Practical Assessment, Research & Evaluation

A peer-reviewed electronic journal.

Volume 20, Number 4, February 2015 ISSN 1531-7714

Hot or Not: The Role of Instructor Quality and Gender on the Formation of Positive Illusions Among Students using RateMyProfessors.com

Katherine C. Theyson, *Sewanee: The University of the South*

Existing literature indicates that physical attractiveness positively affects variables such as income, perceived employee quality and performance evaluations. Similarly, in the academic arena, studies indicate instructors who are better looking receive better teaching evaluations from their students. Previous analysis of the website RateMyProfessors.com confirms this, indicating that instructors who are viewed by students as "hot" receive higher "quality" ratings than those who are "not." However, psychology literature indicates that perceptions of attractiveness are influenced by positive illusions, a property whereby individuals with higher quality relationships view each other more positively than objective observers. This paper uses data from Rate My Professors to investigate the existence of positive illusions in the instructor-student relationship. It finds that positive illusions exist, suggesting that existing literature overestimates the premium associated with physical attractiveness. Furthermore, the source of these illusions varies significantly between male and female instructors with important implications for the role of gender in workplace evaluations, hiring, promotion, and tenure.

A growing economic literature has focused on the subject of perceived physical attractiveness as it relates to variables such as income, perceived employee quality and job performance (See Hamermesh and Biddle, 1994; Hamermesh, Ming & Zhang, 2002, among others). Within this area of study, one line of inquiry investigates the impact of attractiveness on perceived teacher quality. This line of inquiry is important because many institutions use student evaluations of teaching quality in their promotion and tenure decisions and any distortions of these evaluations due to attractiveness could have profound impacts on a faculty member's career trajectory. Furthermore, these results suggest perceived attractiveness may influence the success of employees in a variety of fields where evaluation processes determine raises, promotions, and continued employment. Hamermesh and Parker (2005) investigate this issue using institutional level teaching evaluations and objective physical attractiveness measurements[1] for professors at the University of Texas at Austin. They find that there is a positive relationship between attractiveness and students' perception of class quality from end of course evaluations with marginal benefits for attractiveness accruing more to men than to women. They interpret these results (with some caution) as indicating that better looking individuals are more productive, perhaps because "students simply pay more attention to good-looking instructors." Other researchers have exploited the website RateMyProfessors.com because it provides information from students on 3 aspects of instruction: Helpfulness, Clarity and Easiness. This website is particularly useful because of another question it

[1] Six students at University of Texas at Austin were asked to rate professors at University of Texas at Austin on beauty from 1 to 10 based on pictures publically available on the university's website.

describes as "just for fun" in which students can rate their professor's appearance by designating them as "hot or not." Using RateMyProfessors.com, Felton, Mitchell and Stinson (2004) find that there is a positive and significant correlation between "hotness" and professor quality. Lawson and Stephenson (2005) reconfirm this relationship using regression analysis and assert that these findings indicate that professors gain in perceived quality from hotness. Furthermore, Sen, Voia and Woolley (2010) find that for some midcareer and senior professors as well as male professors in general hotness can result in a "significant earnings premium."

The issue of causality, however, is somewhat thornier. These papers do not attempt to address causality and instead assume that hotness leads to higher teaching evaluations or that students give a "premium" to better looking professors. While it is possible that hotness induces better teaching evaluations from students, it is also possible that students are more likely to view higher quality teachers (and thus teachers they like) as hot. This notion, known as "positive illusions," asserts that individuals frequently exhibit unrealistically favorable impressions of their own personal characteristics and/or the characteristics of those they are in close relationships with. This idea has received significant study in psychology literature (e.g. Murray, Holmes & Griffin, 1996; Murray, Holmes, Dolderman & Griffin, 2000; Sangrador & Yela, 2000; Barelds & Dijkstra, 2009; etc.). Positive illusions have been shown to develop in romantic relationships, and to be positively associated with relationship quality (Barelds & Dijkstra, 2009). Furthermore, Murray *et. al.* (2000) suggest that those in "satisfying relationships" perceive more virtues in their partner than those in less satisfying relationships. These characteristics have also been demonstrated in parent-child and stepparent-stepchild relationships (Cohen & Flowers, 2004) and it seems quite probable that they could exist in the context of normal professor-student relationships.

None of the existing literature, however, investigates the existence of positive illusions and whether or not teacher quality can affect student perceptions of professors' physical attractiveness. While Hammermesh and Parker attempt to generate an "independent measure" of attractiveness, their attractiveness ratings come from students at the same university as the instructors whose attractiveness is being assessed. This may present an issue as the students generating the ratings of physical attractiveness may have knowledge about some of the instructors being rated that could affect their ratings, either because they or their friends had them in class or are otherwise familiar with them and their positive or negative reputation. Other papers simply assume that students are objectively rating whether their instructor is hot and treat this variable as exogenous. This is an important issue because if higher quality teachers are more likely to induce positive illusions amongst their students then they are more likely to be rated as attractive or hot by those same students. In this case, failure to address the existence of positive illusion may lead to an overstatement of the impact of attractiveness on instructional evaluations as well as labor market outcomes when salaries are related to performance or teaching evaluations. In short, attractiveness may not matter as much as previous research has suggested it does.

This paper studies the impact of objective hotness on teaching evaluations, explores the presence of positive illusions in the context of student-instructor relationships and investigates the sources of these positive illusions. This is done using objective hotness data, collected from students at the author's home institution, on instructors from another university located within the same geographic region, combined with data for the same instructors quality, clarity, helpfulness, easiness and hotness from the website RateMyProfessors.com. It finds that, while professors who were objectively rated as better looking receive higher teaching evaluations from their students, there is strong evidence for the existence of positive illusions among students. Specifically, students are more likely to rate their professor as hot if that professor is also rated as high quality, even when controlling for objective attractiveness, suggesting that previous research has overestimated the impact of attractiveness or hotness on teaching evaluations. Furthermore, while both male and female instructors may benefit from positive illusions, the source of these illusions differs significantly across genders with illusions about male instructors' originating from clarity while those about female instructors' stem from helpfulness.

Data

The data for this analysis come principally from two sources: RateMyProfessors.com and objective ratings of attractiveness using pictures from the website of the university from which the 476 instructors were drawn. Teaching evaluations and attractiveness ratings were drawn from RateMyProfessors.com, which allows students to rate their professor or instructor using three metrics: helpfulness, clarity and easiness, all measured on a five-point scale with 5 being the highest rating. Helpfulness and clarity are averaged by Rate My Professors to get a measure of overall quality. A fourth area in which the website collects data occurs in a question labeled "just for fun" that allows students to designate their instructor as "hot" or "not." Instructors with more than 50% of raters designating them as hot have a chili pepper appear by their profile, although none of the underlying data is observable[2]. For the purpose of this study, the average ratings on helpfulness, clarity, easiness and quality were recorded, along with the presence or absence of a chili pepper and the number of student evaluations from which the averages were drawn.

The second source of data is objective evaluations of instructor attractiveness. For this assessment, pictures of faculty members were drawn from departmental websites at the target university. These publically available pictures were downloaded, matched with the Rate My Professor evaluations and placed in a slide presentation with a black background for students to use for attractiveness ratings. Only instructors for which both a picture and a Rate My Professor rating are available are included in this study. This eliminated a large number of the instructors listed on Rate My Professor, some of whom no longer taught at the university and some of whom were likely graduate students or visiting or adjunct faculty who never were included on departmental websites. In the end, the dataset contained data from 476 instructors of which 306 were male and 170 female. Summary statistics for these instructors from RateMyProfessor.com, shown in Table 1, indicate that on average there exist no statistically significant differences between male and female professors with respect to evaluations or the probability of receiving a chili pepper.

Table 1: Summary Statistics from RateMyProfessors.com

	All		Male		Female	
	Mean (sd)	**Range**	**Mean (sd)**	**Range**	**Mean (sd)**	**Range**
Helpfulness	3.82 (0.88)	1-5	3.78 (0.92)	1-5	3.89 (0.79)	1.4-5
Clarity	3.70 (0.87)	1-5	3.66 (0.89)	1-5	3.77 (0.81)	1.2-5
Easiness	3.03 (0.77)	1-5	3.03 (0.77)	1-5	3.06 (0.77)	1.7-5
Quality	3.76 (0.83)	1-5	3.72 (0.86)	1-5	3.83 (0.77)	1.3-5
Pepper	0.20 (0.40)	0-1	0.19 (0.39)	0-1	0.22 (0.41)	0-1
Number of Reviews	13.02 (15.15)	1-171	13.27 (15.40)	1-171	12.57 (14.72)	1-114

To generate appropriate objective attractiveness data, four students were recruited to assess the instructor pictures for attractiveness, two freshmen (one male and one female) and two juniors with the same gender distribution. This mix was chosen to reasonably replicate the age and gender mix at the university from which instructors were drawn. As part of the screening process, students were screened for any contact with the university from which the sample instructors were drawn. The students selected had no contacts with the study university and thus are unlikely to have their opinions swayed by prior experiences.

Table 2: Summary Statistics, Hotness

	Freshman Female	Freshman Male	Junior Female	Junior Male	Composite Rating
Mean	1.99	3.21	1.59	2.70	2.37
(sd.)	(0.90)	(0.93)	(0.90)	(1.05)	(0.76)

Students were shown pictures of people identified to them as professors and asked to rate their hotness on a scale of 1 to 5 with 5 being the hottest without being given any additional information on the scale or the definition of hotness. This method was chosen to most closely replicate the data from Rate My Professor, which also gave no definition of what constituted "hot." In preparing the picture presentation for students to evaluate, there was some concern that the order in which the pictures were presented may impact the hotness ratings. Because of this, the order in which pictures were presented to the raters was randomized. The data collected from these four students raters is

[2] Additional higher ratings of a glowing chili pepper and exploding chili pepper are newer measures available on RateMyProfessors.com, but there is no explanation available to indicate what metric generates these ratings and very few instructors receive these ratings, thus they are not considered here.

summarized in Table 2. The students clearly had some differences in their perceptions of hotness, with both female students giving average ratings below the center of the range and both male students rating, on average, above this center. Despite this, the four students showed strong correlation between their ratings with pairwise correlation coefficients ranging between 0.46 and 0.61. Ratings from the four students were averaged to generate the composite rating of hotness, also shown in Table 2, to reduce measurement error[3].

Methodology

The research herein explores two basic models. The first follows closely from Lawson and Stephenson (2005) and suggests that the "quality" of an instructor depends on that instructor's characteristics including his/her perceived "easiness," his/her gender and whether or not students rated them as "hot." Specifically:

$$Q = \alpha + \beta_1 H + \beta_2 F + \beta_3 (H \times F) + \beta_4 E + \varepsilon \qquad (1)$$

where Q represents instructor quality, E signifies easiness, F is a dummy variable that takes the value of 1 if the instructor is female and 0 otherwise, and H designates instructors that received the chili pepper icon.

Equation 1 is estimated via Tobit because the data on professor quality is bounded within the range 1 to 5. Easiness and quality perceptions may have a complicated relationship because instructors at both extremes of the easiness spectrum may be viewed by their students as lower quality. With no expectation that the effect of easiness would be linear, the model was assessed using several different measures of easiness, including a set of six dummy variables, a spline with knots at 2, 3, & 4, linear and quadratic models.

The second model analyzed is designed to determine the impact of quality teaching on perceptions of attractiveness or hotness in a positive illusions framework. It enables a test of whether the direction of causality between hotness and evaluations of quality may be opposite that suggested in previous literature. Thus, in this model, hotness depends on the characteristics of the instructor including their gender, their objective composite rating of hotness, their perceived easiness, and the quality of their instruction.

$$H = \alpha + \beta_1 Q + \beta_2 B + \beta_3 F + \beta_4 (B \times F) + \beta_5 E + \varepsilon \qquad (2)$$

where H, Q, E and F are as defined in Equation 1 and B represents the objective composite measure of attractiveness or "hotness." Equation 2 is estimated with a Probit model because hotness (H) is a dummy variable designating the overall assessment by RateMyProfessors.com reviewers of whether or not the instructor is hot. All analyses were performed with standard errors clustered at the department level.

Results

Analysis of Equation 1 confirms the findings of Lawson and Stephenson that instructors who are considered hot and easy by their students receive higher overall quality evaluations. Column 1 in Table 3 shows the results of the replication of Lawson and Stephenson. These findings indicate that both male and female instructors benefit from "hotness" and that the impacts on the two groups are not significantly different. Column 2 is restricted to instructors with at least two student evaluations, as ratings based on a single evaluation are the most likely to be biased by student selection. This reduced the number of instructors in the analysis from 476 to 441; however, the results are largely the same, with a slightly higher effect of hotness for both males and females and slightly lower coefficients on all levels of easiness.

Column 3 in Table 3 performs this analysis using a spline with knots at 2, 3, and 4 in place of the dummy variables used previously. This method captures the non-linearity that might exist with respect to the impact of easiness on perceived teacher quality while allowing the analysis of marginal changes within each range. The results still indicate a positive impact of easiness on overall quality, however they show that there is a tendency for the marginal impact of additional easiness to decrease as an instructor becomes easier. Quadratic models of easiness proved insignificant and are not presented here[4].

[3] For further information on the distribution of this data and the standardized ratings calculated from it, please see Appendix 1.

[4] See Appendix 2 for full analysis.

Table 3: Impact of "hotness" on Instructor Quality, Tobit

	(1)	(2)	(3)
Pepper	0.604**	0.647**	0.628**
	(0.110)	(0.094)	(0.099)
Female	0.102	0.040	0.022
	(0.114)	(0.105)	(0.107)
Female*Pepper	-0.038	-0.041	-0.015
	(0.179)	(0.167)	(0.172)
Easiness = 2.0-2.49	0.475**	0.275	–
	(0.169)	(0.172)	
Easiness = 2.5-2.99	0.726**	0.544**	–
	(0.204)	(0.201)	
Easiness = 3.0-3.49	0.989**	0.837**	–
	(0.195)	(0.194)	
Easiness = 3.5-3.99	1.092**	0.903**	–
	(0.217)	(0.224)	
Easiness = 4.0-4.49	1.345**	1.182**	–
	(0.192)	(0.187)	
Easiness = 4.5-5	1.510**	1.310**	–
	(0.197)	(0.229)	
Easiness Spline = 1-2	–	–	0.962**
			(0.255)
Easiness Spline = 2-3	–	–	0.532**
			(0.145)
Easiness Spline = 3-4	–	–	0.317**
			(0.118)
Easiness Spline = 4-5	–	–	0.495**
			(0.215)
N	476	441	441
Pseudo R^2	0.131	0.146	0.149

** Significant at the .05 level.

Continuing with the analysis of Equation 1, it is useful to substitute the objective rating of hotness collected from students who had no contact with the professors for their own students' assessment from RateMyProfessors.com[5]. As shown in Table 4, this analysis reveals a similar pattern with both genders benefiting equally from "hotness," and with easiness being associated with higher quality. Despite the similarity, the difference in the measure of hotness here (on a scale of 1 to 5 rather than a dummy variable) implies a slightly different interpretation. In this case, a movement of 1 point on the 5 point scale generates an increase in quality of 0.155 points (also on a 5 point scale). Thus the hottest instructors would gain an advantage of 0.62 over the least hot. This is comparable to the impact of having a chili pepper in Table 3, however, it is unlikely that all instructors with a chili pepper would receive a 5 on the objective hotness scale and all instructors without a chili pepper would receive a 1. In fact, instructors who received a chili pepper on RateMyProfessor.com averaged 3.02 on the objective hotness scale while instructors without a pepper averaged 2.21, a difference of less than one point on a five-point scale. Thus the objective measure of hotness indicates a smaller premium on hotness than indicated using the initial analysis utilizing the chili pepper. This begs the question of why a difference might exist between these two measures of hotness and leads us to the possible presence of positive illusions among students.

Table 4: Impact of "Objective Hotness" Rating on Instructor Quality, Tobit

	(1)	(2)	(3)
Objective Hotness Rating	0.172**	0.154**	0.155**
	(0.062)	(0.057)	(0.056)
Female	0.113	0.078	0.054
	(0.274)	(0.243)	(0.252)
Obj. Hotness*Female	-0.035	-0.030	-0.027
	(0.098)	(0.088)	(0.087)
Easiness = 2.0-2.49	0.345**	0.307	–
	(0.172)	(0.172)	
Easiness = 2.5-2.99	0.639**	0.605**	–
	(0.200)	(0.200)	
Easiness = 3.0-3.49	0.950**	0.878**	–
	(0.184)	(0.180)	
Easiness = 3.5-3.99	1.041**	0.980**	–
	(0.245)	(0.241)	
Easiness = 4.0-4.49	1.029**	1.252**	–
	(0.195)	(0.194)	
Easiness = 4.5-5	1.386**	1.370**	–
	(0.234)	(0.219)	
Ease 1-2	–	–	1.096**
			(0.327)
Ease 2-3	–	–	0.533**
			(0.142)
Ease 3-4	–	–	0.373**
			(0.129)
Ease 4-5	–	–	0.390
			(0.222)
Picture Quality Controls	No	Yes	Yes
N	441	441	441
Pseudo R^2	0.105	0.119	0.125

** Significant at the .05 level.

The existence of positive illusions among students would suggest that students would view instructors with whom they have a good relationship as "hotter" than an objective viewer and would possibly view

[5] The results shown here use the raw objective hotness data. Standardizing the data does not result in a significant change in the impact of objective hotness. See Appendix 2 for the full analysis.

instructors with whom they had a poor relationship as less hot than an objective individual. Thus, to determine whether or not there may exist positive illusions in a student/instructor relationship, a probit analysis of Equation 2 is used to assess whether or not the qualities of an instructor influence the probability that they receive a pepper, controlling for their objective "hotness." The results of this analysis are shown in Table 5.

Table 5: Impact of Quality on Probability of Receiving a Chili Pepper, Probit

	(1)	(2)
Quality	0.990** (0.137)	–
Helpfulness	–	0.554** (0.190)
Clarity	–	0.437** (0.158)
Objective Hotness Rating	1.026** (0.165)	1.023** (0.165)
Female	0.521 (0.413)	0.502 (0.407)
Female*Objective Hotness	-0.200 (0.143)	-0.194 (0.140)
Ease 1-2	1.576 (1.888)	1.598 (1.882)
Ease 2-3	-0.092 (0.425)	-0.087 (0.428)
Ease 3-4	-0.166 (0.216)	-0.177 (0.222)
Ease 4-5	-1.181 (0.992)	-1.182 (0.991)
Picture Quality Controls	Yes	Yes
N	441	441
Pseudo R^2	0.354	0.354

** Significant at the .05 level.

The significance of instructor quality as a predictor of the probability of obtaining a chili pepper, even when controlling for objective hotness, supports the presence of positive illusions among students. In fact, quality is nearly as strong a predictor of the chili pepper as objective hotness, with a one point increase in quality yielding a 17.7% average marginal increase in the probability of being designated by students as "hot" compared to an 18.4% increase in the same probability from a one point increase in objective hotness. Gender does not significantly affect the probability of having a chili pepper, nor does easiness, indicating that students can have quality relationships with professors of either gender and that being easy is not a contributing factor to developing such a relationship[6].

In considering the factors that may affect relationship development between students and instructors, recall that the measure of quality on RateMyProfessors.com is a composite of two other measures: helpfulness and clarity. Column 2 in Table 5 replaces the measure of overall quality with these two sub-components and indicates that while both helpfulness and clarity are contributing factors to positive illusions, helpfulness is the more important factor, leading to a 10% increase, on average, in the probability of receiving a chili pepper for each additional point, while clarity increases the probability by only 7.9%.

Table 6: Impact of Quality on Probability of Receiving a Chili Pepper by Gender, Probit

	Male		Female	
Helpfulness	0.373	(0.277)	1.033**	(0.268)
Clarity	0.590**	(0.175)	0.029	(0.324)
Objective Hotness Rating	1.071**	(0.154)	0.923**	(0.143)
Ease 1-2	0.544	(1.301)	–†	
Ease 2-3	0.201	(0.306)	-0.574	(0.837)
Ease 3-4	-0.256	(0.293)	-0.213	(0.497)
Ease 4-5	-1.909	(1.412)	-0.423	(1.000)
Picture Quality Controls	Yes		Yes	
N	283		150	
Pseudo R^2	0.376		0.356	

** Significant at the .05 level

† For female instructors, an easiness rating between 1 and 2 perfectly predicted the absence of the chili pepper. For this reason, the category had to be eliminated from the analysis along with eight instructors.

Analyzing the factors that contribute to the chili pepper designation for men and women separately, Table 6 indicates that this difference between the marginal effects of helpfulness and clarity is generated largely by differences between the sexes. For male instructors, positive illusions are generated through the instructor's clarity with each additional point of clarity increasing the probability of receiving the chili pepper by 9.7%. For these instructors, neither helpfulness nor easiness is a significant predictor of chili pepper status. For female instructors, however, the situation is

[6] Because easiness factors significantly into quality ratings (see Tables 3 and 4), it may be possible that easiness' impact on perceived hotness is embedded in the significance of quality. To test this, the same analysis was performed removing quality from the analysis. While this resulted in marginal changes in some coefficients, it did not in any way affect the significance of the covariates confirming that easiness does not contribute to this relationship.

significantly different with helpfulness being significant but clarity insignificant. For female instructors an additional point in helpfulness generates, on average, a 20% increase in the probability of being awarded a chili pepper. While helpfulness and clarity are highly correlated within the dataset as a whole, this result may be a manifestation of differences in gender norms. Research indicates that while men need to exhibit strength to be viewed as an effective leader, women need to also exhibit sensitivity (Johnson, Murphy, Zwedie & Reichard, 2006). In a classroom setting, sensitivity may reasonably be interpreted as helpfulness, thus contributing to quality relationships between female instructors and those they lead (their students), while male instructors may not need to be as helpful to generate similar relationships.

As with any research of this nature a few caveats clearly apply. Firstly, RateMyProfessors.com makes no attempt to collect the views of representative students. Thus, the students who have rated their instructors using this service may represent a non-random sample of students and their opinions. Second, the website offers participants no reference points within the scale for each characteristic rated other than the endpoints of 1 and 5, making ratings completely subjective. Thirdly, while instructors with only one evaluation were excluded from the majority of this analysis, 50% of all included instructors had fewer than 10 evaluations, potentially introducing bias if, as previously noted, these evaluators are not randomly selected. Fourth, while the analyses herein clustered standard errors at the department level, course information from RateMyProfessors.com was not used due to reliability issues. This may bias the results if, as seen in Hamermesh and Parker (2004), students in lower level classes put more emphasis on attractiveness than those in upper level classes and responses from students on RateMyProfessors.com are skewed towards one of these levels or if young instructors are more likely to be teaching lower level classes than their senior peers. Despite these four concerns, the literature shows high levels of correlation between RateMyProfessors.com reviews and student assessments of teaching done within the traditional university setting (Coladarci & Kornfield, 2007) suggesting that these issues may not be as important as in other arenas.

Conclusions

Student evaluations of instructor quality are widely used in universities throughout the United States as part of the promotion and tenure process. Thus, these evaluations impact the career trajectory of faculty members. The results herein reconfirm that evaluations exhibit a positive premium on attractiveness or hotness for both male and female professors. This suggests that better looking instructors receive higher teaching evaluations than their less attractive peers and thus gain advantages in the promotion and tenure process. However, the influence of attractiveness on quality evaluations is not as large as previous research has suggested. This paper's results also indicate that positive illusions exist among students and are associated with instructor quality. High quality instructors are likely to establish a rapport with their students that positively influences those students' perceptions of the instructor's attractiveness. Thus perceptions of physical attractiveness from an instructor's students are likely to be skewed by the instructor-student relationship leading to an overstatement of the attractiveness premium. Accounting for positive illusions among students lessens, but does not eliminate the premium on attractiveness, suggesting that this is still a significant factor in student evaluations of instructors.

In contrast to the previous findings of Hamermesh & Parker (2005) the magnitude of the impact of attractiveness in this study does not differ by gender. However, gender differences do exist with respect to the source of positive illusions that may have implications for the promotion and tenure decision. Specifically, the results indicate that origin of rapport that leads to positive illusions operates through clarity for men and helpfulness for women. While this finding coincides with existing literature on gender stereotypes and perceptions of leadership quality, it is not yet clear what personality traits or time commitments are necessary to be "clear" or "helpful". For example, if helpfulness stems from an instructor being available in his/her office, then female instructors may need to commit more time to office hours than male instructors, which might negatively impact their research productivity. Similarly, if clarity is generated by time spent in class preparation, male instructors may have this same disadvantage. In addition, women may be more likely to be penalized by their students for having a brusque or businesslike

persona, with resulting negative evaluations being disadvantageous to securing promotions. Further research utilizing more detailed institutional level evaluations of instructors by students may allow some of these questions to be answered. It would also be beneficial for additional research to examine the extent to which positive illusions operate in other workplace environments. If attractiveness is similarly affected by employee quality elsewhere in the labor market, then the magnitude of the impact of attractiveness in previous studies may need downward revision[7].

In conclusion, while these results indicate, yet again, that perceptions of performance are enhanced by attractiveness, they offer a ray of light to the less beautiful among us: that quality of instruction is a significant predictor of perceived attractiveness and that it is not different, in the magnitude of its effect, from objective attractiveness or hotness. Thus, if we work hard to build relationships with our students, we will not only be higher quality instructors, we become more attractive as well (at least in their eyes).

References

Barelds, D., & Dijkstra, P. (2009). Positive illusions about a partner's physical attractiveness and relationship quality. *Personal Relationships*, *16*, 263-283.

Bokek-Cohen, Y., & Davidowitz, N. (2008) Beauty in the Classroom: Are Female Students Influenced by the Physical Appearance of Their Male Professors? *Journal of Education and Human Development*, *2*(1).

Cipriani, G., & Zago, A. (2011). Productivity or Discrimination? Beauty and the Exams. *Oxford Bulletin of Economics and Statistics*, *73*, 428-447.

Cohen, J., & Fowers, B. (2004). Blood, Sweat and Tears: Biological Ties and Self-Investment as Sources of Positive Illusions About Children and Stepchildren. *Journal of Divorce & Remarriage*, *42*(1/2), 39-59.

[7] For example, Hamermesh and Biddle (1994) and Hamermesh *et. al.* (2002) use data from surveys in which the survey administrator was asked to rate the beauty or looks of the survey respondent. It is possible that the administrator may have been influenced in their rating by the personality of the respondent or by the nature of some of their responses, which could lead to an overstatement of the beauty premium. This may especially be true of Hamermesh *et. al.* (2002) as in that survey the rating of the respondent's looks took place at the end of the survey. (It is not clear from Hamermesh and Biddle (1994) when in the survey process the beauty rating took place.)

Coladarci, T., & Kornfield, I. (2007). RateMyProfessors.com versus formal in-class student evaluations of teaching. *Practical Assessment, Research and Evaluation*, *12*(6). Available online: http://pareonline.net/getvn.asp?v=12&n=6

Felton, J., Koper, P., Mitchell, J., & Stinson, M. (2006). Attractiveness, Easiness, and Other Issues: Student Evaluations of Professors on RateMyProfessors.com. Central Michigan University.

Felton, J., Mitchell, J., & Stinson, M. (2004). Web-based student evaluation of professors: the relations between perceived quality, easiness and sexiness. *Assessment and Evaluation in Higher Education*, *29*(1), 91-108.

Hamermesh, D., & Biddle, J. (1994). Beauty and the Labor Market. *American Economic Review*, *84*, 1174-1194.

Hamermesh, D., Meng, X., & Zhang, J. (2002). Dress for Success: Does Primping Pay? *Labour Economics*, *9*, 361-373.

Hamermesh, D., & Parker, A. (2005). Beauty in the Classroom: Instructors' Pulchritude and Putative Pedagogical Productivity. *Economics of Education Review*, *24*, 369-376.

Johnson, S., Murphy, S. E., Zewdie, S., & Reichard, R. (2006). The strong sensitive type: Effects of gender stereotypes and leadership prototypes on the evaluation of male and female leaders. *Organizational Behavior and Human Decision Processes*, *106*, 29-60.

Lawson, R., & Stephenson, F. (2005). Easiness, Attractiveness, and Faculty Evaluations: Evidence from RateMyProfessors.com. *Atlantic Economic Journal*, *33*, 485-486.

Murray, S., Holmes, J., Dolderman, D., & Griffin, D. (2000). What the Motivated Mind Sees: Comparing Friends' Perspectives to Married Partners' Views of Each Other. *Journal of Experimental Social Psychology*, *36*, 600-620.

Murray, S., Holmes, J., & Griffin, D. (1996). The Benefits of Positive Illusions: Idealization and the Construction of Satisfaction in Close Relationships. *Journal of Personality and Social Psychology*, *70*, 79-98.

Ponzo, M., & Scoppa, V. (2012) The Good, The Bad, and The Ugly: Teaching Evaluations, Beauty and Abilities. Università Della Calabria, Working Paper 04-2012.

Sen, A., Voia, M., & Wooley, F. (2010). Hot or Not: How appearance affects earnings and productivity in academia. Carleton University, Carleton Economic Papers, CEP 10-07.

Wegner, A., & Fowers, B. (2008). Positive Illusions in Parenting: Every Child is Above Average. *Journal of Applied Social Psychology*, *38*, 611-634.

Appendix 1: Objective Hotness Data Standardization

Objective data on instructor hotness was collected from four students, two male and two female with one of each from the freshman class and one of each from the junior class. Students rated instructors on a scale of 1 to 5 with 1 being the least hot and 5 the most hot. Information on this raw data is available in Table 2.

One issue confronted revolved around whether or not the data collected from these students should be standardized. While the results presented in the main body of the paper use the raw data, the analyses were also performed with data standardized at the evaluator level to a distribution with a mean of zero and a standard deviation of 1. The ranges for this data are shown in Table A1-1 that indicate that three of four students had ratings that skewed to the right. The four standardized student ratings were then averaged into a composite standardized rating with a mean of 0 and a standard deviation of 0.80. This composite rating still exhibits a rightward skew indicating that, perhaps, college instructors are not normally distributed with respect to hotness.

Table A1-1: Standardized Summary Statistics, Hotness

	Freshman Female	Freshman Male	Junior Female	Junior Male	Composite Rating
Maximum	-1.093	-2.388	-0.657	-1.618	-1.439
Minimum	3.333	1.930	3.808	2.185	2.577

Despite this issue, Hamermesh and Parker (2005) note that principle concern with this type of data is whether or not the assessments of hotness were consistent across evaluators. Consistency was analyzed using pair-wise correlation coefficients for the evaluators, which range from 0.46 to 0.61 with an average of 0.53, indicating substantial correlation. Furthermore, the consistency of evaluations across students was evaluated using Cronbach's alpha, a coefficient of internal consistency used to assess the reliability of psychometric data. This value can vary between 0 and 1, with higher values representing greater reliability of the measurement. The value of Cronbach's alpha for the four students evaluators was 0.82, representing considerable agreement between individual raters.

Appendix 2: Evaluation of the Data and Robustness Checks

There were a number of potential robustness checks that were performed to verify the results presented in the main body of the paper. This appendix discusses the additional analyses and robustness checks performed and the results they provide. These analyses included use of standardized objective hotness data in place of raw objective hotness, analyses separated by the gender of the instructor and analyses using linear and quadratic specifications with respect to easiness.

Evaluation of Standardized Objective Hotness Data

Use of the standardized objective hotness rating in place of the raw hotness rating in the analysis of the impact of objective hotness on teaching quality ratings, Table A2-1, has only a small impact on the magnitude of the coefficients and no impact on their signs or the significance of the covariates with the exception of the dummy variable for female in regression 3 which is now negative, but still highly insignificant. Results indicate that an instructor with hotness one standard deviation above the mean receives an increase in teaching quality of approximately 0.15 points on a 5 point scale, a result similar to that seen in Hamermesh and Parker (2005) but smaller in both absolute and relative magnitude.

Table A2-1: Impact of Standardized "Objective Hotness" Rating on Instructor Quality, Tobit

	(1)	(2)	(3)
Standardized Objective Hotness Rating	0.162*** (0.059)	0.145*** (0.054)	0.146*** (0.053)
Female	0.029 (0.102)	0.007 (0.094)	-0.011 (0.093)
Obj. Hotness*Female	-0.032 (0.093)	-0.026 (0.084)	-0.024 (0.082)
Easiness = 2.0-2.49	0.344** (0.172)	0.307* (0.172)	–
Easiness = 2.5-2.99	0.639*** (0.200)	0.605*** (0.201)	–
Easiness = 3.0-3.49	0.949*** (0.184)	0.878*** (0.179)	–
Easiness = 3.5-3.99	1.041*** (0.246)	0.980*** (0.242)	–
Easiness = 4.0-4.49	1.287*** (0.234)	1.251*** (0.194)	–
Easiness = 4.5-5	1.384*** (0.181)	1.368*** (0.220)	–
Ease 1-2	–	–	1.096*** (0.327)
Ease 2-3	–	–	0.533*** (0.142)
Ease 3-4	–	–	0.373*** (0.129)
Ease 4-5	–	–	0.388* (0.223)
Picture Quality Controls	No	Yes	Yes
N	441	441	441
Pseudo R^2	0.105	0.119	0.125

Gender-Specific Analyses

The analysis in the main body of the paper also pooled male and female instructors, accounting for the difference through the use of a dummy variable for female instructors and an interaction between the female dummy and hotness. This method assumes that female and male instructors gain similarly from the other covariates, namely easiness. Since this is not a given, Tables A2-2 and A2-3 display the analysis separated by gender.

Gender-specific analysis reveals some interesting differences in the impact of easiness on perceived quality of an instructor. Table A2-2 displays the results of this analysis using a dummy variable approach where the excluded group is the instructors considered by their students to be the most difficult, with easiness ratings between 1 and 1.99. Male instructors were rewarded with higher quality ratings by their students for being easy with each category of easiness significantly different from the excluded group and increasing quality associated with increasing easiness. Female instructors, however, do not receive the same quality bonuses for their easiness. In fact, female instructors with easiness ratings between 2 and 2.49 saw no increase in their quality evaluations as compared to their most difficult peers. Furthermore, at all levels of easiness, the quality rewards for female professors were smaller than for their male peers and when using the chili pepper to measure hotness, did not gain statistical significance until the female instructor's easiness rating exceeded 3 on a 5 point scale.

Table A2-2: Impact of Hotness on Quality by Gender

	Male		Female	
	(1)	(2)	(3)	(4)
Pepper	0.646*** (0.088)	–	0.606*** (0.125)	–
Objective Hotness Rating	–	0.151*** (0.054)	–	0.108* (0.060)
Easiness = 2.0-2.49	0.421** (0.164)	0.377** (0.190)	0.028 (0.320)	0.112 (0.288)
Easiness = 2.5-2.99	0.572*** (0.184)	0.587*** (0.189)	0.474 (0.332)	0.546* (0.311)
Easiness = 3.0-3.49	0.906*** (0.169)	0.896*** (0.173)	0.678* (0.352)	0.759** (0.302)
Easiness = 3.5-3.99	0.958*** (0.189)	0.960*** (0.220)	0.776** (0.379)	0.906** (0.372)
Easiness = 4.0-4.49	1.235*** (0.152)	1.219*** (0.177)	1.053*** (0.355)	1.179*** (0.367)
Easiness = 4.5-5	1.417*** (0.180)	1.396*** (0.221)	1.097** (0.456)	1.160*** (0.415)
Picture Quality Controls	No	Yes	No	Yes
N	283	283	158	158
Pseudo R^2	0.133	0.117	0.179	0.147

When using a spline with knots a 2, 3 and 4, shown in Table A2-3, the results remain quite similar, with men seeing significant effects of easiness at all levels and women only seeing gains in quality from additional easiness in the middle of the easiness spectrum. These results indicate that students are less likely to confuse easiness with quality for female instructors than for male ones, which may be the results of differences in gender norms. Despite the fact that this separation indicates a substantial gender based difference in the role of easiness on the students' perceptions of instructor quality, it does not yield statistically different estimates for the impact of the chili pepper.

Table A2-3: Impact of Hotness on Quality by Gender

	Male		Female	
	(1)	(2)	(3)	(4)
Pepper	0.630*** (0.098)	–	0.617*** (0.126)	–
Objective Hotness Rating	–	0.155*** (0.055)	–	0.116** (0.057)
Ease 1-2	1.149*** (0.271)	1.180*** (0.379)	0.010 (1.167)	0.442 (1.171)
Ease 2-3	0.449** (0.182)	0.471** (0.184)	0.683*** (0.165)	0.629*** (0.181)
Ease 3-4	0.349** (0.146)	0.359** (0.163)	0.260* (0.155)	0.393** (0.187)
Ease 4-5	0.433* (0.239)	0.359 (0.240)	0.572 (0.374)	0.332 (0.387)
Picture Quality Controls	No	Yes	No	Yes
N	283	283	158	158
Pseudo R^2	0.138	0.126	0.176	0.143

Alternative Functional Forms: Easiness

As noted in the main body of this paper, it was not assumed that there exists a linear relationship between easiness and quality of the presence of a chili pepper. Non-linear specifications of easiness using a series of dummy variables and a spline with knots at 2, 3 and 4 were presented. This section of the appendix presents the results from analyses using linear, quadratic and cubic specifications of easiness with Table A2-4 verifying the robustness of the results presented in Table 3, Table A2-5 corresponding to Table 4, and Table A2-6 doing likewise for Table 5. In each of these tables, regression 1 shows the regression using a linear form of easiness, while regressions 2 and 3 show the results of quadratic and cubic forms respectively. The results indicate that easiness does not follow either a quadratic or cubic form with respect to either quality or the probability of receiving a chili pepper and that changing the functional form of easiness has no significant impact on the value or the significance of the coefficients on other regressors.

Table A2-4: Impact of the Chili Pepper on Quality, Alternative Easiness Form, Tobit

	(1)	(2)	(3)
Pepper	0.634*** (0.098)	0.622*** (0.100)	0.626*** (0.099)
Female	0.030 (0.108)	0.029 (0.107)	0.024 (0.109)
Pepper*Female	-0.021 (0.169)	-0.014 (0.174)	-0.014 (0.173)
Easiness	0.459*** (0.053)	0.952*** (0.354)	2.233*** (0.866)
Easiness Squared	–	-0.080 (0.056)	-0.516* (0.293)
Easiness Cubed	–	–	0.047 (0.032)
N	441	441	441
Pseudo R^2	0.146	0.148	0.149

Table A2-5: Impact of Objective Hotness on Quality, Alternative Easiness Specification, Tobit

	(1)		(2)		(3)	
Objective Hotness Rating	0.164**	(0.065)	0.172***	(0.064)	0.176***	(0.062)
Female	0.084	(0.274)	0.108	(0.269)	0.109	(0.266)
Obj. Hotness*Female	-0.026	(0.101)	-0.036	(0.099)	-0.039	(0.097)
Easiness	0.492***	(0.056)	1.176***	(0.361)	2.389**	(0.961)
Easiness Squared	–		-0.111**	(0.056)	-0.524	(0.342)
Easiness Cubed	–		–		0.044	(0.039)
N	441		441		441	
Pseudo R^2	0.104		0.109		0.110	

Table A2-6: Impact of Quality on Probability of Receiving a Chili Pepper with Alternative Easiness Specifications, Probit

	(1)		(2)	
Quality	1.028***	(0.145)	1.016***	(0.148)
Objective Hotness Rating	1.011***	(0.173)	1.046***	(0.166)
Female	0.388	(0.436)	0.464	(0.433)
Female*Objective Hotness	-0.184	(0.149)	-0.212	(0.147)
Easiness	1.576	(1.888)	1.576	(1.888)
Easiness Squared	–		-0.184	(0.149)
Picture Quality Controls	Yes		Yes	
N	441		441	
Pseudo R^2	0.352		0.358	

Citation:

Theyson, Katherine C. (2014). Hot or Not: The Role of Instructor Quality and Gender on the Formation of Positive Illusions Among Students using RateMyProfessors.com. *Practical Assessment, Research & Evaluation*, 20(4). Available online: http://pareonline.net/getvn.asp?v=20&n=4

Author:

Katherine C. Theyson
Department of Economics
Sewanee: The University of the South
735 University Ave
Sewanee, TN, 37383.

Email:kctheyso [at] sewanee.edu