

**Beauty in the Classroom: Are Female Students Influenced by the Physical Appearance of Their Male Professors?**

Ya'arit Bokek-Cohen, Lecturer and Marketing area coordinator, Ariel University Center of Samaria, Department of Economics and Business Administration, Israel, [bokek@bezeqint.net](mailto:bokek@bezeqint.net)  
Nitza Davidowitz, Lecturer and Director of Academic Development and Assessment Department, Ariel University Center of Samaria, Israel, [d.nitza@ariel.ac.il](mailto:d.nitza@ariel.ac.il)

**Abstract**

Students were asked to rate the physical attractiveness of fifty five professors and these data were correlated with teaching ratings as collected in the previous semester. Male professors considered more attractive received higher teaching ratings, but only from female students. The more attractive female professors did not receive higher ratings from either male or female students. We conclude that male academics enjoy a "beauty premium," while their female colleagues do not. This "discrimination" stems from the contradiction between role images and gender images. When the role image corresponds to the gender image, the "beauty effect" benefits attractive people. When there is a contradiction between these two images, however, the beautiful person does not enjoy the "beauty premium."

**Keywords**

Physical attractiveness, teaching rating, gender image, role image, beauty premium

**Introduction**

Numerous studies have been conducted examining the impact of physical appearance on interpersonal interaction. **Physical attractiveness** is the perception of the physical traits of an individual human person as pleasing or beautiful. Judgment of the attractiveness of physical traits is partly universal to all human cultures, partly dependent on culture or society or time period, and partly a matter of individual subjective preference (Wikipedia, 2008). In an effort to define beauty standards, scientists have found no universal beauties, and have therefore settled on the method of consensus. According to this method, researchers simply ask a number of judges to rate men and women's looks (Hatfield and Sprecher, 1986).

Physical attractiveness can have a significant effect on how people are judged in such fields as employment or social opportunities, friendship, sexual behavior, and marriage. In many cases, humans attribute positive characteristics such as intelligence and honesty to attractive people without consciously realizing it (Wikipedia, 2008). Hatfield and Sprecher (1986) claim that there is a stereotyping process that ensures that beauty equals goodness. People treat attractive people quite differently than ugly people, creating a "self-fulfilling prophecy."

Beauty is an influential factor in all spheres of life: children prefer to associate with attractive children and reject the company of unattractive children (Weiss, 1991). **In marriage**, more attractive people are preferred as spouses (Buss, 1989). More attractive women manage to marry established men and thus advance their own economic status. **In interpersonal relationships**, beautiful people gain more help and cooperation from others (Bull & Rumsey, 1988). The opinions of beautiful people receive greater acceptance, and beautiful people are considered to be nicer, more intelligent, friendlier, and more sensitive. **In the courts**, beautiful people accused of criminal offenses have a greater chance of being acquitted than less attractive people (Efran, 1974). Beautiful people who are convicted receive lighter sentences than less attractive criminals (Landy & Aronson, 1969; Stewart, 1980). **In hospitals**, it has been found that beautiful patients receive more attention and better care from the medical staff, and they also receive a better prognosis. In the case of male attractiveness, height is one of the constituent factors (Hatfield & Sprecher, 1986, Wikipedia, 2008); accordingly, **in**

**the political arena**, the taller of the two main US presidential candidates was elected in all the elections between 1900 and 1968 (Hatfield & Sprecher, 1986, pp. 200). **In the area of employment and salaries**, it has been found that tall people, and tall men in particular, earn higher salaries, on average, and have a better chance of being hired than their shorter peers. Beautiful candidates are perceived as more qualified for the position and receive higher initial salary offers (Dipboye, Arvey & Terpstra, 1977; Hamermesh & Biddle, 1994). Beautiful employees receive better performance evaluations than others, and faster promotion (Morrow, McElroy, Stamper & Wilson, 1990). The more attractive people (men and women alike) are considered to be, the higher their salaries.

Jackson, Hunter and Hodge (1995) employed a meta-analysis to test hypotheses about the relationship between physical attractiveness and intellectual competence. They concluded that attractive people are perceived as more competent than less attractive people; the attractiveness effect was stronger for males than for females.

Hamermesh and Biddle (1994) term all these outstanding benefits the “beauty premium.” Researchers have given little attention to the ramifications of beauty in academic institutions; most of the studies in this field have examined the impact of the students’ attractiveness on their teachers’ evaluations. The purpose of this study is to examine the effect of beauty in Israeli academy. We turn now to a review of the main findings of studies on this issue, and go on to present our research hypotheses.

### **The Effects of Physical Appearance in Higher Education System**

Lombardi and Tocci (1979) found a positive correlation between professor’s attractiveness and his warmth, sensitivity, ability to communicate, knowledge of subject matter and superiority. No interaction effect of sex of rater and sex of ratee was found. Goebel and Cashen (1979) found that attractive teachers were seen as more friendly, better organized, and more likely to encourage students to interact. O’Reilly (1987) asked students to attend a lecture of an attractive or unattractive female teacher. The students were then asked to evaluate the teacher on a 10-item rating form and to indicate their own sex. The findings showed that regardless of sex of the student, the physical attractiveness of the teacher enhanced teaching evaluations. Naumann (1988) showed subjects videotapes of a female instructor: half of them watched an attractive one and half watched an unattractive one. He found no significant effect on ratings for physical attractiveness. Romano & Bordieri (1989) also examined the effect of the physical attractiveness of college professors on the students’ impressions. The researchers asked students to listen to a fifteen-minute audiotape describing typical first-time experiences for college freshmen. As each student listened to the tape, he was shown a black and white facial photograph of an alleged college professor. They found that attractive professors and female professors received the highest teaching scores. The interaction of the sex of the professor and physical attractiveness was not significant. Newsom (1990) also examined the effects of physical attractiveness on teacher performance evaluation. Data were collected from subjects after they had observed one of two simulated preobservation conference videotapes. In one of the videotaped conferences, an actress portrayed an attractive teacher, and in the other one, the actress was less attractive. The results show significant higher evaluations of the attractive “teacher.” Feeley (2002) found significant correlation between the perceived level of attractiveness of instructors and their vocal clarity and such aspects as teaching effectiveness, affective learning, and nonverbal immediacy. Data from the website Ratemyprofessors.com were used by various researchers in order to estimate the influence of the appearance of professors on students’ teaching evaluations. Ratemyprofessors.com is one of the main sites that allow students to post anonymous ratings of college professors in the US and Canada. Students rate professors on three dimensions: easiness, helpfulness, and clarity of teaching. In addition, students can rate attractiveness by assigning a “chili pepper” icon to indicate “hotness,” a concept generally understood as the physical attractiveness of the instructor. Obviously, these data might be biased for many reasons. For example, some of the raters may enter and rate the same professor more than once. Another possibility is that raters may designate a particular instructor as “hot” as a joke. In spite of these potential limitations, some interesting and statistically significant findings were found in the database of this website. Felton, Mitchell, & Stinson (2004) analyzed data from Ratemyprofessors.com in order to determine whether any

relationship exists among perceived quality, easiness, and “hotness” scores. Felton, Mitchell and Stinson (2004) examined data of 3,190 professors and found that the correlation between quality and sexiness is 0.30. Kindred and Mohammed (2005) analyzed the assigned scores, and content analyzed the comments of a sample of 1,054 ratings from the RateMyProfessors.com web site. They found that the correlation between appearance and clarity was 0.34. Hamermesh and Parker (2005) examined the impact of instructors’ looks on their instructional ratings in the courses that they teach. They found that measures of perceived beauty have a substantial and independent positive impact on instructional ratings by undergraduate students. The impact of beauty on instructional ratings was statistically significant for both women and men, but was three times greater for male faculty than for their female colleagues. Hamermesh and Parker discuss their findings and ask whether they imply that beauty itself makes instructors more productive in the classroom, or whether students are merely reacting to an irrelevant characteristic that differs among instructors. We assume that if the latter proposition is true, then we might expect this reaction to differ between both female and male students. A comparison of female and male ratings of both instruction and beauty will allow us to answer to this question. Accordingly, we measured the beauty impact for female and male students separately. Sussmuth (2006) followed the strategy of Hamermesh and Parker and examined whether perceived attractiveness of German university teachers was correlated with the ratings they received from students. He asked 50 students to assess the attractiveness of 50 tenured and non-tenured teaching staff members of a German university. He found that the impact of teacher’s looks on their average instructional ratings for the German sample was about half of the one found in the American university studied by Hamermesh and Parker.

These findings served as the basis for the study, which aims to examine whether the “beauty premium” is also accepted in Israeli academy and, if so, whether there are differences between male instructors and female instructors in terms of the physical appearance effect. Specifically, we sought to examine whether physical attractiveness has an effect on the students’ perception of their teachers in the higher education system in Israel, which presumes to be devoid of such phenomena. In formal terms, beauty or physical attractiveness are irrelevant to the instructor’s performance, and standards for judging and evaluating faculty are purported to be objective and based on instructor’s pedagogical and research abilities. Instructors are expected to furnish their students with theoretical knowledge and the relevant skills and tools, and the students are expected to acquire this knowledge and these tools from their instructors lecturer, as well as to rate their teaching in the feedback questionnaire, regardless of their physical attractiveness.

Gender differences in the beauty effect, were taken into consideration because of the stereotypical gender image of women as unqualified or as less qualified than men for scientific research and creative work. This image is based on the widespread opinion that they have not been imbued with the necessary masculine characteristics such as rationality, abstraction, initiative, assertiveness, and independence (Toren, 2005). Science, scientists, and research carry an imprint of masculinity, and this image is accepted not only by men but also by many women (Ekehammer, 1985). According to figures provided by the Planning and Budget Department of the Council for Higher Education in Israel, it was found that only 24 percent of academic faculty members at all levels are women; the higher the echelon, the lower the number of female faculty members. The small number of women among high-level faculty members is characteristic not only of Israel, but of women in the fields of science all around the world, and in Europe and the US in particular (Messer-Yaron & Cahanovich, 2003). A survey of public attitudes towards women and science commissioned by the Ministry of Science in 2003 revealed that the three fields perceived as scientific areas in which Israel’s achievements are impressive (technology, communications, and weapons) were not cited as fields suitable for women (Messer-Yaron & Cahanovich, 2003).

Gender images constitute a basis for various performance expectations directed at various people, and expectations affect evaluations of others as well as the quality of the individual’s performance (Toren, 2005). On the other hand, physical attractiveness is one of the elements which compose the gender image of women, and it was found that self-attractiveness is more central to the gender role of women than to that of men, and is also cardinal to women’s personal identity, self-esteem, and interpersonal outcomes (Bar-Tal & Saxe, 1976).

Generally, women assign relatively low weight to the physical attractiveness of others (Feingold, 1992). Women's attractiveness enhances their perceived femininity, and femininity is supposedly incongruent with the skill, talent, and job requirements of high status and "masculine-type" jobs.

Based on the combination of the masculine image of the profession of instructor with the great importance of beauty in women's lives, one can predict that the "beauty premium" that women receive will be greater than that of men. Indeed, Heilman and Saruwatari (1979) found that when competing for traditional "male jobs," beauty helped women applicants when the job was clerical. When the job was a high-status managerial one, interviewers preferred the less attractive women.

Thus, it can be hypothesized that attractive female instructors will be perceived as more "feminine" and receive lower ratings than male instructors. On the other hand, the more attractive a female professor, and hence the more she is perceived as "feminine," the more students may feel she is supportive, sensitive, and empathic, therefore raising their evaluations of her teaching. Homely female professors will be "punished" because of this beauty stereotype.

One variable that has been virtually ignored in the cited studies is the student's gender. This variable was taken into consideration in the study of Bachen, McLoughlin, and Garcia (1999), which did not include physical attractiveness as a variable. They found that female students rated female teachers higher than male teachers on several teaching dimensions. They concluded that evaluations of teachers are somewhat guided by students' expectations of appropriate gender role behavior. Davidowitz (2003) has also found that female students give overall higher ratings than their male peers.

In the classic experiment conducted by Dion, Berscheid and Walster (1972) on differences between male and female raters in the evaluation of the beauty of individuals (not instructors), no interaction was found between the rater's gender and the gender of the evaluated image. Assayag's (1998) study on managers' evaluations of job candidates found that the beauty effect is diminished when the rater's gender was identical to that of the ratee. One possible reason for this is jealousy: it is possible that women give lower evaluations to attractive women because they are jealous of them (Dion, Berscheid & Walster, 1972). Another explanation is that all human beings are more sensitive and attentive to the attractiveness of members of the opposite sex, and therefore the effect is weakened in the case of members of the same sex. Helmes and Rode (1982), and Lewis and Walsh (1978) studied sex differences in male and female subjects' evaluations of female counselors. They found that the physical attractiveness of the counselor was a factor for female but not for male subjects. Holahan & Stephan (1981) asked male and female students to read an essay allegedly written by a physically attractive or unattractive female. Subjects were classified as having traditional, moderate, or liberal attitudes toward women (equivalent to gender stereotypes). Female subjects' evaluations were not affected by the physical attractiveness of the writer. Males, however, were influenced by their sex-role attitudes and the physical attractiveness of the writer. The findings showed a reversal of the physical attractiveness stereotype for liberal males.

The purpose of the present study was to replicate and extend the work of Hamermesh and Parker (2005) and Sussmuth (2006) by incorporating student's gender into the same research paradigm and dependent measures. In other words, we looked for an interaction effect between the gender of the student and the gender of the instructor being evaluated. Based on Feingold's (1990) meta-analysis showing that men place greater value on physical attractiveness than women, it is reasonable to predict that the physical attractiveness of instructors will be strongly correlated with their teaching rating, made by male students. Moreover, people are naturally more sensitive to the physical attractiveness of members of the opposite sex, and therefore it can be hypothesized as follows:

## **Research Hypotheses**

1. The positive relationship between the physical attractiveness and teaching ratings of male instructors will be found to be stronger in the female students' feedback than in that of the male students.
2. The positive relationship between the physical attractiveness and teaching ratings of female instructors will be found to be stronger in the male students' feedback than in that of the female students.

### ***Method:***

#### ***Research tools***

##### **Measuring the Teaching Rating Variable**

At the end of each semester, as in most Israeli colleges, faculty are required to give all students an opportunity to respond to a formal questionnaire rating instructors and course on a five-point Likert-scale. The standard feedback questionnaire includes eight items, six of which pertain to the instructor and two to the tutor and the physical conditions in the classroom. Of the six questions about the instructor, one pertains to a general evaluation of the instructor; this question was used as the main indication of the quality of the lecturer's teaching. The questionnaire also poses other questions concerning the student, such as gender, year in program, etc. The questionnaire is distributed during the last two weeks of each semester. Students complete the survey during the class period. The instructor leaves the classroom while the students fill out the questionnaire.

##### **Measuring the Physical Attractiveness Variable**

Most studies in the field used stimuli from either pole of the continuum, attractive or unattractive, whereas our study, similar to that of Hamermesh and Parker (2005) employed a beauty variable presented as a scale rather than a dichotomous variable. Like Morrow (1990), beauty in this study was addressed in a holistic manner, incorporating frame, height, facial expressions and body language, rather than confining ourselves to the evaluation of facial image. The consensus method was used to measure the variable of physical attractiveness, because this method is widely accepted by most researchers in the field (Hatfield & Sprecher, 1986). Researchers simply ask a number of judges to rate men and women's looks. Students who responded to the questionnaire mentioned above were asked to address one additional question: "To what degree is the instructor considered a good-looking person?" This question was written on the blackboard by the person conducting the survey when the students began completing the survey, and they were asked to copy and respond to the question. Mean beauty rankings of all raters were treated as instructor's physical attractiveness score.

This study examines data for teachers of the college for four semesters, from spring 2005 through spring 2006. Because the focus of the study is the relationship between instructor attractiveness and instructor rating, the unit of analysis is the teacher.

### ***Sample***

At the onset of the study, a permission to conduct the study was obtained from the Dean of the Faculty of Social Sciences and Humanities. Instructors were telephoned, using an alphabetical list of all the faculty members. In this manner, without knowing any of them personally, there was an attempt to obtain a heterogeneous sample in terms of the instructors' attitudes towards the subject. The instructors were asked to give their permission to add a question about their physical attractiveness to the rating form. Since a person's physical attractiveness is a sensitive issue related to his self-esteem, it is very surprising that studies on this issue have not reported any difficulty in obtaining the consent of instructors to participate in the study. This may be because no such difficulties were found, but it may also be that difficulties were encountered but were not reported by the researchers. The deliberations of the researchers and instructors and the students' diverse responses to the questions are all part of the phenomenon under study and its significance, and therefore it is important to describe them hereunder.

Extremely diverse responses were obtained from the instructors who were asked to participate in the study. Some instructors were extremely enthusiastic, showed considerable interest in the study, and expressed admiration for its unconventional subject. Others (mainly male) expressed indifference but gave their consent immediately. Polite refusals were also obtained, with explanations such as "I don't think this is appropriate for the institution," as well as vehement refusals, amazement, and anger at the perceived "undignified" subject of the study. One female instructor even tried to organize a petition for the cancellation and immediate termination of the study. It should be noted that the percentage of those who refused to participate was significantly higher among the women: out of the 58 female instructors that we approached, 18 refused (31%), whereas only (10%) of the 59 men that we approached refused. A number of instructors commented that the subject is inappropriate for an academic study.

The first author included herself in the list of instructors whose physical attractiveness was ranked. However, the second author, in spite of her appreciation of the importance of the study, objected to the very idea of linking beauty and physical attractiveness to students' feedback on academic performance. The persons who conducted the feedback survey expressed resentment at being required to write an additional question on the blackboard, and perceived it as a factor that extended the time duration of the task that they had to perform. Many students laughed at the subject of the study and joked while completing the questionnaire. Several students did not even bother to answer the question that we added about beauty<sup>1</sup>.

In total, data was collected on 55 instructors, 35 of them male and 20 female. Of these, only 49 instructors (31 men and 18 women) had scores for the beauty evaluation in two feedback sessions that were conducted at different times. The age range of the male instructors was 31-70 and that of the female instructors was 29-75. The mean age of the male instructors was 47.97, SD=11.81 and that of the women was 45.84, SD=11.80. Twenty-eight instructors were from the Faculty of Social Sciences and Humanities (57.3%), six from the Health Sciences (12.2%), nine from the Faculty of Engineering and Architecture (12.2 %) and six from the Natural Sciences (12.2%). The questionnaires were distributed to a total of 1388 students, including 553 students in the second semester of 2005, 243 students who attended the summer semester in 2005, 321 students in the first semester of 2006, and 271 students in the second semester of 2006. In order to rule out the self-selection effect for instructors, we examined whether the instructors in the sample received higher evaluations than the average general instructor evaluations at the college. In the two years in which the study was conducted, the college mean was 4.016 and the mean for the sample was 4.028; thus there was no significant difference between the means.

### ***Procedure***

When the faculty evaluation survey was administered at the end of the semester, the students were requested to respond to an additional question intended for academic research. The students were asked to address a question that was written on the blackboard: "To what degree is the instructor considered a good-looking person?" In order to analyze the data, the data were grouped by instructor. It is important to note that the low number of subjects was due to the need to examine evaluations at different times by different raters (the instructor's beauty by the students in one class and the instructor's teaching ratings by students of the class that studied with the instructor the previous semester). According to this procedure, it is possible to neutralize the effect of completing the rating questionnaire along with the beauty questions, i.e. the beauty evaluation and the teaching evaluation were taken at different times and from different people.

---

<sup>1</sup> The possibility of the "self-selection" effect of the students will be addressed in the Discussion.

## Results

In order to examine the study hypotheses, a comparison is made in order to examine the sex-based differences in teaching rating means and in the strength of the relationship between teaching and beauty ratings. If there are no differences between the means of the male and female students' teaching evaluations for male instructors, and if there are differences in the strength of the correlation coefficients between instructors' ratings and beauty ratings, then hypothesis 1 has been supported. Similarly, if there is no differences between the means of the male and female students' evaluations of female instructors, and that there are differences in the strength of the correlation coefficients between instructors' ratings and beauty ratings, then hypothesis 2 has been corroborated. Firstly, the "simple" mean differences between instructor evaluations are presented, and secondly the differences in the strength of the correlation coefficients between instructors' ratings and beauty ratings presented.

### Differences in the Teaching Ratings by the Gender of Instructor and Student

To examine these differences, a two factorial analysis of variance was performed on the evaluations of instructors, by the instructor's gender and student's gender, with repeated measurements for the student's gender.

**Table No. 1: Means and Standard Deviations of Teaching Evaluations by the Instructor's Gender and the Student's Gender**

Instructor's gender Student's gender	Male instructors (n=22)		Female instructors (n=13)		Total (n=35) <sup>2</sup>	
	Mean	SD	Mean	SD	Mean	SD
Male students	4.05	0.76	4.26	0.35	4.13	0.64
Female students	4.21	0.61	4.31	0.35	4.24	0.53
Total	4.13	0.68	4.29	0.35	4.21	0.57

It is evident from Table 1 and from the results of the ANOVA that no differences were found in instructors' ratings, either by instructor's gender ( $F(1,33)=0.75$ ,  $p>.05$ ) or by student's gender ( $F(1,33)=1.33$ ,  $p>.05$ ). Additionally, no interaction effect was found ( $F(1,33)=0.40$ ,  $p>.05$ ).

The differences in the strength of the correlation coefficients are shown in Table No. 2:

---

<sup>2</sup> There are beauty evaluations of 49 instructors. However, when we divided the lecturers' evaluations by students' gender, we had evaluations of 25 male instructors by male students and 25 male instructors by female students (but evaluations by male and female students for only 20 male instructors). There are evaluations of 16 female instructors by male students and of 17 female instructors by female students (but evaluations by male and female students for only 15 female instructors). For variance analysis only the data of the 20 male instructors and the 15 female instructors with evaluations by both female and male students could be used.

**Table No. 2: Correlation Coefficients between Instructors' Ratings and Instructors' Beauty, distinguishing between the gender of instructor<sup>3</sup> and student, N=31 male instructors, N=18 female instructors**

Instructor's gender Student's gender	Male	Female
Male	Beauty evaluation 0.19 R <sup>2</sup> = 0.04	Beauty evaluation 0.37 R <sup>2</sup> =0.13
Female	Beauty evaluation *** 0.77 R <sup>2</sup> =0.59***	Beauty evaluation 0.31 R <sup>2</sup> =0.10

\* p<.05 \*\* p<.01 \*\*\* p<.001

**The first study hypothesis:** The positive relationship between physical attractiveness and teaching ratings among male instructors will be found to be stronger in the female students' feedback than in that of the male students.

**Examination of the first study hypothesis:** According to this hypothesis, it would be expected that 1) no differences will be found between the means of the male and female students' ratings of the male instructors; and 2) the correlation coefficient between teaching ratings of male instructors and female students' ratings of instructors' beauty will be higher than this coefficient among male students' ratings.

**In order to test the first research hypothesis,** firstly differences between the means are examined. The ANOVA indicates no differences between male students' and female students' ratings ( $F(1,33)=1.33$ , n.s.). Now the correlation coefficient between the male instructors' teaching and beauty ratings calculated on the female students' data, is compared with the correlation coefficient calculated on the male students' data. The correlation coefficient calculated on the female students was 0.77\*\*\*, whereas the correlation coefficient calculated on the male students is 0.19, n.s.. Hence the relevant findings provide empirical support for the first hypothesis.

**The second research hypothesis:** The positive relationship between physical attractiveness and teaching ratings for female instructors will be found to be stronger in the male students' feedback than in that of the female students.

**Examination of the second research hypothesis:** According to this hypothesis, it would be expected that 1) no differences will be found between the means of the male and female students' teaching ratings of the female instructors, and 2) the correlation coefficient between the teaching ratings and the beauty ratings of female instructors by male student ratings' will be higher than this coefficient among female student ratings.

**In order to test the second research hypothesis,** firstly differences between the means are examined. The ANOVA indicates no differences between the male students' ratings and the female students' ratings ( $F(1,33)=1.33$ , n.s.). Now the correlation coefficients between the female instructors' teaching and beauty ratings, which were calculated on the female students' data, is compared to the correlation coefficient that was calculated on the male students' data. The correlation coefficient that was calculated on the female students is 0.31, n.s., and this coefficient on the male students is 0.37, n.s.. Hence the findings do not provide empirical support for the second hypothesis. No statistically significant correlation coefficients were found either in the female students' evaluations or in the male students' evaluations.

**To summarize the findings:** the first research hypothesis was corroborated while the second was not. Although no differences were found between the male and female students' ratings of the male instructors, a positive relationship was found between physical attractiveness and

---

<sup>3</sup> The correlation coefficient that was obtained without distinguishing between the students's gender: among the male instructors \*0.36, and among the females 0.31, n.s.

teaching ratings among male instructors, which is stronger in the ratings by female students than those of the male students. In the ratings of male instructors by female students, a correlation coefficient of 0.77 was found, and almost 60 percent of the variance of teaching rating is explained by beauty. In contrast, in the evaluation of the male instructors by male students, and in the evaluation of female instructors by students of both sexes, no statistically significant relationship was found, i.e. beauty evaluations do not contribute to explained variance in instructors' teaching ratings.

### ***Discussion***

**The first research hypothesis** claimed that the correlation coefficient between teaching ratings of male instructors and their beauty ratings by female students would be higher than this coefficient among male students. The data provide empirical support for this hypothesis. This sex difference may stem from women's self-perception as lacking expertise for judging beauty (Graziano, Jensen-Campbell, Shebilske and Lundgren, 1993). When asked directly they attach little importance to beauty, but the effect of beauty in their day-to-day interaction is evident from indirect examinations. Assayag (1998) also found that the beauty effect is weakened when the rater's gender is identical to that of the ratee. Assayag suggests that the reason for this is jealousy among people of the same sex. Another explanation of this is that all people are more attentive to the attractiveness of members of the opposite sex, and therefore the effect is weakened in the case of members of the same sex.

**The second research hypothesis** claimed that the correlation coefficient between teaching ratings of female instructors and their beauty ratings by male students would be higher than this coefficient among female students. The data do not provide empirical support for this hypothesis. This can be explained by the diminished importance that men attach to the physical attractiveness of women who are older than they are. When selecting a spouse, the norm in society is for the woman to be some years younger than the man; there are very few couples where the woman is older than her male partner. Men are more sensitive to the physical attractiveness of women who are younger than they are, and virtually indifferent to the appearance of women who are older than they are. For precisely this reason, the first study hypothesis was corroborated: in accordance with the norms of older males, women respond to the physical attractiveness of older males and give higher ratings to attractive male instructors. Based on the average ages of the raters (men 27.60 and women 27.37) and the ratees (men 48 and women 46), it should be noted that a woman of 50 is considered a mother figure and possibly even a grandmother figure for young Israeli males.

In the study conducted by Felton et al. (2004), a general correlation was found between the instructor's physical attractiveness and the ratings of his teaching,  $r = .30$ , and in a study conducted by Kindred et al. (2005), the correlation was  $r = 0.34$ . In our sample we obtained a general correlation of  $r = 0.35$ . However, unlike these studies that did not distinguish between the gender of the lecturer and the gender of the student, we did distinguish between them and our findings were surprising. When the gender of both instructors and students are taken into account, interaction between the variables is revealed: the correlation between beauty and the lecturer's ratings is 0.77 ( $p \leq 0.001$ ) when the evaluated instructors are male and the raters are female students. Moreover, the value of the correlation declines to the point of statistical insignificance in the remaining situations (male instructors and male students, and female instructors and male or female students). In this study, the correlation coefficient obtained among male instructors was  $*0.36$ , and among women 0.31, n.s. (when the student's gender is not taken into consideration). The findings of the present study provide empirical support for the findings in the study conducted by Hamermesh and Parker (2005), and indicate once again that the beauty premium has an effect mainly on the male instructors. The regression coefficient obtained in their study among male instructors was 0.384, and among women 0.128. Sussmuth (2006) found a weaker effect by approximately half than that found among the Americans findings of Hamermesh and Parker (2005). Sussmuth hypothesizes that the reason for this is that the German students are older than the Americans by an average of two years. In the Israeli sample the estimated<sup>4</sup> mean age was 27.78,<sup>5</sup> several years older

---

<sup>4</sup> The mean is estimated and not precise because student indicated age group rather than precise age.

than the American sample and the German one, and nevertheless the impact that was obtained for the lecturer's beauty was highly significant.

Romano and Bordieri (1989) found that attractive professors and female professors received the highest teaching scores. In their study, the interaction of instructor's sex and physical attractiveness was non-significant, which means that physical attractiveness had a main effect on teaching scores. Findings of the present study, however, point to an interaction effect: only attractive male instructors received higher ratings. Attractive female instructors do not receive higher ratings.

The findings of this study are consistent with Assayag's study (1998), about the reduced beauty effect when rater and rate are of the same sex. Assayag also reviewed studies on sex-based differences on the degree of the beauty effect, and found that such differences were found to be more salient in studies that were conducted by men, while female researchers (excluding Assayag 1998) did not find sex differences in this context. The findings of the present study, conducted by two female researchers, cast doubt on Assayag's contention (1998). These findings are consonant with the findings of Holahan and Stephan (1981) that attractive women are not rated as talented, particularly by men who adhere to traditional gender stereotypes. These findings provide additional evidence for the conclusion drawn by Cash and Trimer (1984), whereby feminine beauty confers less of a benefit in relation to the performance of masculine tasks.

In the studies by Hamermesh and Parker (2005) and Sussmuth (2006), beauty rankings were obtained from people who did not take a course given by the lecturer (Hamermesh and Parker presented pictures of the instructors to six raters in order to evaluate their beauty; in Sussmuth's study, beauty ratings were obtained from students at a different university who were also unacquainted with the instructors). In this study the correlation coefficients were calculated based on the beauty ratings that were obtained from each instructor's students, and based on teaching ratings that were obtained from students who were clearly exposed to the instructor's physical attractiveness during the course of their studies. This procedure is a better representation of reality, hence is preferable to the others, because in reality student's evaluations of beauty affect his or her teaching quality ratings and vice versa: The student's acquaintance with the lecturer's character and attitude affect his perception of his beauty.

It is possible that a self-fulfilling prophecy occurs; students expect attractive faculty to be warmer, more sensitive, available, knowledgeable, etc. Instructors unconsciously absorb these expectations and act according to them. If this self-fulfilling prophecy does occur, it will be found among both female and male faculty. Thus this prophecy cannot explain why male but not female instructors benefited from the beauty effect.

The main conclusion of the study is that similarly to administrative and other professions, males in academy also enjoy a "beauty premium," while women do not. This "discrimination" stems from the contradiction between two images: role images and gender images. As Gillen (1981) demonstrated, beautiful people are perceived as more characteristic of their sex, i.e. a beautiful woman is perceived as more feminine and a man as more masculine. When the role image corresponds to the gender image, one can expect the "beauty effect" to benefit beautiful people. However, when such correspondence is absent, as is the case with female instructors, the beautiful person (and, in our case, the beautiful woman) does not enjoy the "beauty premium."

Gillen (1981) suggested that attractiveness enhances gender characterizations, thus an attractive female professor is perceived to be more feminine and an attractive male professor is viewed to be more masculine than their less attractive colleagues. As mentioned in the introduction section, scientific and academic abilities are stereotyped as masculine (Ekehammer, 1985; Toren, 2005). This is another reason why female instructors do not enjoy the "beauty premium." Attractive female instructors work at a masculinely sex-typed job, and

---

<sup>5</sup> The age mean is high since all Jewish high school graduates in Israel must serve in the military immediately after graduation: men serve for 3 years and women for 2 years. Additionally, in the wake of a law that was passed 11 years ago, there is a large number of older students aged 30-50.

their exaggerated feminine attributes are incongruent with those believed necessary for their job<sup>6</sup>. Thus, it may be suggested that when an attractive female instructor receives high ratings, this is not **because** of their appearance, but rather **despite it!**

As to the sex differences in the degree of the beauty effect on female and male students, it is conceivable that men and women differ in their role expectations. Women may be more tolerant to a discrepancy between female instructors' gender image and their role characterizations, but their overall tendency<sup>7</sup> to award higher teaching ratings offsets the beauty premium that they would "award" to more attractive female instructors. This offsetting, combined with jealousy of beautiful women, leads to the absence of a "beauty effect" benefit for female instructors.

Furthermore, women may be less influenced by attractiveness of same-sex objects in forming gender characterizations. Men, on the other hand, may be more influenced by attractiveness in forming such characterizations, and less tolerant to a discrepancy between their gender image and role expectations, thus "punishing" the attractive female instructors.

Following Hamermesh and Parker (2005) and Sussmuth (2006), we took into consideration a self-selection bias: instructors who agreed to permit us ask students to rate their physical appearance during the teaching survey may be better teachers than those who did not agree to do so. In order to control for this bias, a comparison was made between the mean overall teaching evaluation of all the instructors at the college and the mean of our sample. No difference were found, thus it was concluded that the self-selection bias did not exist. Another consideration was taken into account regarding an additional type of self-selection bias: those who agreed to permit us ask students to rate their physical appearance may be better looking than those who did not agree to do so. Unfortunately, in order to control for this bias, beauty ratings for all of the instructors at the College were required, and this was obviously impossible (due to the ethical requirement to obtain the individual consent of all instructors). A self-selection bias of students may also have been at work, based on the reluctance of some students to respond to the additional question about beauty. It is possible that students who intentionally avoided answering this question were more sensitive to the issue of physical appearance. As a result, the correlation coefficients would have been greater had they answered. The size of the present quantitative sample, while useful for initial investigation, may need to be increased for further analysis. There may have been a bias in the beauty and teaching rankings, in the sense that students formed their opinion of the instructor even before ranking his or her beauty or teaching, due to previous information such as gossip or pedagogical reputation (mean grade level in the course). The findings of Cavior, Miller and Cohen (1975), that people judge others as more attractive if they have a longer acquaintance period,<sup>8</sup> should also be taken into consideration. Moreover, it has been found that longer acquaintance with the lecturer reduced the halo effect<sup>9</sup> in the teaching ratings made by students (Jacobs & Kozlowski, 1985).

In order to exclude the possibility that the effect is bidirectional, i.e. beauty affects the teaching ranking but the teaching method and knowledge of the lecturer's personality also affect the evaluation of his beauty,<sup>10</sup> it would be useful to investigate the relationship between beauty and evaluations at two points in time: Before the beginning of the semester – before students are personally acquainted with the instructor through the course – and again at the end of the semester, in order to examine whether acquaintance affects the beauty evaluation, either positively or negatively - in other words, whether the teaching ratings affect the beauty perception.

---

<sup>6</sup> This contradiction exists in addition to the contradiction between beauty and talent among women, which is raised in Holahan and Stephan's study (1981), cited above, which is stronger among men who endorse traditional gender stereotypes about women.

<sup>7</sup> As was found in the data for the college as a whole (Davidowitz, 2003)

<sup>8</sup> See note 4 above

<sup>9</sup> Halo effect is a strong tendency by the rater to think of the rate in general as rather good or rather inferior and to color the judgment of separate qualities by this general feeling (Thorndike, 1920).

<sup>10</sup> Proof of the fact that acquaintance with the lecturer could raise the perception of his beauty was found in our study: a negative correlation was found between attendance in the course and beauty evaluations ( $r=-0.36$ ,  $p<.01$ ). We found that higher the percentage of students who attended up to 40% of all class sections, the lower the instructors' beauty rating.

Finally, conducting a similar study in additional cultures is suggested, in order to examine whether cross-cultural differences exist regarding “sensitivity” to beauty level. Since the ethnic composition of the students is heterogeneous, and includes Jewish students of European and Eastern descent (some religious and some secular, some immigrants from the Former Soviet Union and Ethiopia), as well as Muslim and Christian Arabs, it is possible that there are stronger effects within some cultural groups, but that these were masked by the heterogeneous sample. Therefore it is advisable to study this subject, taking students’ cultural background into account. Similarly, a comparison should be made between the younger and older students, to see whether rater age constitutes a covariate for the beauty-teaching evaluation correlation. A further relevant question for future research is whether attractive male ratees would benefit the “beauty premium” in feminine-typed jobs, e.g. kindergarten teachers or hospital nurses.

In conclusion, this study confirms the findings of past studies on this issue, and sheds light on one variable that has not been addressed in previous studies, specifically the interaction between the professor’s gender and the student’s gender. This interaction is explained by confronting two images: gender stereotypes and role expectations. The combination of these concepts, aligned with sex differences of students in their tolerance to the discrepancy, provides insightful interpretation of the findings.

According to Baker (1984), the first step in escaping the beauty trap is to admit its existence. Hopefully, findings of the present study may help to increase public awareness of the subject and thereby improve the tools for evaluating instructors and other employees.

### **Acknowledgements**

We are grateful to Prof. Dan Zakay, Prof. Noach Milgram, Prof. Yochanan Peres, Dr. Avital Laufer, and Dr. Tali Lev for their valuable suggestions and comments.

### **Bibliography**

- Assayag, M., (1998), *The Impact of Candidates’ Attractiveness and Gender on their Salary and Enrollment*, Master dissertation submitted to Tel Aviv University, Israel (Hebrew).
- Bachen, C. M., McLoughlin, M. M., & Garcia, S. S. (1999). Assessing the role of gender in college students’ evaluations of faculty. *Communication Education*, 48 (3), 194-210.
- Baker, N.C., (1984), *The Beauty Trap-Exploring Woman’s Greatest Obsession*, NY: Franklin Watts.
- Bar-Tal, D., & Saxe, L., (1976), Physical Attractiveness and its Relationship to Sex-Role Stereotyping, *Sex Roles*, Vol. 2, pp. 123-133.
- Bull, R., & Rumsey, N., (1988), *The Social Psychology of Facial Appearance*, Springer-Verlag New York Inc.
- Buss, D.M., (1989), Sex Differences in Human Mate Preferences: Evolutionary Hypotheses Tested in 37 Cultures, *Behavioral and Brain Science*, Vol. 12, pp. 1-49.
- Cavior, N., Miller, K., & Cohen, S.H. (1975), Physical Attractiveness, Attitude Similarity, and Length of Acquaintance as Contributors to Interpersonal Attraction among Adolescents, *Social Behavior and Personality*, Vol. 3, pp. 133-141.
- Davidowitz, N., (2003), *Teaching with Quality: The Quality of Teaching in Judea and Samaria Academic College*, Judea and Samaria Academic College, Israel (Hebrew)
- Dion, K., Berscheid, E., & Walster, E., (1972), What is Beautiful is Good, *Journal of Personality and Social Psychology*, Vol. 24, pp. 285-290.

- Dipboye, R.L., Avrey, R.D., & Terpstra, D.E., (1977), Sex and Physical Attractiveness of Raters and Applicants as Determinants of Resume, *Journal of Applied Psychology*, Vol. 62, No. 3, pp. 288-294.
- Efran, M., (1974), The Effect of Physical Appearance on the Judgment of Guilt, Interpersonal Attraction, and Severity of Recommended Punishment in a Simulated Jury Task, *Journal of Research in Personality*, Vol. 8, pp. 45-54.
- Ekehammer, B., (1985), Women and Research: Perceptions, Attitudes and Choice in Sweden, *Higher Education*, Vol. 14, pp. 693-721.
- Feeley, T. H. (2002). Evidence of halo effects in student evaluations of instruction. *Communication Education*, 51 (3), 225-236.
- Feingold, A., (1990), Gender Differences in Effects of Physical Attractiveness on Romantic Attraction: A Comparison across Five Research Paradigms, *Journal of Personality and Social Psychology*, Vol. 59, pp. 981-993.
- Feingold, A., (1992), Good – Looking People Are Not What We Think, *Psychological Bulletin*, Vol. 111, Iss, 2, pp. 304-341.
- Felton, J, Mitchell, J., & Stinson, M. (2003, June). Web-based student evaluations of professors: The relations between perceived quality, easiness and "hotness." *Social Science Research Network Electronic Library*. Retrieved from <http://ssrn.com/abstract=426763>
- Gillen, B., (1981), Physical Attractiveness: A Determinant of Two Types of Goodness, *Personality and Social Psychology Bulletin*, Vol. 7, pp. 277-281.
- Hamermesh, D.S., & Biddle, J.E., (1994), Beauty and the Labor Market, *American Economic Review*, Vol. 84, No.5, pp. 1174-1194.
- Hamermesh, D.S., & Parker, A.M., (2005), Beauty in the Classroom: Instructors' Pulchritude and Putative Pedagogical Productivity, *Economics of Education Review*, Vol. 24, pp. 369-376.
- Hatfield, E., & Sprecher, S., (1986), *Mirror, Mirror...: The Importance of Looks in Everyday Life*, Albany: State University of New York.
- Heilman, M.E., & Saruwatari, L.R., (1979), When Beauty is Beastly: The Effects of Appearance on Sex on Evaluations of Job Applicants for Managerial and Non managerial Jobs, *Organizational Behavior and Human Performance*, Vol. 23, pp. 360-372.
- Helmes, J., & Rode, L., (1982), Effects of Counselor's Age and Physical Attractiveness on Evaluation of Female Counselors, *Journal of College Student Personnel*, Vol. 23, pp. 418-424.
- Holahan, C.K., & Stephan, C.W., (1981), When Beauty isn't Talent: The Influence of Physical Attractiveness, Attitudes toward Women, and Competence on Impression Formation, *Sex Roles*, Vol. 7, No. 8, pp. 867-876.
- Irilly, J., (1978), Student's Expectations: Ratings of Teacher Performance as Biased by Teacher's Physical Attractiveness, Paper presented at the Annual Meeting of the American Educational Research Association, Toronto.
- Jackson, L.A., Hunter, J.E., & Hodge, C.N., (1995), Physical Attractiveness and Intellectual Competence: A Meta-Analytic Review, *Social Psychology Quarterly*, Vol, 58, No. 2, pp. 108-122.

- Kindred, R., & Mohammed S. (2005). "He will crush you like an academic ninja!": Exploring teacher ratings on ratemyprofessors.com. *Journal of Computer-Mediated Communication*, 10(3), article 9, <http://jcmc.indiana.edu/vol10/issue3/kindred.html>.
- Landy, D., & Aronson, N., (1969), The Influence of the Character of the Criminal and his Victim on the Decisions of Simulated Jurors, *Journal of Experimental Social Psychology*, Vol. 5, pp. 141-152.
- Landy, D., & Sigall, H., (1974), Beauty is Talent: Task Evaluation as a function of the performer's Physical Attractiveness, *Journal of Personality and Social Psychology*, Vol. 29, pp. 299-304.
- Lewis, K.N., & Walsh, W.B., (1978), Physical Attractiveness: Its Impact on the Perception of a Female Counselor, *Journal of Counseling Psychology*, Vol. 25, pp. 210-216.
- Lombardi, J., & Tocci, M., (1979), Attribution of Positive and Negative Characteristics of Instructors as a Function of Attractiveness and Sex of Instructor and Sex of subject, *Perceptual and Motor Skills*, Vol. 48, pp. 491-494.
- Messer-Yaron, H. & Cahanovich, S. (2003) *Women and Science in Israel – a Snapshot*, The Council for the Advancement of Women in Science and Technology, Ministry of Science, Israel (Hebrew).
- Morrow, P.C., McElroy, J.C., Stamper, B.G., & Wilson, M.A., (1990), The Effects of Physical Attractiveness and Other Demographic Characteristics on Promotion Decisions, *Journal of Management*, Vol. 16, No. 4, pp. 723-736.
- Naumann, J.J., (1988), *The Effects of Teacher Expressiveness, Physical Attractiveness and Self-Disclosure on Student Ratings of Teaching*, Doctoral Dissertation submitted to Fordham University.
- Newsum, R.L., (1990), *Surgency Bias: Effects on Teacher Performance Evaluation*, Doctoral Dissertation submitted to Iowa State University.
- O'Reilly, M.T., (1987), Relationship of Physical Attractiveness to Students' Ratings of Teaching Effectiveness, *Journal of Dental Education*, Vol. 51, No. 10, pp. 600-602.
- Romano, S.T., & Bordieri, J.E., (1989), Physical Attractiveness Stereotypes and Students' Perceptions of College Professors, *Psychological Reports*, Vol. 64, Iss. 3, Part 2, pp. 1099.
- Stewart, J., (1980), Defendant's Attractiveness as a Factor in the Outcome of Criminal Trials: An Observational Study, *Journal of Applied Social Psychology*, Vol. 10, pp. 348-361.
- Sussmuth, B., (2006), Beauty in the Classroom: Are German Students Less Blinded? Putative Pedagogical Productivity due to Professors' Pulchritude: Peculiar or Pervasive?, *Applied Economics*, Vol. 38, pp. 231-238.
- Thorndike, E.L., (1920), A Constant Error in Psychological Ratings, *Journal of Applied Psychology*, Vol. 4, pp. 25-29.
- Toren, N. (2005), *Women in Israeli academy: Images, numbers, discrimination*, Ramot Publishing House, Tel Aviv University (Hebrew).
- Weiss, M., (1991), *Conditional Love: The Handicapped child as Perceived by his Parents*, Tel Aviv, Hapo'alim Pub. Co.(Hebrew).
- Wikipedia, (2008), [http://en.wikipedia.org/wiki/Physical\\_attractiveness](http://en.wikipedia.org/wiki/Physical_attractiveness)

**Appendix A:**

**Differences in of instructors' beauty ratings by instructor's gender and student's gender**

To examine these differences, a bidirectional ANOVA was performed on the lecturer's beauty ratings by the lecturer's gender and the student's gender, with repeated measurements for the student's gender.

**Table 5: Means and Standard Deviations of the Beauty Evaluations by the Lecturer's Gender and the Student's Gender**

The lecturer's gender The student's gender	Male instructors (n=22)		Female instructors (n=13)		Total (n=35)	
	Mean	SD	Mean	SD	Mean	SD
Male students	3.90	0.73	3.42	0.74	3.66	0.76
Female students	3.42	0.94	3.92	0.79	3.67	0.90
Total	3.66	0.72	3.67	0.63	3.67	0.68

It is evident from Table 5 and from the results of the ANOVA that no differences were found in the lecturer's beauty ratings, neither according to the lecturer's gender ( $F(1,33)=0.002$ ,  $p>.05$ ) nor according to the student's gender ( $F(1,33)=0.01$ ,  $p>.05$ ). However, statistically significant interaction according to the lecturer's gender and the student's gender was found ( $F(1,33)=12.74$ ,  $p>.001$ ) It was found that the female students rate the beauty of the female instructors as higher ( $M=3.92$ ) than the beauty of the male instructors ( $M=3.42$ ). Moreover, the male students rate the beauty of the male instructors more highly ( $M=3.90$ ) than the beauty of the female instructors ( $M=3.42$ ). No differences were found between male students and female students in the evaluation of the male instructors' beauty and no differences were found between male students and female students in the evaluation of the female instructors' beauty.