.21)	2.68 (1.78)	2.95 (2.01)

Note: Standard deviations are listed in parentheses. Significant main effects of factors are displayed in text above. No interaction effects were significant.

Student ratings of the likelihood of taking a class with the professor also showed a main effect of Characteristics. Students were more likely to take a class with professors described with warm characteristics ($M = 6.22$, $SD = 1.37$) than cold characteristics ($M = 3.42$, $SD = 1.92$), $F(1, 100) = 77.04$, $p < .001$, $\eta^2_p = .43$. The main effect of Gender, $F(1,$

$100) = 1.44$, $p = .23$, and the interaction effect were not significant, $F(1, 100) = 1.83$, $p = .18$.

Twenty five students provided open-ended comments. Most wrote that it was important for their professors to be warm ($N = 11$). However, some equated warmth with lack of difficulty in the course ($N = 2$). Further, two students stated that they wanted additional information about how difficult or easy the course was to make an informed decision.

DISCUSSION: STUDY 1

In Study 1, we did not find evidence in favor of the gender stereotype content model. That is, students were not less likely to take a class from a cold female professor than a cold male professor. Further, competence ratings were greater for professors with warm compared to cold characteristics, and for women compared to men. This trend implies that competence ratings were affected by judgment about warmth. In contrast to the stereotype content model, the evidence appears to favor an alternative explanation whereby students simply prefer a warm compared to cold professor. The data suggest that those who were well-liked due to their warmth were seen as more competent and were strongly preferred by students. Although more female than male students participated in our study, we did not evaluate the gender of the perceiver for two reasons. First, Fiske et al.'s (2006) work suggests that sexist beliefs are shared by male and females approximately equally. Second, we had no specific hypotheses regarding gender of the perceiver on our dependent measures.

* STUDY 2 *

In Study 1, contrary to expectation, female professors were seen as more competent than male professors. Further, the decision to take a class was affected by the characteristics of the professor. Drawing from the grading leniency hypothesis, however, we predicted that students were more likely to take a class with a warm professor because he or she was perceived to be more grading lenient. In other words, we hypothesized that course difficulty would influences one's likelihood to take a class and that easy classes would be preferred. However, in line with the stereotype content model, a target person's gender may still influence judgments. To test these hypotheses, a male or female professor was described as having a difficult or easy class. According to the stereotype content model, women who teach difficult classes should be punished more (be seen as less warm, and less competent) than men who teach difficult classes. However, there should be no differences between male and female professor who teach easy classes. In contrast, the grading leniency

hypothesis suggests a main effect of difficulty whereby easy courses are preferred regardless of gender. We reasoned that those with a low GPA may be especially likely to provide positive ratings of a professor who taught an easy class. To examine this possibility, we measured students GPA to remove the variance associated with academic achievement.

METHOD

Participants

As in Study 1, participants were recruited online from the university psychology subject pool ($N = 43$) as well as the Psychological Research on the Net site ($N = 11$). As in Study 1, the university subject pool was older on average (M = 31.07) than those recruited from the online site ($M = 22.55$), but no other differences emerged. After removing participants who indicated that they were not students ($N = 10$), our sample consisted of 54 students (45 female) who were 29 years of age on average ($SD = 9.71$). Participants were Caucasian or White (70%), Hispanic or Latino (20%) or Black or African American, Asian, or mixed race (10%). Further, the average GPA was 3.38 on a 4 point scale ($SD = 0.51$).

Materials and Procedures

Participants were directed to the Survey Monkey website to complete a study about what kinds of information students use to choose classes. Participants were asked to read eight statements and try to form an impression of a professor. The statements represented comments left by other students on sites like "ratemyprofessors.com." Gender was manipulated via professors names and the pronouns used in the statements. In the difficult course condition, participants read statements such as, "Too much homework and lots of tests. Lucky enough to have gotten a C-." In the easy course condition statements were listed such as, "Really easy class, easiest A I have ever gotten." In sum, a 2(Gender: male vs. female) x 2(Difficulty: easy vs. difficulty) between-participant design was used.

As in Study 1, participants rated how warm, friendly, and caring the professor was. They also rated how competent, smart, and intelligent they viewed the professor. They indicated the likelihood that they would take a class from the professor and were asked if they had any comments they would like to share about their impression of the professor. Ratings were made on 7 point scales ranging from 1 (*not very*) to 7 (*very*).They also completed demographic questions which included a question about their GPA.

RESULTS

Warmth and competence scales were created by combining and averaging the three warmth ($\alpha = .92$) and three competence items ($\alpha = .95$). Each analysis was performed with participant GPA as a covariate. However, GPA was not significant and did not diminish the effect of the independent variables. Therefore analyses are reported without GPA as a covariate. Factorial ANOVAs with Difficulty and Gender were used to analyze the data. Attributions about competence were analyzed first. A main effect of Difficulty emerged, $F(1, 50) = 13.02$, $p = .001$, $\eta^2_p = .21$, indicating that professors who taught easy courses were seen as less competent ($M = 3.66$, $SD = 1.62$) than professors who taught difficult courses ($M = 5.16$, $SD = 1.31$). The main effect of Gender, $F(1, 50) = .12$, $p = .73$, and interaction effect were not significant, $F(1, 50) = .07$, $p = .80$.

Inferences about warmth followed a different pattern (see Table 2). The main effect of Difficulty, $F(1, 50) = 41.01$, $p < .001$, $\eta^2_p = .45$, indicated that professors who taught easy courses were seen as possessing more warmth ($M = 5.18$, $SD = 1.43$) than those teaching difficult courses ($M = 2.59$, $SD = 1.44$). Again, the main effect of Gender, $F(1, 50) = .04$, $p = .85$, and interaction effect were not significant, $F(1, 50) = .04$, $p = .83$.

Student course preferences supported the grade leniency hypothesis and showed a main effect of Difficulty, $F(1, 50) = 16.42$, $p < .001$, $\eta^2_p = .25$. Students were more likely to take a class that was described as easy ($M = 5.10$, $SD = 1.18$) than difficult ($M = 2.72$, $SD = 1.86$). There was no main effect of Gender, $F(1, 50) = .53$, $p = .47$, or interaction effect, $F(1, 50) = .32$, $p = .57$. As stated above, these effects persisted when GPA was used as a covariate. In other words, student GPA did not influence warmth or competence ratings or one's likelihood to take a class.

Table 2. Descriptive Statistics by Condition for Study 2

Gender & Course Difficulty	Study 2		
	Competence	Warmth	Likelihood to Enroll in Class
Female			
Easy	3.77 (1.49)	5.19 (1.50)	5.44 (2.10)
Difficult	5.18 (1.32)	2.48 (1.87)	2.78 (2.05)
Male			
Easy	3.51 (1.81)	5.18 (1.39)	4.69 (2.29)
Difficult	5.14 (1.34)	2.64 (1.20)	2.69 (1.81)

Note: Standard deviations are listed in parentheses. Significant main effects of factors are displayed in text above. No interaction effects were significant.

Nineteen students provided open-ended comments. An evaluation of the content of comments revealed a number of student concerns. Four students suggested that because grades matter to those outside of the academy more than one's actual growth or intellectual development, it makes sense to take a class where one can earn an easy "A." One respondent suggested that an easy grade does not mean one may not learn, and others remarked that making classes unnecessarily difficult demonstrates a professor's lack of consideration for students ($N = 3$). A few students stated that they would not want an easy A because they would not feel challenged ($N = 3$). However, average ratings suggest that this sentiment was not reflected in their likelihood to take difficult classes.

DISCUSSION: STUDY 2

In contrast to the hypotheses drawn from the stereotype content model, student ratings of professors were not influenced by gender. However, supporting the leniency hypothesis, student ratings of their professors were highly influenced by course difficulty. Students rated those with easy courses as less competent, but warmer than those with difficult courses. Further, students preferred easy courses to difficult ones. Although students recognized that a difficult professor may be more competent, they believed that a professor with an easy course would be more caring, warm, and friendly and thus they preferred the easy course to the difficult one. Further, students' academic achievement (e.g. GPA) did not alter this pattern.

GENERAL DISCUSSION

When it comes to choosing a class, our data suggest that students are more likely to choose an easy course and warm professor compared to a difficult course or cold professor. Surprisingly, Study 2 showed that while easy professors were seen as less competent, they were still preferred, in part, due to their perceived warmth. While we did not examine professor evaluations directly, the likelihood of taking a course showed that ease of one's course, and inferences about warmth drove ratings more than gender. Thus, contrary to the stereotype content model tested in Study 1, we did not find effects of gender. However, warm professors were perceived as more competent and preferred over their cold counterparts. Study 2 provided support for the grade leniency hypothesis. Students' likelihood to enroll in a course was highly influenced by course difficulty. Students preferred a less competent professor who taught an easy course compared to a more competent professor who taught a difficult course (e.g., competence ratings were high the difficult course conditions and lower in the easy course conditions).

The most intriguing finding comes from the differences in perceptions of competence in Study 1 and 2. In Study 1 a warm professor was seen as more competent than a cold professor, while in Study 2 a professor with a difficult course was seen as more competent, and less warm than one with an easy course. Study 1 results may have stemmed from student schemas that suggest that if one is warm, then one is also likely to have other positive characteristics such as competence. In contrast, a difficult course may be taught by a person who is extremely competent, but unable to take students lack of exposure to the material into account while teaching. More research is needed to explore student schemas and implicit personality theories regarding professors by course difficulty.

Course difficulty decreased the likelihood that a student would take a class, and also influenced perceptions of competence. Easy courses, were perceived to be taught by relatively incompetent professors, but were preferred anyway. These results are startling because ideally, competence should matter a great deal to students. That is, faculty in higher education are judged by their peers, colleagues, and administrators according to their competence. However, to win over students one need only be warm and provide easy courses. While one's warmth can win the approval of those he or she works with, it is not often thought of as essential to perform one's job in academia. The data reported here suggest otherwise. If students have voice, and influence administrator's decision making about their employees in higher education, this research suggests that warmth is necessary, at least, if one wishes to earn positive evaluations or have students enroll in one's classes. Unfortunately, the discrepancy between positive reviews and difficult courses (those with difficult course were seen as less warm) could prove problematic for those interested in providing challenging and difficult course work. That is, it is not clear how faculty in higher education can excel both in winning student approval via high student evaluations while at the same time providing challenging course work that is likely to make students struggle, and therefore, grow intellectually.

Our findings are consistent with Patrick's (2011) work on professor evaluations. Patrick showed that while one's expected grade in a course did not influence professor ratings, it influences one's rating of the course. Those who expected worse grades evaluated the course more negatively than those who expected good grades. The data reported here also showed that difficult courses were not highly desired, however, it also suggests that course difficulty may generalize to inferences about an instructor's competence and warmth. More research is needed to examine the relationship between course difficulty and professor evaluations.

Remedios and Lieberman (2008) showed that mastery oriented students, or those who find learning for learning sake important, evaluate classes in terms of how stimulating, interesting, and useful they are. This can be compared to achievement-oriented students who are more concerned with course outcomes (i.e., grades). In their study, grades were less important in student evaluations of courses. Further, for mastery oriented students, data suggest that an earned grade would be more important than an easy grade. We did not

measure mastery orientation, and thus our findings may have stemmed from an achievement-oriented sample since students focused more on outcomes (e.g. grades). However, we doubt this interpretation. Instead, our data may differ because unlike many studies of student impressions, we asked students for their impressions about fictional classes. It is possible that students respond differently to hypothetical courses, or courses they have not taken, than to classes for which they are currently enrolled. For example, students may believe easy classes are more desirable, but feel differently about courses they enjoy, courses in their major, or courses for which they are currently enrolled. We hope future researchers investigate these possibilities by seeing which variables account for the most variance in course rating and professor ratings.

Our data suggest that stereotypes about gender did not influence students' course preferences. This finding aligns with Vaillaincourt's (2013) work in which gender had no effect on the ratings of female professors who gave negative feedback to students. That is, female professors in Vaillaincourt's (2013) study were not punished for deviating from the gender stereotype of warmth. It is possible that in educational settings women are already conforming to traditional notions of femininity simply by teaching, and are therefore given more latitude to deviate from other stereotypical norms. Vaillaincourt (2013) has gone further, suggesting that gender stereotypes in academia is a dated idea. We possess a more cautious perspective and believe more research is needed to detect subtle evaluative differences in how men and women in academia are perceived and rewarded by both students and their peers.

A big question that looms after this study is "So what"? As Kindred and Mohammed (2005) showed, students care more for written comments left on these professor evaluation sites rather than the numerical rating. Taking this information and applying it to our results means comments from a past student such as "easy professor, easy A" can be more persuasive than a comment such as "tough professor but learned a lot," when it comes to course selection time. The same can be applied to comments that range from "really caring professor" to "really cantankerous professor," with warmth characteristics having a larger impact on enrollment plans. While some might view comments left on these websites as either extremely vengeful or extremely positive, they do matter when trying to form an impression of a future professor.

As a final note, this research points to the importance of examining the types of questions that are asked of students in instructor evaluation forms. The data reported here indicate that there may be a discrepancy between what students desire (i.e., they want to feel good about the educational experience via easy assignments and course expectations) and what educators and the public find desirable in higher education (i.e., a challenging educational environment that promotes and enhances the best and brightest). If this assessment is accurate, using student evaluations to make employment decisions focuses on the consumer rather than the goals of the education institution as a whole and thus may be misguided.

As with much research on social judgments, our studies used vignettes which may limit the generalizability. However, we believe our methods were ecologically valid in that many students take advantage of posted professor ratings to make course decisions. Students in our studies only made impressions and judged the likelihood of taking a class from a professor. Thus, it is not clear whether our results would generalize to end-of-the-semester course evaluations. Future research should explore this possibility. Researchers should also examine whether gender and course difficulty effects are more pronounced for online compared to face to face courses. We suspect that for online students, gender and the personality characteristics associated with a professor matter even less, and that course difficulty matters more.

The implications of our study have significant impact on academic culture and what it means to be a competent teacher. Academia relies on student evaluations to make decisions about current professors and graduate students readying themselves for a job in academia. If students are seeking only those professors who give out "easy A's," and administrator's reinforce this behavior via more positive performance reviews of easy versus difficult faculty this has the power to transform our institutions and decrease their educational impact. If college is only about getting the A's necessary to get a bachelor's degree then this signifies the failure on any institution of providing an adequate undergraduate education experience.

REFERENCES

Abele, A. E. (2003). The dynamics of masculine-agentic and feminine-communal traits: Findings from a prospective study. *Journal Of Personality And Social Psychology*, *85*, 768-776.

Asch, A. E. (1946). Forming impressions of personality. *The Journal of Abnormal And Social Psychology*, *41*, 258-290.

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.

Best, J. B., & Addison, W. E. (2000). A preliminary study of perceived warmth of professor and student evaluations. *Teaching of Psychology*, *27*, 60-62.

Buchert, S., Laws, E. L., Apperson, J. M., & Bregman, N. J. (2008). First impressions and professor reputation: Influence on student evaluation of instruction. *Social Psychology of Education*, *11*, 397-408.

Centra, J. A., & Gaubatz, N. B. (2000). Is there gender bias in student evaluations of teaching? *Journal of Higher Education*, *70*, 17-33.

Fiske, T. S., Cuddy, J. C. A., & Glick, P. (2006). Universal dimensions of social cognition: Warmth and competence. *TRENDS in Cognitive Sciences*, *11*, 77-83. doi:10.1016/j.tics.2006.11.005

Greenwald, A. G., & Gillmore, G. M. (1997). Grading leniency is a removable contaminant of student ratings. *American Psychologist, 52,* 1209–1217.

Gutek, B. A., & Morasch, B. (1982). Sex ratios, sex-role spillover, and sexual harassment of women at work. *Journal of Social Issues*, *38*, 55-74.

Jewell, T. R., & McPherson, M. A. (2012). Instructor-specific grade inflation: Incentives, gender, and ethnicity. *Social Science Quarterly*, *39*, 95-109. doi: 10.1111/j.1540-6237.2011.00827.x

Kelley, H. H. (1950). The warm-cold variable in first impressions of persons. *Journal Of Personality*, *18*, 431-439.

Kindred, J., & Mohammed, S. N. (2005). "He will crush you like an academic ninja!": Exploring teacher ratings on Ratemyprofessors.com. Journal of Computer Mediated Communication, 10, 00. doi:10.1111/j.1083-6101.2005.tb00257.x

Marsh, H. W., & Roche, L. A. (1997). Making students' evaluations of teaching effectiveness effective. *American Psychologist, 52,* 1187-1197.

Patrick, C. L. (2011). Student evaluations of teaching: Effects of the Big Five personality traits, grades and the validity hypothesis. *Assessment & Evaluation in Higher Education*, *36*, 239-249. doi: 10.1080/02602930903308258

Remedios, R., & Lieberman, D. A. (2008). I liked you course because you taught me well: the influence of grades, workload, expectations and goals on students' evaluations of teaching. *British Educational Research Journal*, *34*, 91-115. doi: 10.1080/01411920701492043

Sidanius, J., & Crane, M. (2006). Job evaluations and gender: The case of university faculty. *Journal of Applied Social Psychology, 19,* 174-197. doi: 10.1111/j.1559-1816.1989.tb00051.x

Vaillancourt, T. (2013). Students aggress against professors in reaction to receiving poor grades: An effect moderated by student narcissism and self-esteem. *Aggressive Behavior*, *39*, 71-84. doi: 10.1002/ab.21450

Widmeyer, W. N., & Loy, J. W. (1988). When you're hot, you're hot! Warm-cold effects in first impressions of persons and teaching effectiveness. *Journal of Educational Psychology*, *80*, 118-121.

AUTHOR INFORMATION:

Ivan Carbajal is a doctoral student in Behavioral Science at the University of North Texas. His current research interests include psychophysiology, specifically pertaining to circadian match/mismatch influences on effort intensity, terror management theory applied to behavioral restraint and implicit attitudes, and minority wellness research. Address: Ivan Carbajal, Department of Psychology, 1611 W Mulberry St #171, Denton, TX 76201, USA. Email: IvanCarbajal@my.unt.edu

Jamie S. Hughes is an Associate Professor at The University of Texas of the Permian Basin. She received her Ph.D. in Social Psychology at New Mexico State University in 2010. Her research interests include moral psychology, person perception, and social issues. Address: Dr. Jamie S. Hughes, 4901 E. University Blvd., Odessa, TX 79701. Email: hughes_j@utpb.edu