

Gender differences in students' ratings of university teachers in the Iranian education system

Abbas Zare-ee

University of Kashan, Iran

Zuraidah Mohd Don

University of Malaya, Malaysia

Iman Tohidian

University of Kashan, Iran

Abstract

University students' ratings of teaching and teachers' performance are used in many parts of the world for the evaluation of faculty members at colleges and universities. Even though these ratings receive mixed reviews, there is little conclusive evidence on the role of the intervening variable of teacher and student gender in these ratings. Possible influences resulting from gender-related differences in different socio-cultural contexts, especially where gender combination in student and faculty population is not proportionate, have not been adequately investigated in previous research. This study aimed to examine Iranian university students' ratings of the professional performance of male and female university teachers and to explore the differences in male and female university students' evaluation of teachers of the same or opposite gender. The study was a questionnaire-based cross-sectional survey with a total of 800 randomly selected students in their different years of undergraduate study (307 male and 493 female students, reflecting the proportion of male and female students in the university) from different faculties at the University of Kashan, Iran. The participants rated male and female teachers' performance in observing university regulations, relationship with colleagues, and relationships with students. The researchers used descriptive statistics, means comparison inferential statistics and focus-group interview data to analyze and compare the students' ratings. The results of one-sample t-test, independent samples t-test, and Chi-square analyses showed that a) overall, male university teachers received significantly higher overall ratings in all areas than female teachers; b) male students rated male teachers significantly higher than female students did; and c) female students assigned a higher overall mean rating to male teachers than to female teachers but this mean difference was not significant. These results are studied in relation to the findings in the related literature and indicate that gender can be an important intervening variable in university students' evaluation of faculty members.

Introduction and background

The evaluation of the effectiveness of higher education and the performances of university teachers is an enormous task because of the multidimensionality of the teaching/learning processes and the involvement of individually and socially complicated human beings. In spite of this complexity, university teachers' performances are constantly evaluated based on difference criteria to increase the standards of higher education, professional development, promotion, payment, tenure, grants and other similar

good intentions. Formative and summative evaluations of teachers' performance take place in different ways in different institutions of higher education (Peterson, 2000). Students' ratings, usually reflected in their responses to questionnaire items, are one of the widely used and common forms of evaluation. Student evaluations are a primary measure of the teaching performance of college and university faculty members in the US (Campbell, Steiner, & Gerdes, 2005, p. 211), Europe and many other parts of the world (Basow, 2000; Radmacher & Martin, 2001).

In the Iranian system of university teacher evaluation, for example, students are required to respond to a teacher evaluation questionnaire constructed and approved by the Ministry of Science, Research, and Technology (MSRT) at the end of each semester. The questionnaire includes items on the teacher's appearance, his/her academic competence, his/her teaching skills, and his/her class behavior. Data collected through this questionnaire are stored, analyzed, and used for the purposes mentioned above (see Ahanchian, 2003). It should be noted that the teaching profession in higher education in Iran is predominantly male dominated at present even though the population of female students is significantly higher than male students. Based on the latest formal statistics from the Institute of Research and Planning in Higher Education (IRPHE, 2010) a total of around 21,000 (12 percent female and 88 percent male) university teachers (lecturers, assistant professors, associate professors, and professors teach at different public institutions of higher education controlled by the Iranian Ministry of Higher Education (MSRT). These teachers teach a total of 1713652 students attending public colleges and universities including 668588 (39 percent) male and 1045064 (61 percent) female students. The total number of teachers and students in all ranks is much bigger in the private section of Iranian higher education but reflects almost the same gender compositions-the male dominated teacher population and the female dominated student population.

This combination of the teaching staff is very different from other countries. According to Carrington et al. (2007), in recent years, policy-makers in England, Australia and some other countries have called for measures to increase male recruitment to the teaching profession in compulsory schooling because of the lower number of male teachers and beliefs in the unexamined assumptions of the benefits of matching teachers and pupils by gender.

Under such circumstances, students' ratings have great consequences for university teachers, decision makers, and students. Teachers may be underpaid or demotivated as a result of students' ratings. In fact, tenure and promotion decisions are among the most common applications of the results of students' evaluation of university teachers (Dukes & Victoria, 1989; Schwier, 1982; Seldin, 1989; Tieman & Rankin-Ullock, 1985). Opinions about the value of student ratings are divided. Some policy makers may receive faulty information communicated through biased students' evaluation of teachers. As students' ratings may be affected by irrelevant variable such as grades or gender, they may have unintended adverse consequences such as grade inflation referred to by Langbein (2007), as student rating can give faculty a clear incentive to inflate grades to get higher course ratings.

On the other hand, some previous research shows that student ratings of teachers may be – at least to some extent – reliable, valid and generalizable (Craig, Redfield, & Galluzzo, 1986; Konrad, 1986; Schwier, 1982). Commenting on the related body of research from North America, Morgan and Davies (2006) conclude that

student ratings of instructor effectiveness contained some valid information even if imperfectly measured and imperfectly defined. (p. 3)

Marsh (1984) and Cashin (1988) also contend that students' ratings of university teachers are positively correlated with teachers' self-rating of teaching effectiveness and therefore students' ratings are

dependable. However, there are also many studies (e.g. Langbein, 2007; Andersen & Miller, 1997) that cast doubts on the validity and reliability of students' evaluation of university teachers or at least do not view these ratings as the only source of evaluation data because of the intervening variables that may affect them.

Many variables are mentioned that might affect students' rating of university teachers. Grades that students receive from their teachers are an important variable in many contexts. According to Andersen and Miller (1997), students' expectations can also affect their teacher evaluation. The way they are treated by the teacher, the teacher's personality variables, affiliation, rank, and many other factors may affect students' evaluations of teachers. One possible intervening variable in students' evaluation of teachers in higher education is gender. Teaching effectiveness may have the greatest influence on students' ratings; however, gender bias is also an important part of evaluating teachers by their students (Dukes and Victoria, 1989). So, the central purpose of this study is to explore the role of gender in students' evaluation of teachers in higher education.

Gender and university students' evaluation of teachers

Gender bias in students' rating of university teachers has been the focus of many studies and the results have so far been rather controversial and contradictory. Siskind and Kearns (1997) contend that previous research has come up with drastically different results on the role of gender in students' evaluation of faculty members at colleges and universities. As summarized in the following paragraphs, previous research has shown results as varied as no difference in overall students' ratings of male and female faculty; male faculty superiority in students' ratings; female faculty superiority; same-gender rating superiority; opposite-gender superiority; male superiority in some dimensions of teaching effectiveness; and female superiority in some dimensions (See also Huston, 2005).

Some researchers, such as Hamermesh and Parker (2005) and Miller and Chamberlin (2000), found that female faculty members received lower total course evaluation ratings. Similarly, Burns-Glover and Veith (1995) found that university students who rated the suitability of three fictitious applicants in terms of behaviors and traits necessary for "great professors" showed clear biases in their ratings. All student raters favored male teacher traits in this study and male student raters showed greater bias toward male teachers. In another study, Basow and Silberg (1987), who studied the ratings of over 1,000 undergraduates for 16 male and female professors in terms of teaching effectiveness and sex-typed characteristics, found that male students gave female professors significantly poorer ratings than male professors on teaching evaluation measures. Female students also evaluated female professors less favorably than male professors on some measures. This clearly points to a less favorable condition for female teachers in general in this group of studies.

Basow (1998) argued that even though the overall effect of gender on the students' evaluations of teachers may not be great (only 3 percent of the variance observed in this study), in interaction with other related variables, this bias can eventually make student rating biased against female teachers. In a study by Bennett (1982), it was concluded that students form their own criteria for rating male and female teachers expecting higher standards from female professors than male professors. The study showed that students used similar criteria for rating male and female instructors; however, female teachers were seen as charismatic but not authoritarian with higher ratings in terms of professionalism, presentation skills, and accessibility.

In contrast with the studies that report higher students' ratings for male teachers, there is a second group of studies whose findings point to higher ratings for female teachers. Results of some of these

studies show that both male and female students tend to rate their female teachers higher (Costa, Terracciano & McCrae, 2001; Kierstead, D'Agostino & Dill, 1988). On the other hand, some studies show that only female students rate female teachers higher. For instance, Bachen, McLoughlin, and Garcia's (1999) found that compared with male teachers, female teachers attained higher rankings from female students. In this study male student raters did not rate female and male teachers significantly differently. The reported bias was attributed partly to students' gender schema referring to "basic cognitive structures concerning gender that help guide their perceptions, interpretations, and recall of people and events" (p. 193). This study used students' self-reported perceived differences between male and female faculty members as "general cases" (p. 198) rather than actual teachers which might reflect schema-affected evaluations.

In a third group of studies, statistically significant differences in evaluations for male and female instructors have not been reported (Ahanchian, 2003; Centra & Gaubatz, 2000; Feldman, 1993; Franklin & Theall, 1992; Santhanam & Hicks, 2002). Ahanchian (2003) studied 371 students using the teacher evaluation questionnaire approved by the Iranian Ministry of Science, Research and Technology (MSRT) and administered by the evaluation unit of public universities and added up scores on all its items without considering individual items or the classification of items to the aspects of teachers' behavior that they measured. He reported no significant effect for student gender. Similarly, Hameed, Ali, Hameed, Saleem, and Javed (2015) studied the relationship between gender and teacher evaluation in the case of 250 students from four different universities of Pakistan and found that teachers were evaluated equally by the male and female students in terms of organizing and planning courses and communication skills. In students rating of teaching quality at the university level, Fernandez and Mateo (1997) found no relationships between ratings and teachers' or students' gender. Feldman (1992, 1993) argued that even though a small number of related studies show higher student ratings for male university teachers, global evaluations of male and female teachers are not radically different.

Same-gender preferences in students' rating of teachers have also been noted in a number of studies, most of which show that female university student raters rank their female instructors higher (Basow & Silberg, 1987; Feldman, 1993; Centra & Gaubatz, 2000). Examining the effects of gender on students' evaluation of teaching at university, Centra and Gaubatz (2000) found that female instructors obtained higher ratings on some scales by their female students, while their male colleagues attained higher ratings by male students on just one scale of course organization and planning, showing that same-gender preferences were stronger for female students. Similarly, Basow (1995) reported that female members of university received higher ratings from their female students and lower ratings from their male students. A somewhat different observation was reported in a study by Lueck, Endres, and Caplan (1993). This study showed that both male and female students assigned higher ratings to their same-gender instructors. Overall, there seems to be less empirical support for same-gender preferences in the case of male students.

Research on teacher evaluation has shown that teachers may be rated by their students based on how they are perceived in terms of warmth, interpersonal contact, and interest in their students (Kierstead, D'Agostino, & Dill, 1988; Bennett, 1982; Downs & Downs, 1993). These variables may affect female teachers more seriously in societies where the norms restrict their contact with the opposite sex. Biased male and female students' evaluation of teachers may also interact with other variables such as course level, discipline, or dominance in male-dominated or female-dominated contexts of education. Previous experience with teachers has also been shown to increase their ratings in the case of female teachers (Fandt & Stevens, 1991).

What is evident in the research literature is that gender of the students and teachers can influence students' teacher evaluation. Conclusions about the directions of effect are, of course, difficult to draw because many variables are at work and major studies from some contexts are lacking. Clayson (2013) found significant links between end-of-the-semester teacher evaluations and students' first impressions of the instructor and the instructor's personality. Students' expected grades, the level of study, and instructor's status also need to be taken into account in relation to students' ratings of instructors (Goldberg & Callahan, 1991). Kierstead, D'Agostino, and Dill (1988) studied the effects of variables such as social contacts between instructor and students, instructor's facial expressions, and instructor's gender on students' evaluation of teaching. They reported that instructor sex interacted with the other variables (student-instructor contacts and facial expressions). They argued that students' higher expectations of female instructors' performance led to their lower rating. That is, a male and a female teacher who are of the same standard might receive different ratings just because of their students' expectations.

Student groups in higher education are not homogenous and may therefore evaluate teachers differently using different evaluation instruments. The research reviewed above indicates that great differences in rating faculty might be due to variables such as the gender of the students and/or the gender of the teachers. As the review shows, the criteria that students form in their minds about teachers, their gender schema, their expectations from their teachers, and their sense of identity with their teachers as in reported higher same-gender evaluations underline the undeniable role of gender in students' evaluations of teachers. Since these factors are culture and context dependent, data from various socio-cultural contexts can contribute to the better understanding of the role of gender in teacher evaluation. The purpose of the present research was to explore possible gender-related differences in teacher evaluation by students in an Iranian context. In previous research examining gender-related differences in students' teacher evaluation, students have been asked to rate real (Bennett, 1982), hypothetical (Rubin, 1981), or general case (Bachen, McLoughlin, & Garcia, 1999) male and female teachers on different instruments eliciting information on various teaching behaviors, social contacts, and traits. This study asked male and female students to rate actual specific teachers who taught the evaluators rather than collecting data on the evaluation of hypothetical teachers or general cases. There is actually no agreement on the construct validity of the instruments administered to measure students' ratings of instructional effectiveness or professional performance of teachers. Clayson (2014), for instance, found major evidence that student evaluations on *ratemyprofessors.com* were "biased by a halo effect" in measuring teaching effectiveness. On the other hand, as Clayson admits,

to say that SET [students' evaluation of teaching] is invalid as a measure of effective teaching does not mean that the instruments do not validly measure something else. (p. 696).

The main reason for this problem is that the instruments used in previous research are context- and purpose-specific. The instrument constructed for data collection in the present study elicited data on how male and female teachers were evaluated in terms of observing university regulations, relationships with colleagues, and with students. The findings can, therefore, be used to complement previous research findings. The study specifically addressed the following research questions:

1. Are male and female university teachers rated differently in terms of observing university regulations, relationship with colleagues and with their students?
2. Do male and female students rate same-gender and opposite gender teachers differently in terms of observing university regulations, relationship with colleagues and with their students?

Methods

Participants

The researchers administered the questionnaire to a randomly selected sample of 1000 university students at the faculties of humanities, science, engineering, and arts at the University of Kashan. The purpose was to select data from a large-enough sample that adequately represented the total population of undergraduate students (n= 6000 students taught by a total of 300 teachers). Only participants from the second, third, and fourth year were selected to make sure they had enough contact with instructors to be able to complete the evaluation questionnaire. Because of the much lower number of female teachers at the university (10%), the researchers made sure that each questionnaire was completed by a participant who had academic interactions with female teachers as well. Some targeted students had left many items of the questionnaire unanswered (including gender specification) or taken the questionnaire away. The researchers, therefore, decided to stop data collection once 800 completed questionnaires had been accumulated (307 male and 493 female 38.4% and 61.9% respectively).

Data collection instruments

The main instrument of the study was a researcher-created questionnaire and the research questions were answered based on the data collected through this questionnaire. Even though the study did not contain any significant qualitative part, interview with a sub-sample was also considered for further clarification of the quantitative data provided by the whole sample. To measure the participants' views about male and female university teachers' professional performance in terms of observing university regulations, relationships with colleagues, and relationships with students, a five-point Likert scale questionnaire was developed based on an examination of the contents of items on the questionnaire given to Iranian university students at the end of each semester for teacher evaluation by the Iranian Ministry of Science, Research, and Technology. The researcher also used the University of Texas guidelines for Intern teacher/student self-evaluation as well as comments by the colleagues in the preparation of the questionnaire. For validity considerations, the items of the questionnaire were written in the respondents' mother tongue, Persian, to make sure that they were fully comprehensible for all the participants. The questionnaire was also pilot tested with a group of 35 students in a class to revise and modify ambiguous items and to prepare the final form of the questionnaire.

The final version of the questionnaire (Appendix A) consisted of two parallel sets of 24 items that elicited information on students' ratings of their male and female teachers' professional behavior (in terms of observing university regulations, relationships with colleagues, and with students). Students were asked to assign overall ratings based on all the teachers they had had at the university. Each item/statement was mentioned twice: once for male, and once for female professors. Items were written in five-point Likert scale ranging from a score of 1 to 5, with 1 representing "never" and 5 representing "always". The questionnaire included ten items on professors' observing university regulations (items 1-10), seven items on their relationships with colleagues (items 11-17), and seven items on their relationships with students (item 18-24). The Cronbach alpha reliability for the questionnaire was 0.93 (number of items=48). Back-translation by an expert colleague at the English Department was used to validate the wording and translation of items administered in the participants' mother tongue, Persian.

Data analysis method

The data for this research were collected over a period of 2 months in early 2010 with the help of student assistants using three methods of questionnaire distribution and collection: through students' institutional emails, through the university postal service for university accommodation and through direct contacts in classrooms. When the data on students' ratings of male and female teachers were collected, they were checked and entered into SPSS data files. These ratings were analyzed and the results were compared using descriptive and means comparison inferential statistics of one-sample t-test, independent samples t-test, and Chi-square analyses. The findings were finally checked with the common themes extracted from a focus-group interview with 20 randomly selected participants during which they expressed their reasons for choosing either male or female instructors as better teachers.

Results

During the analyses of the data and for better classification and summarization purposes, mean ratings of both male and female teachers in three subgroups of items were considered apart from the analyses of means on each individual item. The new means were the new dependent variables in the data. These included observing university regulations (10 items), relationships with their colleagues (7 items), and on relationships with students (7 items). Initial descriptive statistics showed that male university teachers achieved higher means than female teachers on all the items and as Table 1 shows, male university teachers were rated higher in observing university regulations, relationships with their colleagues, and relationships with students. Their total scores were also higher than those for female teachers.

Table 1: Total ratings for male and female university teachers by all students.

| Variables | Male teachers | | Female teachers | |
|------------------------------------------------------------------|---------------|------|-----------------|------|
| | Mean | SD | Mean | SD |
| University professors' mean score: observing regulations | 3.34 | .745 | 2.98 | .969 |
| University professors' mean score: relationships with colleagues | 3.34 | .762 | 2.98 | .828 |
| University professors' mean score: relationships with students | 3.26 | .753 | 2.89 | .841 |
| Total scores | 9.94 | 2.05 | 8.85 | 2.4 |

One-sample T-test was performed in order to check the statistical significance of these differences between the means produced by male and female teachers and the results summarized in Table 2 indicate that all were significant: ratings in each category and in total were significantly higher for male teachers. The first research question of the study therefore received a positive response, that is male and female university teachers receive significantly different ratings.

Table 2: One-sample t-test comparing total mean ratings for male and female university teachers (df=799).

| Variables | Male | | Female | |
|------------------------------------------------------------------|---------|------|---------|------|
| | t | Sig | t | Sig |
| University professors' mean score: observing regulations | 126.920 | .000 | 86.999 | .000 |
| University professors' mean score: relationships with colleagues | 124.006 | .000 | 101.802 | .000 |
| University professors' mean score: relationships with students | 122.356 | .000 | 97.341 | .000 |
| Total scores | 136.819 | .000 | 103.435 | .000 |

Male and female students' ratings of males and female teachers were also looked at separately to see whether students rated same-gender and opposite gender teachers differently in terms of observing

university regulations and relationship with students and relationship other professors. Descriptive statistics summarized in Table 3 show that male students consistently rated female teachers lower than male teachers. The mean rating of female teachers by male students was 7.85 in contrast to a mean of 9.5 assigned to them by their same-gender students. Female teachers were rated lower by male students in all the three areas of observing university regulations, relationships with their colleagues, and relationships with students. Even though female students assigned a higher total mean rating to male teachers than to female teachers, this mean difference was not significant. Therefore, it was down to the male students that male professors were rated more highly than female professors overall.

Table 3: Mean ratings for male and female university teachers across student genders (male 307, female 493).

| Variables | Student Gender | Female Professors | | Gender | Male Professors | |
|--------------------------------------------------------|----------------|-------------------|-------|--------|-----------------|-------|
| | | Mean | SD | | Mean | SD |
| Professors' mean scores: observing regulations | Male | 2.57 | .782 | Male | 3.36 | .633 |
| | Female | 3.23 | .989 | Female | 3.34 | .808 |
| Professors' mean scores: relationships with colleagues | Male | 2.65 | .696 | Male | 3.35 | .630 |
| | Female | 3.18 | .838 | Female | 3.33 | .834 |
| Professors' mean scores: relationships with students | Male | 2.62 | .714 | Male | 3.25 | .680 |
| | Female | 3.06 | .870 | Female | 3.26 | .796 |
| Total score | Male | 7.85 | 1.952 | Male | 9.96 | 1.739 |
| | Female | 9.48 | 2.473 | Female | 9.93 | 2.230 |

Comparisons of means for teachers by gender across the two groups of male and female students indicated that female teachers are rated significantly lower by male students ($p= 0.001$). As Table 4 shows, the mean differences in the ratings received by male and female teachers are significant.

Table 4: Independent samples t-test comparing teachers' means across student genders (male 307, female 493).

| Variable | Female | | | | | Male | | | | |
|------------------------------------------|--------|------|---------|--------|------|--------|------|------|--------|------|
| | F | Sig. | t | df | Sig. | F | Sig. | t | df | Sig. |
| Teacher observing university regulations | 45.053 | .000 | -9.901 | 798 | .000 | 24.262 | .000 | .393 | 798 | .694 |
| | | | -10.448 | 754.04 | .000 | | | .415 | 756.98 | .678 |
| Teacher relationships with colleagues | 18.193 | .000 | -9.359 | 798 | .000 | 27.452 | .000 | .311 | 798 | .756 |
| | | | -9.767 | 735.08 | .000 | | | .331 | 769.00 | .741 |
| Teacher relationships with students | 15.372 | .000 | -7.409 | 798 | .000 | 8.180 | .004 | - | 798 | .866 |
| | | | -7.753 | 739.83 | .000 | | | .168 | 723.23 | .861 |
| Total scores | 31.379 | .000 | -9.809 | 798 | .000 | 20.075 | .000 | .196 | 798 | .845 |
| | | | -10.355 | 754.52 | .000 | | | .207 | 758.79 | .836 |

To compare the differences in the expected frequencies of responses to items on the evaluation questionnaire and to confirm the results of the independent sample t-test, Chi square was also calculated, the result of which are summarized in Table 5. As the table shows, mean differences in the ratings of female teachers by male and female students are significant ($\chi^2=552.62, p\leq 0.05$)

Table 5: Chi square analyses.

| Test | Variable | Female | | | Male | | |
|--------------------|-----------------|---------|-----|------|---------|-----|------|
| | | Value | df | Sig | Value | df | Sig |
| Pearson Chi-Square | Teachers' score | 552.615 | 488 | .023 | 459.881 | 457 | .453 |

The data collected for the purposes of this study also included a recorded focus-group interview in which 20 students expressed their reasons for choosing either male or female teachers as professionally better. Although the analysis of the contents of this interview was not directly relevant to the research questions and was not used to test the research hypotheses, the themes appearing in its coding were in line with the quantitative findings. Based on the recurrent themes observed in both male and female students' responses, obstacles such as expected younger age of marriage, pregnancy, and tougher employment conditions were main obstacles for the career prospect of female teachers. Both male and female students mentioned that male teachers were expected to be more successful because they were taken more seriously than women, they were the responsible and careful bread-winners of the family, and because they acted and decided less emotionally. These recurrent themes in the coded interview data support the quantitative findings of the study that female students assign a higher total rating to male teachers.

Discussion

The data and the analyses in this study indicate that male university teachers are rated significantly higher by male student evaluators. Male and female university teachers are rated by students differently in terms of observing university regulations, relationship with students, and relationship with other professors and in all these areas male teachers receive higher ratings from the total population of students irrespective of same-gender preferences. These findings support the ones found by scholars like Hamermesh and Parker (2005) and Miller and Chamberlin (2000), who also found that female faculty members received lower total course evaluation ratings. In contexts such as Iran where teaching in higher education is still male dominated in spite of the greater population of female students, women teachers form a minority in the community of higher education teachers, which can affect their group dynamics. What Ahanchian (2003) found in the university of Mashhad in Northern Iran was not supported by this study. In spite of his strong arguments that student gender can be a source of bias in teacher evaluation, he found no difference between male and female students in their evaluations of their teachers. One strength of the present study was that it analyzed the evaluation of male and female teachers separately, a point which was overlooked in Ahanchian's study of the evaluation of Iranian university teachers by students. Male and female students' evaluation of same or different gender teachers may not be the same. The findings of this study showed that male students evaluated same gender (male) teachers significantly higher than they rated female teachers. A reanalysis of his data and data from other major universities will certainly shed more light on the issue. With the student population of over 60 percent female students and a teacher population of around 10 percent female teachers in Iranian public universities, with the student population of over 60 percent female students and a teacher population of around 10 percent female teachers in Iranian public universities, It seems that less bias should be exercised in offering employment conditions to women teachers in Iranian higher education. The findings of the study point to the fact that female students, who form the majority of the student population in the context of the study, do not make significantly different evaluation of same and different gender teacher. Therefore, equal employment opportunities for male

and female teachers in institutions of higher education can gradually improve the gender composition of the community of higher education teachers.

The study also showed that even though female students tend to rate same-gender teachers higher in some areas, they actually assign ratings to opposite gender teachers that are not significantly lower; whereas, male students rate female teachers significantly lower in all areas including observing university regulations, relationship with students, and relationship with other professors. In other words, in line with the findings of studies by Hamermesh and Parker (2005) and Miller and Chamberlin (2000) female faculty members receive lower total course evaluation ratings. However, these lower ratings are because their opposite gender students, male students, rate them lower.

In the context where this study was carried out, male students' evaluation significantly favored male teachers. However, female students' evaluation was not significantly different for same and different gender teachers. These findings will certainly not be final decisions unless replications in other universities with larger samples and national surveys confirm them.

Conclusions

Student ratings, now one of the most widely used sources of information on teaching effectiveness and teacher evaluation in higher education, are sometimes not welcome by teachers and scholars as they can be biased, are easy to abuse, and do not always lead to improvements in teaching. In using faculty evaluating systems, gender of the teacher and students has often been shown to be a relevant factor. The magnitude of the effect of this variable may vary based on the context, the role of the faculty, the expectation of the students, and many other variables. Student ratings alone can, therefore, not be considered enough in teacher evaluation (Seldin, 1993). Evaluators need to look at more sources of information and assign correct weightings to these sources of information (Arreola, 2000) and as Cashin (1995) argues, viewing student ratings as data rather than evaluations may help to put them in proper perspective. As our study indicates, male and female students may have different ratings of their same and different gender teacher. In other word, both the gender of the student evaluator and that of the teacher evaluated are factors worth considering. Viewed as data, students' ratings of their teachers can be cautiously used along with other kinds of data such as ratings by other faculty members and course results to make better evaluations of teachers in higher education. This can be one implication of the present study for the teacher Supervision evaluation units of Iranian universities known as Nezarat-Arzyabi office.

Acknowledgement

This study has been partially supported by grant number 9955-89/05/30 from the University of Kashan to the first author. The author is grateful to the Vice-Chancellor of Research of the university for support and encouragement and for offering leave to complete the study. Special thanks also go to student assistants who helped in the process of data collection for this research.

References

Ahanchian, M. R. (2003). The effects of same and opposite student gender on university faculty evaluation. *Journal of Psychology and Education*, 33(1), 183-199.

Zare-ee, A., Mohd Don, Z. & Tohidian, I. (2016). Gender differences in students' ratings of university teachers in the Iranian education system. *Learning and Teaching in Higher Education: Gulf Perspectives* 13(1).
<http://dx.doi.org/10.18538/lthe.v13.n1.232>

Anderson, K., & Miller, E. D. (1997). Gender and student evaluations of teaching. *PS: Political Science and Politics*, 30(2), 216-219. <http://dx.doi.org/10.2307/420499>

Arreola, R. A. (2000). *Developing a comprehensive faculty evaluation system: A handbook for college faculty and administrators on designing and operating a comprehensive faculty evaluation system (2nd ed.)*. Bolton, MA: Anker Publishing Company, Inc.

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), 193-210. <http://dx.doi.org/10.1080/03634529909379169>

Basow, S. A. (1995). Student evaluations of college professors: When gender matters. *Journal of Educational Psychology*, 87(4), 656-665. <http://dx.doi.org/10.1037/0022-0663.87.4.656>

Basow, S. A. (1998). Student evaluations: Gender bias and teaching styles. In L. H. Collins, Chrisler, J.C., & Quina, K. (Eds.), *Career strategies for women in academe: Arming Athena*. (pp. 135-156). Thousand Oaks, CA: Sage.

Basow, S. A. (2000). Best and worst professors: Gender patterns in students' choices. *Sex Roles*, 43(5-6), 407-417.

Basow, S. A., & Silberg, N. T. (1987). Student evaluations of college professors: Are female and male professors rated differently? *Journal of Educational Psychology*, 79(3), 308-314. <http://dx.doi.org/10.1037/0022-0663.79.3.308>

Bennett, S. K. (1982). Student perceptions of and expectations for male and female instructors: Evidence relating to the question of gender bias in teaching evaluation. *Journal of Educational Psychology*, 74(2), 170-179. <http://dx.doi.org/10.1037/0022-0663.74.2.170>

Burns-Glover, A. L., & Veith, D. J. (1995). Revisiting gender and teaching evaluations: Sex still makes a difference. *Journal of Social Behavior & Personality: Special Issue: Gender in the workplace*, 10(6), 69-80.

Campbell, H., Steiner, S., & Gerdes, K. (2005). Student evaluations of teaching: How you teach and who you are. *Journal of Public Affairs Education*, 11(3), 211-231.

Carrington B., Francis, B., Hutchings, M., Skelton, C. Read, B. & Hall, I. (2007). Does the gender of the teacher really matter? Seven- to eight-year-olds' accounts of their interactions with their teachers, *Educational Studies*, 33(4), 397-413.

Cashin, W. E. (1988). *IDEA technical report no. 20: Student ratings of teaching: A summary of the research*. Manhattan: Kansas State University, Center for Faculty Evaluation and Development.

Cashin, W. E. (1995). Student ratings of teaching: The research revisited. *IDEA paper No. 32*. Retrieved January 23, 2015 <http://eric.ed.gov/?id=ED402338>

Centra, J. A., & Gaubatz, N. B. (2000). Is there gender bias in student evaluations of teaching? *The Journal of Higher Education*, 71(1), 17-33. <http://dx.doi.org/10.2307/2649280>

Clayson, D. E. (2013). Initial impressions and the student evaluation of teaching. *Journal of Education for Business*, 88(1), 26-35. <http://dx.doi.org/10.1080/08832323.2011.633580>

Clayson, D. E. (2014). What does ratemyprofessors.com actually rate? *Assessment & Evaluation in Higher Education*, 39(6), 678-698. <http://dx.doi.org/10.1080/02602938.2013.861384>

Zare-ee, A., Mohd Don, Z. & Tohidian, I. (2016). Gender differences in students' ratings of university teachers in the Iranian education system. *Learning and Teaching in Higher Education: Gulf Perspectives* 13(1). <http://dx.doi.org/10.18538/lthe.v13.n1.232>

- Costa, P., Jr., Terracciano, A., & McCrae, R. R. (2001). Gender differences in personality traits across cultures: Robust and surprising findings. *Journal of Personality and Social Psychology*, *81*(2), 322-331. <http://dx.doi.org/10.1037/0022-3514.81.2.322>
- Craig, J. R., Redfield, D. L., & Galluzzo, G. R. (1986). *Evaluating effective teaching in colleges and universities: How far have we come?* Paper presented at the annual meeting of the American Evaluation Association, Kansas City, MO. Retrieved January 23, 2015 <http://eric.ed.gov/?id=ED282888>
- Downs, V. C., & Downs, T. M. (1993). *An exploratory and descriptive study identifying communicative behaviors associated with effective college teaching*. Paper presented at the annual meeting of the International Communication Association. May 1993, Washington, D.C.
- Dukes, R. L., & Victoria, G. (1989). The effects of gender, status, and effective teaching on the evaluation of college instruction. *Teaching Sociology*, *17*(4), 447-457. <http://dx.doi.org/10.2307/1318422>
- Fandt, P. M., & Stevens, G. E. (1991). Evaluation bias in the business classroom: Evidence relating to the effects of previous experiences. *Journal of Psychology: Interdisciplinary & Applied*, *125*(4), 469-478. <http://dx.doi.org/10.1080/00223980.1991.10543309>
- Feldman, K. A. (1992). College students' views of male and female college teachers: Part 1—evidence from the social laboratory and experiments. *Research in Higher Education*, *33*(3), 317-375. <http://dx.doi.org/10.1007/BF00992265>
- Feldman, K. A. (1993). College students' views of male and female college teachers: Part II--Evidence from students' evaluations of their classroom teachers. *Research in Higher Education*, *34*(2), 151-211. <http://dx.doi.org/10.1007/BF00992161>
- Fernandez, J., & Mateo, M. A. (1997). Student and faculty gender in ratings of university teaching quality. *Sex Roles*, *37*(11-12), 997-1003. <http://dx.doi.org/10.1007/BF02936351>
- Franklin, J., & Theall, M. (1992). *Student ratings of instructions and gender differences revisited*. Paper presented at the 75th Annual Meeting of the American Educational Research Association, New Orleans, LA.
- Goldberg, G., & Callahan, J. (1991). Objectivity of student evaluations of instructors. *Journal of Education for Business*, *66*(6), 377-78. <http://dx.doi.org/10.1080/08832323.1991.10117505>
- Hameed, F., Ali, A., Hameed, A., Saleem, Z., & Javed, Y. (2015). Teacher evaluation: the role of gender. *Quality & Quantity*, *49*(5), 1779-1789. <http://dx.doi.org/10.1007/s11135-014-0054-3>
- Hamermesh, D. S. & Parker, A. M. (2005). Beauty in the classroom: Instructors' pulchritude and putative pedagogical productivity. *Economics of Education Review*, *6*(2), 377-78. <http://dx.doi.org/10.1016/j.econedurev.2004.07.013>
- Huston, T. (2005). Research report: Race and gender bias in student evaluations of teaching. Retrieved January 23, 2015 from <http://www.seattleu.edu/cetl/RaceGenderEvalTeaching.doc>
- IRPHE. (2010). Iranian Research and Planning in Higher Education: *the statistics of higher education*. Retrieved January 23, 2015 http://www.irphe.ir/FA/Statistics/Stat_index.aspx
- Kierstead, D., D'Agostino, P., & Dill, H. (1988). Sex role stereotyping of college professors: Bias in students' rating of instructors. *Journal of Educational Psychology*, *80*(3), 342-344. <http://dx.doi.org/10.1037/0022-0663.80.3.342>
- Zare-ee, A., Mohd Don, Z. & Tohidian, I. (2016). Gender differences in students' ratings of university teachers in the Iranian education system. *Learning and Teaching in Higher Education: Gulf Perspectives* *13*(1). <http://dx.doi.org/10.18538/lthe.v13.n1.232>

Konrad, A. G. (1986). *Student evaluation of teaching: Research and recommendations*. Retrieved January 23, 2015 from <http://eric.ed.gov/?id=ED279658>

Langbein, L. (2007). Management by results: Student evaluation of faculty teaching and the mis-measurement of performance. *Economics of Education Review*, 27(4), 417-428.
<http://dx.doi.org/10.1016/j.econedurev.2006.12.003>

Lueck, T. L., Endres, K. L., & Caplan, R. E. (1993). The interaction effects of gender on teaching evaluations. *Journalism & Mass Communication Educator*, 48(3), 46-54.
<http://dx.doi.org/10.1177/107769589304800307>

Marsh, H. W. (1984). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases, and utility. *Journal of Educational Psychology*, 76(5), 707-754.
<http://dx.doi.org/10.1037/0022-0663.76.5.707>

Miller, J. & Chamberlin, M. (2000). Women are teachers, men are professors: A study of student perceptions. *Teaching Sociology*, 28(4), 283-298. <http://dx.doi.org/10.2307/1318580>

Morgan, J., & Davies, T. (2006). Analysis of bias in student evaluations of faculty at an all female Arab university in the Middle East. *Learning and Teaching in Higher Education: Gulf Perspectives*, 3(2), 1-20.

Peterson, K. D. (2000). Teacher evaluation. In *A comprehensive guide to new directions and practices* (2nd ed.). Thousand Oaks, CA: Corwin Press, Inc. Retrieved January 28, 2015 from http://www.seattleu.edu/cetl/cetl_006.htm

Radmacher, S. A., & Martin, D. J. (2001). Identifying significant predictors of student evaluations of faculty through hierarchical regression analysis. *The Journal of Psychology*, 135(3), 259-268.
<http://dx.doi.org/10.1080/00223980109603696>

Rubin, R. B. (1981). Ideal traits and terms of address for male and female college professors. *Journal of Personality and Social Psychology*, 41(5), 966-974. <http://dx.doi.org/10.1037/0022-3514.41.5.966>

Santhanam, E. & Hicks, O. (2002). Disciplinary, gender, and course year influences on student perceptions of teaching: Explorations and implications. *Teaching in Higher Education*, 7(1), 17-31.
<http://dx.doi.org/10.1080/13562510120100364>

Schwier, R. A. (1982). Design and use of student evaluation instruments in instructional development. *Journal of Instructional Development*, 5(4), 28-34. <http://dx.doi.org/10.1007/BF02905229>

Seldin, P. (1989). How colleges evaluate professors. 1988 vs. 1983. *AAHE Bulletin*, 41(7), 3-7. Retrieved January 23, 2015 from <http://eric.ed.gov/?id=ED305872>

Seldin, P. (1993). The use and abuse of student ratings of professors. *The Chronicle of Higher Education*, 39(46): A40.

Siskind T. G., & Kearns, S. P. (1997). Gender bias in the evaluation of female faculty at the Citadel: A qualitative analysis. *Sex Roles*, 37(7-8), 495-525.

Tieman, C. R., & Rankin-Ullock, B. (1985). Student evaluations of teachers. *Teaching Psychology*, 12(2), 177-191.

Zare-ee, A., Mohd Don, Z. & Tohidian, I. (2016). Gender differences in students' ratings of university teachers in the Iranian education system. *Learning and Teaching in Higher Education: Gulf Perspectives* 13(1).
<http://dx.doi.org/10.18538/lthe.v13.n1.232>

Appendix A

Questionnaire Items given to students in Likert format.

Dear Student: The following questions have been designed to reflect your evaluation of university professors' professional behavior and performance in your idea. There are 24 statements about female professors and the same 24 statements about male professors. Using the best of your knowledge of how male and female teachers at your university behave, have behaved, and/or might behave, please mark the appropriate box to indicate your level of agreement or disagreement with each statement. The items are on male and female teachers in general and you are not evaluating any specific teacher. Please do NOT write your names on the questionnaire. The information that you provide is kept completely anonymous and confidential.

Student's gender: Field of study: Year at university:

Statements/items on Female Professors' Professional Behavior

1. Female university teachers observe policies and regulations of their university.
2. Female university teachers wear clothes that match the standards of their university.
3. Female university teachers are neat and clean in their records, files, and lesson plans.
4. Female university teachers report absences of their students regularly.
5. Female university teachers use students' personal information discreetly.
6. Female university teachers appropriately report problems to the officials for consideration and attention.
7. Female university teachers regularly attend meetings and extracurricular activities.
8. Female university teachers observe the rules of using computers, projectors, copying machines, and other tools.
9. Female university teachers keep their classrooms and offices clean.
10. Female university teachers are punctual in attending classes and meetings.
11. Female university teachers are friendly with university administrative staff.
12. Female university teachers become friendly with other academic staff members and colleagues.
13. Female university teachers are careful when using other colleagues' rooms, materials, or equipment.
14. When making decisions that affect other colleagues, female university teachers inform them in advance.
15. Female university teachers respect their colleagues' personal attitudes.
16. Female university teachers do NOT gossip about other staff members.
17. Female university teachers are careful in referring students to other faculty members.
18. Female university teachers do NOT gossip about their students.
19. In dealing with students, female university teachers are fair and pleasant.
20. Female university teachers help students to realize their potentialities in learning.
21. Female university teachers consider individual differences and needs in their dealing with students.
22. Female university teachers are genuinely concerned about the progress of their students.
23. Female university teachers show interest in students' class and extracurricular activities.
24. Female university teachers have social contacts with their students in keeping with standards of the university.

Statements/items on Male Professors' Professional Behavior

1. Male university teachers observe policies and regulations of their university.
2. Male university teachers wear dresses that match the standards of their university.
3. Male university teachers are neat and clean in their records, files, and lesson plans.
4. Male university teachers report absences of their students regularly.
5. Male university teachers use students' personal information discreetly.
6. Male university teachers appropriately report problems to the officials for consideration and attention.
7. Male university teachers attend meetings and extracurricular activities.
8. Male university teachers observe the rules of using computers, projectors, copying machines, and other tools.
9. Male university teachers keep their classrooms and offices clean.
10. Male university teachers are punctual in attending classes and meetings.
11. Male university teachers are friendly with university administrative staff.
12. Male university teachers become friendly with other staff members and colleagues.
13. Male university teachers are careful when using other colleagues' rooms, materials, or equipment.
14. When making decisions that affect other colleagues, male university teachers inform them in advance.
15. Male university teachers respect their colleagues' personal attitudes.
16. Male university teachers do NOT gossip about other staff members.
17. Male university teachers are careful in referring students to other faculty members.
18. Male university teachers do NOT gossip about their students.
19. In dealing with students, male university teachers are fair and pleasant.
20. Male university teachers help students to realize their potentialities in learning.
21. Male university teachers consider individual differences and needs in their dealing with students.
22. Male university teachers are genuinely concerned about the progress of their students and show this to them.
23. Male teachers show interest in students' class and extracurricular activities.
24. Male university teachers have social contacts with their students in keeping with standards of the university.