

Students' Perceptions of Expressiveness: Age and Gender Effects on Teacher Evaluations

Julianne Arbuckle^{1,2} and Benne D. Williams¹

In this study we investigated the relationship between college students' perceptions of professors' expressiveness and implicit age and gender stereotypes. Three hundred and fifty-two male and female students watched slides of an age- and gender-neutral stick figure and listened to a neutral voice presenting a lecture, and then evaluated it on teacher evaluation forms that indicated 1 of 4 different age and gender conditions (male, female, "old," and "young"). Main and interaction effects indicated that students rated the "young" male professor higher than they did the "young" female, "old" male, and "old" female professors on speaking enthusiastically and using a meaningful voice tone during the class lecture regardless of the identical manner in which the material was presented. Implications of biased teacher-expressiveness items on student evaluations are discussed.

KEY WORDS: student evaluations; teacher expressiveness; students' stereotypes.

Research indicates that certain personality characteristics influence student evaluations of college professors. From the students' points of view, teacher-expressive characteristics such as warmth, enthusiasm, and extroversion apparently separate effective from ineffective teachers (Basow, 2000; Basow & Silberg, 1987; Best & Addison, 2000; Bousfield, 1940; Cravens, 1996; Feldman, 1986; Guerrero & Miller, 1998; Marsh & Roche, 1997; Radmacher & Martin, 2001). Consequently, college faculty who may wish to increase their student ratings on certain items on the evaluation form (Williams & Ceci, 1997; Wilson, 1998), and perhaps improve their teaching effectiveness, could upgrade their presentation styles with warmth, enthusiasm, and extroversion (Leblanc, 1998; Upham, 2001). However, the research does not clearly indicate specifically who can or cannot benefit from displaying these necessary characteristics, particularly when age or gender variables are isolated and examined (Basow, 1990; Basow & Distenfeld,

1985; Basow & Silberg, 1987; Bennett, 1982; Kalavar, 2001; Perdue & Gurtman, 1990; Peterson, 1980). A substantial issue to explore is the degree to which students' perceptions of a professor's age and gender influence their perceptions of the professor's warmth and enthusiasm.

The purpose of this study was to determine whether the perceptions of college professors' expressiveness in the classroom are strongly associated with students' implicit attitudes toward age and gender. Although the student evaluation may disguise flagrant age and gender bias, there is evidence that college students rate male and female faculty according to subtle culturally conditioned age and gender stereotypes (Bennett, 1982; Kite, 2001). Centra and Gaubatz (2000, p. 31) reported that "bias occurs when characteristics such as age and gender impact evaluations systematically but do not impact student learning." Age and gender characteristics may not impact student learning (Abrami & d'Apollonia, 1999; d'Apollonia & Abrami, 1997), but they can impose an undesirable, costly effect on unfairly evaluated faculty members and their subsequent tenure, promotion, salary, retirement, and quality of life (Basow, 1998).

¹Department of Psychology, Weber State University, Ogden, Utah.

²To whom correspondence should be addressed at Department of Psychology, 1202 University Circle, Weber State University, Ogden, Utah 84408-1202; e-mail: jarbuckle@weber.edu.

Despite a large contemporary literature on student evaluations, age effects are infrequently mentioned, and the available studies on gender effects have been controversial and contradictory (Basow, 1998, 2000; Freeman, 1994; Marsh & Roche, 1997). Many researchers, mindful of the controversy, have concluded that most evaluation instruments are multidimensional and result in global scores that can be misleading and difficult to interpret (Barnes & Barnes, 1993; Greenwald, 1997; Marsh & Roche, 1997; McKeachie, 1997). But studies that focus on subjective teacher-expressive dimensions of student evaluations may illuminate why male and female faculty who teach the same courses are rated so differently even when they are matched in rank and teaching experience (Basow, 1998, 2000; d'Apollonia & Abrami, 1997; Greenwald & Gillmore, 1997). Furthermore, parsing and analyzing particular items in a multidimensional evaluation instrument can pinpoint potential biases in the evaluation instrument (Wilson, 1998), or, more notably, the implicit biases of the student evaluators (Basow, 1998, 2000; Unger, 1979).

Differences detected in male and female professors' evaluations suggest that gender stereotyping by college students results, at times, in more favorable evaluations of women than men, and, at other times, in more favorable evaluations of men than of women. The differences seem to occur more often on certain subjective, teacher-expressive factors and less often on objective content (Basow, 1990; Bennett, 1982; Kierstead, D'Agostino, & Dill, 1988). Teacher expressiveness is important to students (Basow, 1998; Basow & Silberg, 1987). However, men's and women's expressiveness is defined differentially, because people typically consider "warmth" and qualities such as tenderness, responsiveness, and sympathy to be feminine characteristics (Basow & Silberg, 1987; Costa, Terracciano, & McCrae, 2001; Kierstead et al., 1988; Martin, 1984), whereas they consider "dynamism," enthusiasm, passion, fervor, and eagerness to be masculine characteristics (Helgeson, 1994; Mulac & Lundell, 1982).

Attitudes about emotional expressiveness stem from culturally driven implicit gender stereotyping (Banaji, Hardin, & Rothman, 1993; Kawakami & Dovidio, 2001; Swim & Cohen, 1997). Specifically, a man is expected to express "masculine" emotions when battling in warfare (see Bendya, Finucane, Kirby, & O'Donnell, 1996) and a woman is expected to express "feminine" emotions when tending a baby; but a reversal of those roles is usually deemed socially discomforting. That is, men who gently speak "moth-

ere" to babies, and women who harshly shout war cries in battle behave inconsistently with their gender stereotypes. Therefore, preconditioned, internalized, and automatically activated responses to men and women are common denominators of gender stereotyping (Bargh, 1999; England & Manko, 1991; Glick & Fiske, 1999; Nosek, Banaji, & Greenwald, 2002). Thus, it can be argued that male and female faculty's expressiveness differentially activates students' automatic perceptual filters (Basow, 1990).

A few studies support the idea that age variables negatively influence students' perceptions of personal appeal (Chabot, 2000; Levin, 1988; Perdue, & Gurtman, 1990; Peterson, 1980). For example, 24 speakers, ages 41–83, all tape-recorded a paragraph that was then listened to by 10 college-aged students, along with 40 nonstudents, ages 41–80. In analyzing the results of the study, Deal and Oyer (1991) discovered an inverse relationship between the age of the speakers' voice tones and the college-aged students' perceptions of their pleasantness. Student listeners judged the younger speakers more pleasing to listen to than they did the older ones (Deal & Oyer, 1991). Moreover, Unger and Saundra (1993) found that college students believed that younger adults are more likely to persist and achieve, or to adapt to a changing environment, than are older adults. Ageist attitudes raise underlying questions as to older adults' mental acuity and likability (Danziger & Welfel, 2000; Levy, 1996, 2001). Such studies suggest that students subscribe to a societal norm that negatively categorizes "old people" (Bargh, Chen, & Burrows, 1996; Levy, 2001; Levy & Banaji, 2002; Perdue & Gurtman, 1990).

Dunning and Sherman (1997) posed the idea that some people's stereotypes influence them to make tacit, biased inferences about members of a social category who are described in vague language (e.g., young man, young woman, old man, old woman). In the absence of relevant details, automatic processing activates unconscious attitudes toward a person in a particular category (Hopkins & Moore, 2001; Kawakami, Young, & Dovidio, 2002; Pittinsky, Shih, & Ambady, 2000). Thus, students' assessments of "old," "young," "women," and "men" agree with their schemata of socially constructed age and gender roles for each, which show up in their evaluations of male and female professors in the form of stereotyping (Basow, 1998) most often on subjective, expressive, noncontent items.

To determine whether implicit attitudes toward categories of people ("old," "young," men, women) affect students' perceptions of their professors'

enthusiasm and voice tone, we examined student evaluations of a picture-slide-audio taped lecture of a speaker whose identity was not revealed, represented by a computer generated gender- and age-neutral stick figure that we arbitrarily named in four distinct ways, one for each of four subsets of students: a "male professor under the age of 35" (YM), a "female professor under the age of 35" (YF), a "male professor over the age of 55" (OM), or a "female professor over the age of 55" (OF). The evaluation contained six teacher-expressive and six objective content items that are typically found in university student rating forms (see the evaluation form in the Appendix). In our analyses, we examined items that suggested teacher-expressiveness: professor enthusiasm (Items 1—"Spoke enthusiastically about the subject" and 9—"Appeared interested in the subject"), warmth (Item 6—"Made me feel accepted and included"), voice tone (Item 8—"Used voice tone to identify important concepts"), and confidence (Items 5—"Seemed to be conscientious" and 10—"Seemed to be relaxed and confident"). Accordingly, our expectations for student responses to these items were tested via the following hypotheses:

Hypothesis 1. Students will rate male professors higher than female professors. Gender will influence student ratings of the professors on the teacher-expressive items.

Hypothesis 2. Students will rate male and female professors under the age of 35 higher than male and female professors over the age of 55. Age will specifically influence student ratings of the professors on teacher-expressive Items 1 and 8.

Hypothesis 3. Student ratings on the teacher-expressive items will be qualified by an interaction of gender and age. Students will rate YM higher than YF, OM, and OF on the six teacher-expressive items in the evaluation, in particular, Items 1 ("Spoke enthusiastically about the subject") and 8 ("Used voice tone to identify important concepts").

METHOD

Participants

Participants were 198 female and 154 male college students from a medium-sized western university ($N = 352$). The students were enrolled in six

Introductory Psychology classes that were randomly selected for the study. Introductory Psychology fills a general education requirement, and participants enrolled in the six classes listed a wide range of academic majors: Undeclared ($n = 109$), Business and Economics ($n = 59$), Social and Behavioral Sciences ($n = 54$), Health Professions ($n = 36$), Natural Sciences ($n = 27$), Arts and Humanities ($n = 21$), Health and Physical Education ($n = 18$), Technology ($n = 16$), Education ($n = 17$), and Bachelor of Integrated Studies ($n = 2$). Exactly 313 students gave their ages as between 17 and 24 years; 22 were between 25 and 30 years old; and 17 were over 30 years of age. We did not ask them to state their ethnicity or race; however, the university student population is composed of 1% African Americans, 2% Asian/Pacific Islanders, 78% European Americans, 3% Hispanics, 1% Native Americans, 1% Non-U.S. Citizens, and 14% Other and Unknowns. All participants were treated in accordance with the American Psychological Association's ethical guidelines (APA, 2002).

Materials

All students watched the same 35-min picture-slide-audio taped presentation on "Stages of Relationship Building." The slide was a picture of a computer-generated genderless stick figure; the audiotape was read by a 45-year-old woman whose voice did not suggest her age or gender, the result of a previous manipulation check as we developed our materials. The manipulation check involved 122 students in two introductory psychology and two history classes, who listened to an audio tape of the woman reading verbatim two short paragraphs on research methods common to both academic disciplines. The students were asked to guess, with as much certainty as possible, the gender of the speaker. Fifty-nine (48.4%) students guessed the speaker was a woman, and 61 (50%) of the students guessed the speaker was a man. Two students (1.6%) indicated that they were unsure and preferred not to guess.

A list of learning objectives, a lecture outline, a 25-question multiple-choice test, and a 12-item teacher evaluation form were given to each of the participants. The teacher evaluation form was a modified College of Behavioral and Social Science instrument that included subjective (teacher-expressive) and objective (content) items that students are asked to complete at the end of every term. The instrument assesses student evaluations of the course, the teacher, and overall instructional effectiveness on a

scale from 1 (*never*) to 6 (*always*). Internal reliability (Cronbach's alpha) of the 12 items across conditions was .84. The Appendix shows an evaluation form that contains an example of one of four different professor descriptions.

Procedure

This study was conducted at the beginning of the semester to avoid potential response sets based on the age and gender of the regular course professors. The six regularly assigned professors for each of the classes were: (a) a 50-year-old woman, (b) two women, 30 and 35 years old, (c) two men, 57 and 60 years old, and (d) a 34-year-old man. In the absence of the regular course professor, the authors (a woman and a man, both over 45 years of age), who were not teaching Introductory Psychology that semester, administered the procedures in each of the six classes. We introduced ourselves and invited the students in each of the classes to participate in a study about classroom lectures. The students were informed that they could elect not to participate and that, if they chose that option, they would be excused from class. All of the students chose to participate.

The students were told that they would watch an audiovisual lecture on "Stages of Relationship Development." This topic was selected because it was thought to be of general interest to students who choose to enroll in a psychology course. The same automated 35-min slide-taped lecture was presented in each of the classes. During the audio taped lecture, a slide of a genderless stick-figure "professor," holding a pointer, standing by an overhead projector in front of a blackboard, was shown six times for 10 s each time. After the presentation, the students completed a 25-item, 10-min multiple-choice test on the material. Students in each of the six classes, unknown to them, had been previously separated into four randomly selected, approximately equal groups. They were also unaware that the groups had been set up to test the relative value of the age and gender of the "professor." In the final 5 min of class time, each participant received an evaluation form on which was printed a description of one of four different gender-age professors.

Description 1: "The professor was a female under age 35."

Description 2: "The professor was a male under age 35."

Description 3: "The professor was a female over age 55."

Description 4: "The professor was a male over age 55."

We withheld the experimental manipulation (the professor descriptions) until after the lecture and test to determine whether information about the age and gender of the professor would operate as instant primes to activate the students' implicit age and gender stereotypes (Kawakami & Dovidio, 2001). Fazio, Sanbonmatsu, Powell, and Kardes (1986) reported that the presentation of a prime automatically activates attitudes stored in memory and guides behavior outside conscious awareness. The activated attitudes serve as human perception filters that influence subsequent behavior (See Fazio, Roskos-Ewoldsen, & Powell, 1994, for an extended discussion). Moreover, many implicit attitudes preclude conscious choice (Bargh et al., 1996; Bargh & Ferguson, 2000), and may actually determine prejudicial responses (Bargh et al., 1996; Devine & Monteith, 1999). As Bargh (1999) stated: "Once a stereotype is so entrenched that it becomes activated automatically, there is really little that can be done to control its influence" (p. 378). Our goal for this study was to test for students' stereotypes toward the age and gender of college professors. They were asked to read the introductory statements and instructions carefully and then to respond thoughtfully to all of the items on the evaluation form. Students also noted their own academic major, age, and gender.

RESULTS

A $2 \times 2 \times 2$, professor gender (male and female) \times professor age (under 35 or over 55) \times student gender (male or female) between-subjects multivariate analysis of variance (MANOVA) performed on the 12 dependent variables in the student evaluation revealed negative skews and a heterogeneity of variance on Items 3, 11, and 12. Because of problems with the assumptions of normality for these items, we eliminated them and reanalyzed the data.

Hypothesis 1 predicted that professor gender would influence student ratings of the professors on the teacher-expressive items. A 3-way MANOVA (professor gender \times professor age \times student gender) on the subsequent 9-item student evaluation disclosed that gender significantly impacted the ratings, $F(9, 330) = 2.63$, $p = .006$. Follow-up univariate analyses revealed significant gender effects on Item 1 (enthusiasm), $F(1, 338) = 19.97$, $p = .001$; Item 6 (felt accepted), $F(1, 338) = 7.95$, $p = .005$;

Table I. Means and Standard Deviations by Professor Gender^a on Nine Items on the Student Evaluation

Item	Male		Female	
	M	SD	M	SD
1 Enthusiasm	2.40***	1.52	1.67	1.34
2 Precise teaching	3.39	1.14	3.20	1.32
4 Logical and organized	4.03	1.06	3.96	1.17
5 Seemed conscientious	3.29	1.34	3.25	1.34
6 Felt accepted	2.91*	1.42	2.40	1.56
7 Used scientific terminology	3.98	1.03	3.92	1.18
8 Meaningful voice tone	2.62**	1.56	2.12	1.50
9 Showed interest	2.94***	1.46	2.37	1.55
10 Relaxed and confident	3.20	1.29	3.06	1.39

Note. The numbers indicate the mean score on a 6-point Likert scale. Higher numbers indicate higher ratings.

^an = 174.

*p = .005, **p = .003, ***p = .001.

Item 8 (meaningful voice tone), $F(1, 338) = 8.86$, $p = .003$; and Item 9 (showed interest in subject), $F(1, 338) = 10.80$, $p = .001$. Students rated male professors higher on these four items. See Table I for the means and standard deviations by gender for the nine items on the evaluation form.

Hypothesis 2 predicted that students would rate male and female professors under the age of 35 higher than male and female professors over the age of 55. Age would specifically influence student ratings of the professors on teacher-expressive Items 1 and 8. The MANOVA revealed a significant effect of professor age, $F(9, 330) = 2.03$, $p = .04$, and follow-up univariate ANOVA divulged significant age effects on Item 1, $F(1, 338) = 5.07$, $p = .025$, and Item 8 $F(1, 338) = 5.43$, $p = .02$. The younger professors were rated higher than the older professors on enthusiasm (younger professor $M = 2.24$, $SD = 1.56$; older professor $M = 1.84$, $SD = 1.37$) and meaningful voice tone (younger professor $M = 2.57$, $SD = 1.62$; older professor $M = 2.17$, $SD = 1.45$).

Hypothesis 3 predicted that student ratings on the six teacher-expressive items would be qualified by an interaction of gender and age. The MANOVA revealed a significant professor age \times professor gender interaction, $F(9, 330) = 1.98$, $p = .04$. As predicted, follow-up univariate ANOVAs indicated significant gender-age interactions on Item 1 (enthusiasm), $F(1, 338) = 7.53$, $p = .006$; Item 5 (seemed conscientious), $F(1, 338) = 5.82$, $p = .02$; Item 6 (felt accepted), $F(1, 338) = 5.05$, $p = .03$; Item 8 (meaningful voice tone), $F(1, 338) = 10.17$, $p = .002$; Item 9 (showed interest in subject), $F(1, 338) = 9.423$, $p = .002$; and Item 10 (relaxed &

confident), $F(1, 338) = 5.06$, $p = .03$. See Table II for the means and standard deviations of the interactions by professor age and gender. A scrutiny of the means discloses that students consistently rated the male professor under age 35 highest on all six of the items except Item 10 ("Seemed to be relaxed and confident"). However, the YM virtually tied with the OF for highest student rating and the YF came in last on this particular item. Moreover, Fig. 1 illustrates that the effect of professor age depended on the effect of professor gender on the ratings of Items 1 (enthusiasm) and 8 (meaningful voice tone). Specifically, students rated a young professor higher on these items only when the professor was a man.

DISCUSSION

Although most of the students' ratings hovered around a low 2 or 3 out of 6 on the Likert scale, a clear age–gender interaction pattern emerged on all of the evaluation items, particularly on the six teacher-expressive statements. Students varied their ratings according to their perceptions of *who* was exhibiting enthusiasm, conscientiousness, acceptance, desirable voice tone, interest, and confidence. Even though they watched and listened to an identical stick figure and recorded voice, students evaluated it higher if they thought it was a man under age 35. The mere description of the professor's age and gender on the evaluation form after the presentation and test could have activated their automatic stereotype of a college professor and determined their biased responses (Bargh, 1999). In support of this proposition is Messner's (2000) finding that students expect their professors to look like a younger man.

Basow (1998), Helgeson (1994), and Mulac and Lundell (1982) have all proposed that dynamism (enthusiasm) is more consistent with the masculine gender role stereotype; thus, students may not have perceived dynamism in the female professor. Moreover, our findings support Dijksterhuis and van Knippenberg (1996), who argued that priming a social category activates the schemata for stereotype-consistent characteristics and obstructs the schemata for stereotype-inconsistent characteristics. The female professor's description on the evaluation form may have blocked students' schema for "woman as enthusiastic, eloquent college professor," and evoked an unconscious devaluation of the female professors' expressiveness, even during a lecture on a stereotypical feminine topic like relationship development.

Table II. Means and Standard Deviations of the Interactions by Professor Age and Professor Gender^a of Nine Items on the Evaluation

Item	Professor age	Professor gender	M	SD
1 Enthusiasm	Under 35	Female	1.66	1.32
		Male	2.83***	1.56
	Over 55	Female	1.70	1.36
		Male	1.99	1.37
2 Precise teaching	Under 35	Female	3.27	1.31
		Male	3.39	1.07
	Over 55	Female	3.25	1.32
		Male	3.39	1.20
4 Logical and organized	Under 35	Female	3.92	1.23
		Male	4.08	0.92
	Over 55	Female	4.00	1.22
		Male	3.99	1.19
5 Seemed conscientious	Under 35	Female	3.22	1.33
		Male	3.61**	1.20
	Over 55	Female	3.28	1.35
		Male	2.98	.140
6 Felt accepted	Under 35	Female	2.27	1.55
		Male	3.13*	1.31
	Over 55	Female	2.54	1.31
		Male	2.70	1.50
7 Used scientific terminology	Under 35	Female	3.99	1.31
		Male	3.99	0.97
	Over 55	Female	3.86	1.24
		Male	3.98	1.09
8 Meaningful voice tone	Under 35	Female	2.06	1.54
		Male	3.08****	1.53
	Over 55	Female	2.18	1.18
		Male	2.16	1.44
9 Showed interest	Under 35	Female	2.27	1.58
		Male	3.34****	1.32
	Over 55	Female	2.47	1.52
		Male	2.54	1.49
10 Relaxed & confident	Under 35	Female	2.90	1.44
		Male	3.21*	1.25
	Over 55	Female	3.23	1.31
		Male	3.18	1.34

Note. The numbers indicate the mean score on a 6-point Likert scale. Higher numbers indicate higher ratings.

^an = 87.

*p < .03. **p = .02. ***p = .006. ****p = .002.

The male and female stick figures identified as over age 55 could have received lower evaluations on enthusiasm and voice tone than the male and female under age 35 because of intergroup bias effects (Hewstone, Rubin, & Willis, 2002). Ashburn-Nardo, Voils, and Monteith (2001) indicated that intergroup bias occurs automatically, even in unlikely conditions. According to this theory, people who are members of an outgroup lose their personal status and become homogenous (e.g., adults over the age of 55 corre-

spond to the “senior citizen” group label and receive an abundance of negatively valenced dispositional and intellectual qualities that are imagined to be internal, stable, and global). The precise number for a professor’s age in our study (i.e., professors over the age of 55) may have activated the negative stereotype that old people are less pleasing to listen to (Deal & Oyer, 1991), and this may have inhibited students’ from detecting an interesting voice tone in the older professors. In general, people perceive and

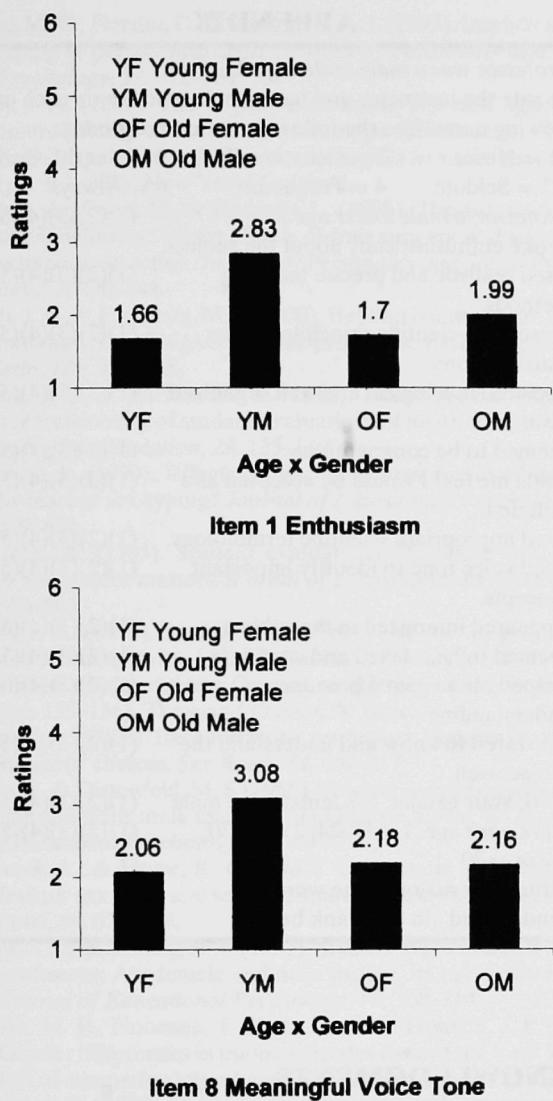


Fig. 1. Means of the interactions by professor age and professor gender on Items 1 (Spoke enthusiastically about the subject) and 8 (Used voice tone to identify important concepts).

interpret a situation that concurs with their belief systems. Previous experiences feed belief systems and then influence perceptions and interpretations of situations (Lachman, 1996; Macrae, Bodenhausen, Milne, Thorn, & Castelli, 1997). In short, the students probably perceived the lecture in response to negative interpretations of previous encounters with older people (Levy, 2001).

Students' apparent preference for the YM's enthusiasm did not increase their test scores. Their test-score means differed by only one- or two-points; out of 25 points, students who believed that the stick figure was a YM had a mean of 19; students who believed that the stick figure was an OM had a mean of 21; students who believed that the figure was a YF or OF, each, had a mean of 20. Note that students

who rated the younger male professor as the most enthusiastic displayed the lowest test scores. Our findings are consistent with Wilson's (1998) report that the perception of enthusiasm *alone* raises the total scores on student evaluations, but does not consequently raise test scores or enhance student learning.

Our participants rated the professors similarly on the objective, content items. They acknowledged a correspondence among the ratings of the YM, YF, OM, and OF with respect to the ability to organize and logically deliver, with appropriate scientific terminology, precise lectures (Items 2—"Used realistic and precise teaching methods," 4—"Lectured in a logical and well-organized manner," and 7—"Used appropriate scientific terminology"). Our results concur with other studies that indicated that significant rating differences disappeared from the evaluations of male and female faculty with the elimination of items that described teacher-expressiveness (Marsh & Roche, 1997; McKeachie, 1997). Our data suggest that student evaluations cannot objectively measure vague instructor variables, such as enthusiasm, warmth, confidence, and voice tone, particularly when they are filtered through biased, sociocultural lenses (cf., Ekman, 1994; Scherer, 1992).

Unlike previous researchers (Amin, 1994; Basow, 1995, 1998, 2000; Basow & Howe, 1987; Basow & Silberg, 1987; Das & Das, 2001; Tatro, 1995), we found no significant professor gender \times student gender interactions. In general, those researchers used real professors, fictitious scenarios, and teacher-characteristics preference surveys, which possibly evoked explicit social desirability effects and politically correct responses. The presentation of the stick figure and neutral voice allowed us to control for teaching style, physical appearance, gender, age, and voice tone, which possibly activated implicit stereotyped responses. Our results suggested that both men and women students not only implicitly preferred expressive, enthusiastic professors, they also implicitly expected young men to display the expressiveness and enthusiasm.

A number of limitations are apparent in this study. Whereas we performed a manipulation check on the neutrality of the speaker's voice, we did not perform a similar check on the stick figure, which could have appeared masculine to the students. Moreover, Basow (1998) reported that male professors are inclined more than female professors to use the lecture style. The stick figure "professor" in this study lectured exclusively during the slide-tape

recorded presentation, which could have influenced the students to perceive it as a man. However, the analyses indicated significant differences in the ratings between the younger man and the older man and between the younger and older women; if students believed that the stick figure was a man, then gender and age remain important discriminating factors. In addition, we did not check for section effects before combining our data. Further, we modified a standardized evaluation form, but included in it typical statements that relate to objective teaching characteristics and statements that refer to seemingly stereotype-tapping personality characteristics.

In sum, the results of the present study suggest that students may expect their professors to express themselves in the classroom according to socio-cultural scripts that limit full human development and influence people to appraise men and women differently even when they are displaying identical behaviors (Aries, 1996). An implicit sociocultural edict that pertains to roles in life is problematic because it distorts our perceptions and memories of people in ways that are detrimental to men and women and to older adults, regardless of gender. In the higher education arena, the problematic effects include practical, economic ramifications. When professors receive evaluations biased by implicit stereotyping, their current and future lifestyles are adversely affected in terms of potential tenure and promotion possibilities, along with the uncertainty of salary increases. Further, the professors may realize that something is wrong, but they are usually not able to identify what it is, and thus may feel confused or inept, and eventually become what students expect them to be (McKeachie, 1997).

The replication of our results by future researchers may encourage educators to construct a multidimensional student evaluation instrument that eliminates teacher expressiveness items that are potentially biased against the age and gender of the professor. As previous researchers have noted, a professor's enthusiasm does not necessarily predict teaching competence (Basow, 1990; Ogden, Chapman, & Doak, 1994) or increased student learning (Basow & Distenfeld, 1985; Williams & Ceci, 1997); "no single criterion of effective teaching is sufficient" (Marsh & Roche, 1997, p. 187). Future researchers could also examine the extent to which students' stereotypical judgments accumulate through time to influence the judgments of rank and tenure committees and the salary assessments university administrators make.

APPENDIX

The professor was a male under age 35.

Please rate the instructor you have just listened to on each of the following items. Use the following scale to respond.

1 = Never	3 = Occasionally	5 = Nearly Always
2 = Seldom	4 = Frequently	6 = Always

The professor, a male under age 35: (1)(2)(3)(4)(5)(6)

1. Spoke enthusiastically about the subject. (1)(2)(3)(4)(5)(6)
2. Used realistic and precise teaching methods. (1)(2)(3)(4)(5)(6)
3. Presented scientific principles rather than opinions. (1)(2)(3)(4)(5)(6)
4. Lectured in a logical and well organized manner. (1)(2)(3)(4)(5)(6)
5. Seemed to be conscientious. (1)(2)(3)(4)(5)(6)
6. Made me feel I would be accepted and included. (1)(2)(3)(4)(5)(6)
7. Used appropriate scientific terminology. (1)(2)(3)(4)(5)(6)
8. Used voice tone to identify important concepts. (1)(2)(3)(4)(5)(6)
9. Appeared interested in the subject. (1)(2)(3)(4)(5)(6)
10. Seemed to be relaxed and confident. (1)(2)(3)(4)(5)(6)
11. Helped me to gain a broader understanding. (1)(2)(3)(4)(5)(6)
12. Appeared to know and understand the subject well. (1)(2)(3)(4)(5)(6)
13. Mark your gender. 1 = female; 2 = male (1)(2)(3)(4)(5)(6)
14. Mark your age. 1 = 17-24; 2 = 25-30; 3 = over 30 (1)(2)(3)(4)(5)(6)
15. Write your major or the word "undeclared" in the blank below.

ACKNOWLEDGMENTS

We acknowledge valuable input and encouragement from individuals who read our revised manuscript: Eric Amsel, Brooke Arkush, Lauren Fowler, Richard Grow, Robert Hogge, and Scott Richards.

REFERENCES

- Abrami, P. C., & d'Appollonia, S. (1999). Current concerns are past concerns. *American Psychologist*, 54, 519-520.
- American Psychological Association. (2002). Ethical principles of psychologists and code of conduct. Retrieved October 4, 2002, from <http://WWW.apa.org/ethics>
- Amin, M. E. (1994). Gender as a discriminating factor in the evaluation of teaching. *Assessment and Evaluation in Higher Education*, 19, 135-143.
- Aries, E. (1996). *Men and women in interaction: Reconsidering the differences*. New York: Oxford University Press.
- Ashburn-Nardo, L., Voils, C. I., & Monteith, M. J. (2001). Implicit associations as the seeds of intergroup bias: How easily do they take root? *Journal of Personality and Social Psychology*, 81, 789-799.

- Banaji, M. R., Hardin, C., & Rothman, A. J. (1993). Implicit stereotyping in person judgment. *Journal of Personality and Social Psychology, 54*, 272–281.
- Bargh, J. A. (1999). The cognitive monster: The case against the controllability of automatic stereotype effects. In S. Chaiken & Y. Trope (Eds.), *Dual process theories in social psychology* (pp. 361–382). New York: Guilford.
- Bargh, J. A., Chen, M., & Burrows, L. (1996). The automaticity of social behavior: Direct effects of trait concept and stereotype activation on action. *Journal of Personality and Social Psychology, 71*, 230–244.
- Bargh, J. A., & Ferguson, M. J. (2000). Beyond behaviorism: On the automaticity of higher mental processes. *Psychological Bulletin, 126*, 925–945.
- Barnes, L. L., & Barnes, M. W. (1993). Academic discipline and generalizability of student evaluations of instruction. *Research in Higher Education, 23*, 135–149.
- Basow, S. A. (1990). Effects of teacher expressiveness: Mediated by teacher sex-typing? *Journal of Educational Psychology, 82*, 599–602.
- Basow, S. A. (1995). Student evaluations of college professors: When gender matters. *Journal of Educational Psychology, 87*, 656–665.
- Basow, S. A. (1998). Student evaluations: The role of gender bias and teaching styles. In L. H. Collins, J. C. Chrisler, & K. Quina (Eds.), *Arming Athena: Career strategies for women in academe* (pp. 135–156). Thousand Oaks, CA: Sage.
- Basow, S. A. (2000). Best and worst professors: Gender patterns in students' choices. *Sex Roles, 34*, 407–417.
- Basow, S., & Distenfeld, M. S. (1985). Teacher expressiveness: More important for male teachers than female teachers? *Journal of Educational Psychology, 77*, 45–52.
- Basow, S. A., & Howe, K. G. (1987). Evaluations of college professors' sex-type, and sex, and students' sex. *Psychological Reports, 60*, 671–678.
- Basow, S. A., & Silberg, N. T. (1987). Student evaluations of college professors: Are female and male professors rated differently? *Journal of Educational Psychology, 79*, 308–314.
- Bendya, M. E., Finucane, T., Kirby, L., & O'Donnell, J. P. (1996). Gender differences in public attitudes toward the Gulf War: A test of competing hypotheses. *Social Science Journal, 33*, 1–22.
- Bennett, S. K. (1982). Student perceptions of and expectations for male and female instructors: Evidence relating to the question of gender bias in teaching evaluations. *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.
- Bousfield, W. A. (1940). Students' ratings of qualities considered desirable in college professors. *School and Society, 51*, 253–256.
- Centra, J. A., & Gaubatz, N. B. (2000). Is there gender bias in student evaluations of teaching? *Journal of Higher Education, 71*, 17–33.
- Chabot, H. F. (2000). Predicting ageist and sexist attitudes and the conditions for their existence (Doctoral dissertation, University of New Hampshire). *Dissertation Abstracts International, 60*, 12-B. (UMI No. 9953418)
- Costa, P., Jr., Terracciano, A., & McCrae, R. R. (2001). Gender differences in personality traits across cultures: Robust and surprising findings. *Journal of Personality and Cognition, 81*, 322–331.
- Cravens, T. F. (1996, March). *Students' perceptions of the characteristics of teaching excellence*. Paper presented at the National Social Science Conference, Reno, NV.
- Danziger, P. R., & Welfel, E. R. (2000). Age, gender, and health bias in counselors: An empirical analysis. *Journal of Mental Health Counseling, 22*, 135–149.
- d'Appolonia, S., & Abrami, P. C. (1997). Navigating student ratings of instruction. *American Psychologist, 52*, 1198–1208.
- Das, M., & Das, H. (2001). Business students' perceptions of best university professors: Does gender role matter? *Sex Roles, 45*, 665–676.
- Deal, L. V., & Oyer, H. J. (1991). Ratings of vocal pleasantness and the aging process. *Folia Phoniatrica, 43*(1), 44–48.
- Devine, P. G., & Monteith, M. J. (1999). Automaticity and control in stereotyping. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 339–360). New York: Guilford.
- Dijksterhuis, A., & van Knippenberg, A. (1996). The knife that cuts both ways: Facilitated and inhibited access to traits as a result of stereotype activation. *Journal of Experimental Psychology, 32*, 271–288.
- Dunning, D., & Sherman, D. A. (1997). Stereotypes and tacit inference. *Journal of Personality and Social Psychology, 73*, 459–471.
- Ekman, P. (1994). Strong evidence for universals in facial expressions: A reply to Russell's mistaken critique. *Psychological Bulletin, 115*, 268–287.
- England, E. M., & Manko, W. (1991, April). *College student evaluations of stereotypic female subcategory members*. Paper presented at the meeting of the Eastern Psychological Association, New York, NY.
- Fazio, R. H., Roskos-Ewoldsen, D. R., & Powell, M. C. (1994). Attitudes, perception, and attention. In P. M. Niedenthal & S. Kitayama (Eds.), *The heart's eye: Emotional influences in perception and attention* (pp. 197–216). San Diego, CA: Academic Press.
- Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology, 50*, 229–238.
- Feldman, K. A. (1986). The perceived instructional effectiveness of college teachers as related to their personality and attitudinal characteristics: A review and synthesis. *Research in Higher Education, 24*, 139–213.
- Freeman, H. (1994). Student evaluations of college instructors: Effects of type of course taught, instructor gender and gender role, and student gender. *Journal of Educational Psychology, 86*, 627–630.
- Glick, P., & Fiske, S. T. (1999). Gender, power dynamics, and social interaction. In M. M. Ferree & J. Lorder (Eds.), *Revisioning gender: The gender lens* (pp. 365–398). Thousand Oaks, CA: Sage.
- Greenwald, A. G. (1997). Validity concerns and usefulness of student ratings of instruction. *American Psychologist, 52*, 1182–1186.
- Greenwald, A. G., & Gillmore, G. M. (1997). Grading leniency is a removable contaminant of student ratings. *American Psychologist, 52*, 1209–1217.
- Guerrero, L. K., & Miller, T. (1998). Associations between nonverbal behaviors and initial impressions of instructor competence and course content in videotaped distance education courses. *Communication Education, 47*(1), 30–42.
- Helgeson, V. S. (1994). Prototypes and dimensions of masculinity and femininity. *Sex Roles, 31*, 653–682.
- Hewstone, M., Rubin, M., & Willis, H. (2002). Intergroup bias. *Annual Review of Psychology, 53*, 575–604.
- Hopkins, N., & Moore, C. (2001). Categorizing the neighbors: Identity, distance, and stereotyping. *Social Psychology Quarterly, 64*, 239–252.
- Kalavar, J. M. (2001). Examining agesim: Do male and female college students differ? *Educational Gerontology, 27*, 507–513.
- Kawakami, K., & Dovidio, J. F. (2001). The reliability of implicit stereotyping. *Personality and Social Psychology Bulletin, 27*, 212–225.
- Kawakami, K., Young, H., & Dovidio, J. F. (2002). Automatic stereotyping: Category, trait, and behavioral activations. *Personality and Social Psychology Bulletin, 28*, 3–15.

- Kierstead, D., D'Agostino, P., & Dill, P. (1988). Sex role stereotyping of college professors: Bias in students' ratings of instructors. *Journal of Educational Psychology, 80*, 342-344.
- Kite, M. E. (2001). Changing times, changing gender roles: Who do we want women and men to be? In R. K. Unger (Ed.), *Handbook of the psychology of women and gender* (pp. 215-227). New York: Wiley.
- Lachman, S. J. (1996). Processes in perception: Psychological transformations of highly structured stimulus material. *Perceptual and Motor Skills, 83*, 411-418.
- LeBlanc, R. (1998, June/July). Good teaching: The top 10 requirements. *The Teaching Professor*, p. 1.
- Levin, W. C. (1988). Age stereotyping: College student evaluations. *Research on Aging, 10*(1), 134-148.
- Levy, B. R. (1996). Improving memory in old age through implicit self-stereotyping. *Journal of Personality and Social Psychology, 71*, 1092-1107.
- Levy, B. R. (2001). Eradication of ageism requires addressing the enemy within. *Gerontologist, 41*, 578-579.
- Levy, B. R., & Banaji, M. R. (2002). Implicit ageism. In T. D. Nelson (Ed.), *Ageism: Stereotyping and prejudice against older persons* (pp. 49-75). Cambridge, MA: MIT Press.
- Macrae, C. N., Bodenhausen, G. V., Milne, A. B., Thorn, T. M. J., & Castelli, L. (1997). On the activation of social stereotypes: The moderating role of processing objectives. *Journal of Experimental Psychology, 33*, 471-489.
- Marsh, H. W., & Roche, L. A. (1997). Making students' evaluations of teaching effectiveness effective: The critical issues of validity, bias, and utility. *American Psychologist, 52*, 1187-1197.
- Martin, E. (1984). Power and authority in the classroom: Sexist stereotypes in teaching evaluations. *Signs, 9*, 482-492.
- McKeachie, W. J. (1997). Student ratings: The validity of use. *American Psychologist, 52*, 1218-1225.
- Messner, M. A. (2000). White guy habitus in the classroom: Challenging the reproduction of privilege. *Men and Masculinities, 2*, 457-469.
- Mulac, A., & Lundell, T. L. (1982). An empirical test of the gender-linked language effect in a public speaking setting. *Language and Speech, 25*, 243-256.
- Nosek, B. A., Banaji, M., & Greenwald, A. G. (2002). Harvesting implicit group attitudes and beliefs from a demonstration website. *Group Dynamics, 6*(1), 101-115.
- Ogden, D., Chapman, A. D., & Doak, L. (1994, November). *Characteristics of good/effective teachers: Gender differences in student descriptors*. Paper presented at the meeting of the Mid-South Educational Research Association, Nashville, TN.
- Perdue, C. W., & Gurtman, M. B. (1990). Evidence for the automaticity of ageism. *Journal of Experimental Social Psychology, 26*, 199-216.
- Peterson, C. C. (1980). Are young people biased against older teachers? *Journal of Genetic Psychology, 136*, 309-310.
- Pittinsky, T. L., Shih, M., & Ambady, N. (2000). Will a category cue you? Category cues, positive stereotypes, and reviewer recall for applicants. *Social Psychology of Education, 4*(1), 53-65.
- Radmacher, S. A., & Martin, D. J. (2001). Identifying significant predictors of student evaluations of faculty through hierarchical regression analysis. *Journal of Psychology, 135*, 259-268.
- Scherer, K. R. (1992). On social representations of emotional experience: Stereotypes, prototypes, or archetypes? In M. von Cranach, W. Doise, & G. Mugny (Eds.), *Social representations and the social bases of knowledge* [Swiss Monographs in Psychology] (Vol. 1, pp. 30-36). Lewiston, NY: Hogrefe & Huber.
- Swim, J. K., & Cohen, L. L. (1997). Overt, covert, and subtle sexism: A comparison between the attitudes toward women and modern sexism scales. *Psychology of Women Quarterly, 21*, 103-118.
- Tatro, C. N. (1995). Gender effects on student ratings of faculty. *Journal of Research and Development in Education, 28*, 169-173.
- Unger, R. K. (1979). Sexism in teacher evaluation: The comparability of real life to laboratory analogs. *Academic Psychology Bulletin, 1*(2), 163-170.
- Unger, R. K., & Saundra. (1993). Sexism: An integrated perspective. In F. L. Denmark & M. Paludi (Eds.), *Psychology of women: A handbook of issues and theories* (pp. 141-188). Westport, CT: Greenwood Press.
- Upham, D. (2001, November). *Celebrating teachers*. Paper presented at the meeting of Kappa Delta Pi, Orlando, FL.
- Williams, W. M., & Ceci, S. J. (1997). "How'm I doing?" Problems with student rating of instructors and courses. *Change, 29*(5), 12-23.
- Wilson, R. (1998, January 16). New research casts doubt on value of student evaluations of professors. *Chronicle of Higher Education, 44*, 12-14.

Copyright of *Sex Roles* is the property of Springer Science & Business Media B.V. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.