



# I checked the prof on ratemyprofessors: effect of anonymous, online student evaluations of professors on students' self-efficacy and expectations

Stefanie S. Boswell<sup>1</sup> · Sara L. Sohr-Preston<sup>2</sup>

Received: 16 May 2019 / Accepted: 15 May 2020  
© Springer Nature B.V. 2020

## Abstract

Little is known about the effects of Ratemyprofessors.com teaching evaluations relative to university-administered teaching evaluations in students. Therefore, this study investigated differences in use and consideration of these evaluation sources. It also investigated the effects of evaluation source, evaluation valence (mostly positive or mostly negative), and their interaction on students' decision making, self-efficacy for a course, and expectations about academic outcomes. Participants were 223 undergraduates who viewed either mostly positive or mostly negative teaching evaluations about a fictitious professor and were told that were from either university-administered teaching evaluations or Ratemyprofessors.com. Evaluation valence, but not source, affected participants' decision making, self-efficacy, and expectations about the professor's course. Implications for student self-efficacy and expectations are discussed. Additionally, recommendations for students who choose to use Ratemyprofessors.com to inform their course decision making are provided.

**Keywords** Ratemyprofessors.com · Self-efficacy · Student evaluations of teaching · Student attitudes

## 1 Introduction

On Ratemyprofessors.com (RMP), college students anonymously provide informal, online evaluations of professors. Additionally, students may access professor reviews to inform their course decision making (Brown and Kosovich 2015; Hartman et al. 2013). RMP is popular with students and use of the website has grown

---

✉ Stefanie S. Boswell  
ssboswel@uiwtx.edu

<sup>1</sup> Department of Psychology, University of the Incarnate Word, 4301 Broadway, CPO #9, San Antonio, TX 78232, USA

<sup>2</sup> Department of Psychology, Southeastern Louisiana University, Box 10831, Hammond, LA 70402, USA

quickly; since it began in 1999, RMP has received over 19 million evaluations of 1.7 million professors in the United States, the United Kingdom, and Canada (RMP 2020a).

The qualities for which students rate their professors on RMP have changed over time. Formerly, students rated professors on easiness, helpfulness, and clarity; the average of the clarity and helpfulness ratings created professors' overall quality scores (RMP n.d.). At the time of writing, students continue to rate professors on their easiness and now provide a singular overall quality rating using a scale ranging from 1 (*low*) to 5 (*high*). Current RMP guidelines encourage students to draw upon their beliefs about a professor's clarity and helpfulness when forming that professor's overall quality score (RMP 2020b). The overall quality score determines the type of "smiley face" displayed next to the professor's name; "good quality" professors receive a green smiling face, "average quality" professors receive a yellow neutral face, and "poor quality" professors receive a red frowning face (RMP, n.d.). In addition to numeric ratings, students may also provide open-ended, descriptive comments about their impressions of the professor and the course. Until recently, students also evaluated the "hotness" of professors; hot professors were denoted by a chili pepper icon. Although RMP did not operationally define the chili pepper as a symbol of physical attractiveness, many students and faculty alike believed this to be its meaning (Bleske-Rechek and Fritsch 2011; Landry et al. 2010; Theyson 2015). Under extensive criticism that the chili pepper degraded professors, particularly women, and was immaterial to teaching, RMP removed it from its ratings in June 2018 (Flaherty 2018; Simonton 2018).

RMP is a rather new resource, established in 1999, compared to university-administered student evaluations of teaching (UASET) conducted by universities, which can be traced back to the 1920s (Coladarci and Kornfeld 2007). Although many have questioned whether RMP evaluations function differently than UASET, there is little research comparing the effects of the two. Given that, this study investigated the effect of evaluation source (UASET versus RMP) and valence (positive versus negative) on undergraduates' decision making and self-efficacy for a course.

## 1.1 Why RMP?

Anonymous professor-rating websites are an international phenomenon; examples include MisProfesores.com Mexico, RateMyTeachers.com Australia, Canada, Ireland, New Zealand, and United Kingdom, kr.RateYourProf.com South Korea, MeinProf.de Germany, StudenterTyckerTill.se Sweden, Hodnoceniskol.cz Czech Republic, and Campus.nikki.ne.jp Japan. Although they vary in the qualities on which users rate professors (e.g., easiness, helpfulness, clarity, knowledge), the websites share features. The ratings are anonymous and open access; moreover, verification that the rater was ever the professor's student is not required (Hinz, 2011). Among the professor-rating websites, RMP touts itself as the standout due to its size; it is the largest of its type internationally (RMP 2020a). RMP has gained great popularity since its introduction; at the time of writing, approximately 4 million college students utilize the site every month to read about and

rate professors. According to RMP, the site was built to enable college students to select the “best courses and professors” using feedback from their peers (RMP 2020a, para 3). Compared to the word-of-mouth which informed course decision making prior to the inception of professor-rating websites, RMP provides a greater breadth of information from a larger sample to satisfy college students longing for more data. Thus, the evaluations available on RMP allow college students to act as informed consumers when decision making about education (Ritter 2008). The evaluations provide college students with information about the course outcomes that they most value, be it long-term knowledge gain or seeming assurance of an A course grade (Brown and Kosovich 2015; Hayles and Gessler 2005).

Not only is RMP the largest international professor-rating website in size, it is also the largest in terms of scholarly attention; its validity has been questioned by many scholars (Buck 2004; Clayson 2014; Davison and Price 2009; Murray and Zdravkovic 2016). A self-selection bias in the sample is a common concern; the overall population of professors’ students may significantly differ from those individuals who leave anonymous evaluations (e.g., Legg and Wilson 2012). For example, individuals who leave RMP evaluations may possess polarized attitudes toward the professors while those with temperate views eschew the site (Bleske-Rechek and Fritsch 2011; Legg and Wilson 2012). Moreover, unlike UASET, RMP offers no assurance that a professor’s evaluations are provided by that professor’s students (Carnevale 2006; Johnson and Crews 2013; Otto et al. 2008; Timmerman 2008). Evaluations may be provided by, for example, hostile colleagues (Caron 2006), helpful friends and family members, and the professors themselves (Jaschik 2006; Montell 2006).

In addition to apprehension about RMP’s sample, researchers have also questioned RMP’s validity due to concerns about bias in its evaluations related to professorial characteristics. For example, evidence suggests gender-bias with mixed findings in favor of men (e.g., Boehmer and Wood 2017) or women (e.g., Sohr-Preston et al. 2016). Professor age potentially creates additional bias; older professors are evaluated more harshly on RMP (Sohr-Preston et al. 2016; Stonebraker and Stone 2015). Professors’ perceived race also influences RMP ratings; Black and Asian professors are rated as poorer quality than their White counterparts (Reid 2010; Subtirelu 2015).

Findings regarding relationships between the RMP variables of easiness, clarity, quality, and the now defunct chili pepper rating suggest additional evaluation bias. For example, “hot” instructors were rated higher in helpfulness and clarity than their “not hot” counterparts (Felton et al. 2004; Freng and Webber 2009; Liu et al. 2013; Rosen 2017). Moreover, easier professors also receive higher quality ratings (Boehmer and Wood 2017; Lawson and Stephenson 2005; Rosen 2017). These sources of bias lead some scholars to argue that RMP does not accurately reflect instructors’ performance or student learning. Rather, it represents a consumer-culture mentality in higher education (e.g., Ritter 2008). Nevertheless, other researchers suggest that RMP evaluations correlate significantly with UASET (Brown et al. 2009) and are representative of student learning (Otto et al. 2008).

## 1.2 Effects of RMP on college students

While much of the literature has questioned the legitimacy and potential biases in its evaluations, RMP is extremely popular with students whether it is valid or not (Davison and Price 2009). Given this, increasing research interest has focused on the effects of RMP use on college students. For example, RMP evaluations influence students' perceptions of the professor. Lewandowski et al. (2012) found that previous exposure to RMP evaluations shaped students' evaluation of a live professor's teaching. Given that these evaluations shape students' expectations about professors, they stand to influence whether students ever enroll in the RMP-evaluated professors' courses. Early research found this to be the case; RMP-style evaluations significantly affected college students' course selection decisions when they received this information in an academic advising setting (Hayes and Prus 2014). Moreover, Brown and Kosovich (2015) found that course enrollment is correlated with RMP ratings; courses taught by professors with higher easiness and overall quality ratings filled faster than those taught by their lower-rated counterparts.

Findings about RMP and course enrollment, coupled with those questioning the validity of RMP, suggest that students are basing significant course decisions upon flawed information. However, students perceive their peers' RMP evaluations as a valid (Field et al. 2008) and trustworthy (Landry et al. 2010) resource when choosing their courses. Moreover, they perceive RMP evaluations to be as valid and beneficial of a tool to inform their decision making as official university resources such as academic advisors, other faculty members, and the university course catalog (Hayes and Prus 2014).

While still limited, the extant literature also reveals that exposure to informal online professor evaluations like those on RMP can affect students' affect about an upcoming class as well as expectations about their course grade (Edwards and Edwards 2013; Edwards et al. 2007; Edwards et al. 2009; Kowai-Bell et al. 2011; Kowai-Bell et al. 2012). The positivity or negativity (i.e., valence) of the evaluations, for example, is an important influence. Students who view positively valenced evaluations of the professor report looking forward to the professor's course (Kowai-Bell et al. 2011). Moreover, students exposed to positively valenced evaluations perceive greater control over their grade in the professor's course; this effect occurs in both undergraduate (Kowai-Bell et al. 2011) and graduate students (Kowai-Bell et al. 2012).

### 1.2.1 Self-efficacy

Given that RMP evaluations can influence students' perceived control over their grade in a professor's class, it stands to reason that these evaluations could also influence students' self-efficacy for other abilities related to academic success. Self-efficacy refers to individuals' beliefs that they can effectively perform the behaviors necessary to achieve a goal (Bandura 1977). Self-efficacy affects individuals' motivation to undertake a new goal, the amount of effort they are willing to devote toward the goal, and how long they will persist despite difficulties achieving that goal (Bandura 1982, 1989, 1997). Self-efficacy is dynamic; it

is contingent upon the goal to be completed (Bandura 1977; Pajares 1997). For example, college students possess college self-efficacy, or beliefs about how their personal efforts contribute to their success in college courses. Those high in college self-efficacy believe that they can competently problem-solve the challenges they encounter to achieve their desired college-related goals (Solberg et al. 1993; Solberg and Villarreal 1997).

Self-efficacy is affected by multiple factors; however, the influence of vicarious experiences on self-efficacy is particularly significant in the context of RMP evaluations. Through vicarious experiences, individuals develop self-efficacy for a goal when they observe others successfully achieve that goal (Bandura 1977). The effect of vicarious experience on self-efficacy is the greatest when individuals have limited information about their ability to perform in a particular setting (Pajares 1997); for example, when enrolling in a course with a novel professor. Through RMP evaluations, students may learn about peers' perceived successes with the professor, thus bolstering their own self-efficacy for the professor's course. In turn, students' bolstered self-efficacy may influence their willingness to enroll in the course and execute the behaviors necessary to achieve college success. Conversely, vicarious exposure to peers' perceived failures with the professor may lower self-efficacy and in turn, lower students' willingness to engage behaviors predictive of college success. For example, course behaviors such as class attendance (e.g., Gump 2005; Kwak et al. 2018), class note-taking (e.g., Reed et al. 2016), and study time (e.g., George et al. 2008), collectively referred to as course effort, are predictive of success (Svanum and Bigatti 2006). Moreover, interactions between students and professors are predictive of success (e.g., Tinto 1999). For example, interpersonal behaviors such as contact with professors and the quality of the student-professor relationship are strongly related with college success (Good and Adams 2008; Katreovich and Aruguete 2017;), sometimes more so than students' characteristics and behaviors such as engagement in study groups (Schudde 2019). Any influence that RMP may exert on self-efficacy for these behaviors stands to have significant impact on course performance.

## 2 The current study

This study investigated the effects of evaluation source (RMP or UASET) on student self-efficacy and expectations. Students perceive RMP to be a valid source to inform their course decision making (Brown et al. 2009; Field et al. 2008; Hayes and Prus 2014). However, some scholars argue that students turn to RMP because they have no alternative information source because results from UASET are rarely made public. Without access to UASET, students use RMP as an information tool for decision making about classes and professors (e.g., Kindred and Mohammed 2005). However, if students gained access to traditional UASET results, UASET may be preferred and students may be more influenced by these professor ratings. This informed the first research question:

## 2.1 Research question 1

Will students complete, read, and consider more seriously RMP evaluations than UASET?

Because evaluation valence affects students' expectations and intentions about a course, we also investigated if evaluation source (RMP or UASET) interacts with evaluation valence in its effects on student beliefs and behaviors. Specifically,

## 2.2 Research question 2

Will evaluation source, valence, or their interaction influence students' decision making about a course?

## 2.3 Research question 3

Will evaluation source, valence, or their interaction influence students' self-efficacy for course and interpersonal behaviors related to course success?

## 2.4 Research question 4

Will evaluation source, valence, or their interaction influence students' expected academic outcomes for the professor's course?

The study investigated questions 3 through 4 while controlling for participants' current course and interpersonal college self-efficacy.

# 3 Method

## 3.1 Study design

The study employed a fully between-groups multivariate design, crossing evaluation source (RMP or UASET) with evaluation valence (mostly positive or mostly negative) to create four conditions: RMP mostly positive evaluation, RMP mostly negative evaluations, UASET mostly positive evaluation, and UASET mostly negative evaluations. Participants were randomly assigned to view evaluations from one of these four conditions.

## 3.2 Participants

Sample size planning involved power analysis using G\*Power 3 (Erdfelder et al. 1996; Faul et al. 2007). A power of .80 and an alpha level of .05 were used to calculate the minimum number of participants needed to detect a medium effect size (Cohen 1988, 1992). The analysis indicated that data from a minimum of 190 participants would be necessary. Participants were recruited through a behavioral sciences participant pool at a large-sized, public university in the western United

States. Potential participants received a brief written description of the study; individuals interested in participation then completed a written informed consent process. Of the 236 individuals who consented to participate in the study, 13 were excluded due to incompleteness of study measures; therefore, the final sample included 223 participants. The majority of the sample identified as female ( $n = 146$ , 61.86%) and participants ranged in age from 18 to 60 years ( $M = 21.37$ ,  $SD = 5.28$ ).

Regarding race/ethnicity, 81.27% of participants ( $n = 181$ ) identified as White or Caucasian, 9.4% ( $n = 21$ ) identified as Hispanic or Latinx, 2.2% ( $n = 5$ ) identified as Asian, and 1.8% ( $n = 4$ ) identified as Native American. Three participants each identified as Black or African American (1.3%), Pacific Islander (1.3%), or other (1.3%). One participant (.4%) identified as multiethnic and one (.4%) provided no information about racial or ethnic identity. These demographics of the sample are representative of the university's overall student population. Finally, the sample was comprised of 114 freshmen/first-year students (51.1%), 58 sophomores (26.0%), 40 juniors (3.6%), and 8 seniors (3.6%). Five students did not provide college class information. University of the Incarnate Word's Human Subjects Institutional Review Board reviewed and approved the study.

### 3.3 Measures and materials

#### 3.3.1 College self-efficacy

Participants completed the College Self-Efficacy Inventory (Solberg et al. 1993). The College Self-Efficacy Inventory has three subscales: course efficacy, social efficacy, and roommate efficacy.

All roommate efficacy items and one social efficacy item about “getting a date” were removed to better reflect the experience of non-traditional and commuter students. Zajacova et al. (2005) made a similar modification in their study of self-efficacy in college students. After these modifications, the scale included seven course self-efficacy items and seven social self-efficacy items rated on a 10-point, Likert-type scale ranging from 1 (*not at all confident*) to 10 (*extremely confident*). Higher scores indicated greater college self-efficacy. The course self-efficacy subscale represents confidence for coursework and study-skills behaviors; for example, “take good class notes” (p. 88). The social self-efficacy subscale represents confidence for social behaviors; for example, “ask a professor a question” (p. 88). Solberg et al. (1993) reported the measure has strong internal reliability and good validity properties. In the present study, the course self-efficacy internal reliability coefficient was high ( $\alpha = .90$ ). The social self-efficacy internal reliability coefficient was good ( $\alpha = .86$ ).

#### 3.3.2 Evaluation stimuli

Participants received either mostly positive evaluations or mostly negative evaluations about a fictitious professor. The mostly positive evaluations contained four positive evaluations and one negative evaluation; the mostly negative evaluations

contained four negative evaluations and one positive evaluation. This is consistent with Kowai-Bell et al.'s (2011) study of RMP's effect on student affect. Participants in the RMP conditions were informed that the evaluations originated from RMP; participants in the UASET conditions were informed that the evaluations originated from UASET. Evaluations were copied from the RMP profiles of professors at the primary author's institution. RMP categorizes professors as good, average, or poor based on their overall quality numeric rating (RMP 2020b). The evaluations in the positive conditions originated from profiles with RMP's *good quality* ranking (score 3.5 to 5.0); evaluations in the negative conditions originated from profiles with RMP's *poor quality* ranking (score: 0 to 2.4; RMP 2020b).

To maintain external validity, grammatical errors and casual writing style were retained in the evaluations. However, slight changes were made to the evaluations to remove information about professor name, age, and gender as well as information about course discipline (ex: mathematics, history). Similar to previous studies of RMP evaluation effects (Boswell 2016; Kowai-Bell et al. 2011, 2012), the number of comments was held constant between the mostly positive and mostly negative conditions; five evaluations were included for each condition. Average evaluation word count was the same in each condition.

A sample negative evaluation was: "The class is difficult. This class had me pretty lost (along with most others I asked). Rambled a bit and didn't really have definite answers for anything, so it was tough to know how to answer exam questions and do assignments."

A sample positive evaluation was: "instructions were always crystal clear. Teaches the material very well and easily- it's hard not to learn from in this class. The prof was funny and I always looked forward to going to class. Down to earth, passionate about the subject, always willing to help outside of class."

### 3.3.3 Manipulation check

To check that participants read the evaluations and perceived the valence of the comments presented, participants utilized a 9-point, Likert-type scale ranging from 1 (*very negative*) to 9 (*very positive*) to describe the professor's evaluations.

### 3.3.4 Decision making

Participants utilized a 9-point, Likert-type scale ranging from 1 (*strongly disagree*) to 9 (*strongly agree*) to rate their agreement with statements regarding their decision about the professor's course: (1) how seriously they would take these evaluations when making a decision to take the professor's course, and (2) intention to take the professor's course.

### 3.3.5 Self-efficacy

Using the same 9-point, Likert-type scale above, participants rated their self-efficacy for four course-related behaviors (attend each class session, take class notes, study for the course, and invest effort into the course) and four interpersonal-related

behaviors (ask the professor a question during class, answer the professor's question during class, ask the professor for help, and form a good relationship with the professor).

### 3.3.6 Expected academic outcomes

Participants also rated their expectation that they could earn an A in the professor's course and learn in the course using the same 9-point, Likert-type scale as above.

### 3.3.7 Teaching evaluation behaviors and explicit beliefs

Participants utilized a 9-point, Likert-type scale ranging from 1 (*never*) to 9 (*always*) to indicate the frequency at which they have completed or plan to complete (1) UASET and (2) RMP evaluations. They also used this scale to indicate how often they would read evaluations from (1) UASET and (2) RMP during their course decision making. Moreover, they utilized a 9-point, Likert-type scale ranging from 1 (*not seriously*) to 9 (*very seriously*) to indicate how seriously they would consider (1) UASET and (2) RMP evaluations in their course decision making.

### 3.3.8 Demographic questionnaire

Participants also completed a demographic questionnaire that contained open-ended questions about age, sex, ethnicity, and college class.

## 3.4 Procedure

To mimic real-world RMP use, data collection occurred online via Qualtrics, an online tool for the administration of research studies (Qualtrics 2019). Following informed consent, participants completed the College Self-Efficacy Inventory and were then randomly assigned to the RMP mostly positive, RMP mostly negative, UASET mostly positive, or UASET mostly negative condition. Participants then received a set of five evaluations of a fictitious professor. Participants were instructed to imagine that they were using these evaluations to inform their decision making and self-efficacy about an upcoming course with the professor and then complete the rating scales. The vignette and role-play method has been utilized in other studies investigating the effect of RMP (Boswell 2016; Kowai-Bell et al. 2012). This method increases the external validity of the experimental design while maintaining high internal validity (Atzmüller and Steiner 2010). Moreover, role-play addresses ethical considerations associated with exposure of students to an experimental condition involving a purposefully underperforming professor; they also control for potential bias a student may have from previous interactions with a particular real-life professor (Gotlieb 2011). Following completion of the rating scales regarding their decision-making, self-efficacy, and expected academic outcomes for the professor's course, participants completed measures of teaching evaluation beliefs and behaviors as well as the demographic questionnaire.

### 3.5 Statistical analyses

To investigate research question 1, a series of *t*-tests examined how often participants complete and would read UASET compared to RMP as well as how seriously they would consider each source. To investigate question 2, a multivariate analysis of variance (MANOVA) investigated question 2 regarding the effects of evaluation source, valence, and their interaction on participants' decision making about a professor's course. Finally, a multivariate analysis of covariance (MANCOVA) investigated questions 3 and 4 regarding the effects of evaluation source, valence, and their interaction on participants' course and interpersonal self-efficacy as well as their expected academic outcomes for a professor's course.

## 4 Results

### 4.1 Teaching evaluations: behaviors and explicit beliefs

Table 1 presents descriptive statistics for how often participants complete or plan to complete RMP and UASET, how often they read RMP and how often they would read UASET if it were available, and how seriously they would consider RMP and UASET evaluations in their course decision making. A series of *t*-tests (see Table 1) found that when compared to RMP, participants complete or plan to complete UASET more often, would read UASET more often, and would consider UASET more seriously in their course decision making.

### 4.2 Manipulation check

A manipulation check determined if participants perceived the valence of the comments presented. Item content was, "How would you describe the professor's evaluations?" with close-ended response options ranging from 1 (*very negative*) to 9 (*very positive*). A one-way analysis of variance (ANOVA) evaluated the manipulation, comparing responses to this single item between conditions. The ANOVA was

**Table 1** Descriptives and differences in completion, viewing, and consideration

	<i>M</i>	<i>SD</i>	<i>n</i>	<i>t</i>	<i>p</i>
How often do you complete			222	10.79	<.001
RMP	3.63	2.63			
UASET	5.94	2.71			
How often do you view			222	4.11	<.001
RMP	6.32	2.60			
UASET	7.01	2.06			
How seriously do you consider			222	4.86	<.001
RMP	6.56	1.87			
UASET	6.98	1.77			

significant,  $F(1, 221) = 190.07, p < .001$ ; as predicted, participants rated the evaluations for the mostly positively evaluated professor ( $M = 7.05$ ) as significantly higher than those of the mostly negatively evaluated professor ( $M = 3.75$ ).

### 4.3 Source and valence effects on decision making

A MANOVA then examined the effect of source of evaluations and their valence on participants' decision making about the course (how seriously they would consider the evaluations when making a decision to take the professor's course and intention to take the course). Evaluation source was coded as either RMP or UASET and evaluation valence was coded as either mostly positive or mostly negative. Because Box's test was significant, Pillai's trace was selected as the coefficient for interpretation. Results indicated marginal (i.e., nonsignificant using standard criterion) difference based on source (Pillai's trace = .03;  $p = .05$ ; partial  $\eta^2 = .03$  or small effect size) but significant difference by valence (Pillai's trace = .48;  $p < .01$ ; partial  $\eta^2 = .48$  or large effect size). Follow-up testing using one-way ANOVA indicated that significant difference by valence was only found for intention to take the class ( $F = 190.45$ ;  $p < .01$ ). The interaction between source and valence was included in the analysis, but results were not significant (Pillai's trace = .00).

### 4.4 Source and Valence Effects on Self-Efficacy and Academic Outcomes

Next, MANCOVA with college self-efficacy as a control variable investigated the effect of evaluation source and valence on the course-related self-efficacy dependent variables (intention to attend every class, take notes, study, and invest effort), interpersonal-related self-efficacy (ask the professor a question in class, answer the professor's question in class, ask for help, form a good relationship), and expected academic outcomes (earn an A, and learn). In preliminary analysis, Box's test again suggested potential violation of the assumption of homogeneity of covariances, so Pillai's trace was again selected as the coefficient for interpretation. Results revealed significant group differences based on the covariate of college self-efficacy (Pillai's trace = .29;  $p < .01$ ; partial  $\eta^2 = .29$  or large effect size). Valence significantly differentiated responses (Pillai's trace = .26;  $p < .01$ ; partial  $\eta^2 = .26$  or large effect size); however, source did not (Pillai's trace = .05). The interaction between source and valence was explored but was not significant (Pillai's trace = .04).

Follow-up testing employing a series of univariate analyses of covariance (ANCOVA) indicated that significant differences based on valence (using alpha of .005 based on a Bonferroni correction) were present for "I will study for the professor's class" ( $F = 9.89$ ), "I will invest effort into this class" ( $F = 13.59$ ), "I will ask the professor for help if I need it" ( $F = 9.53$ ), "I will have a good relationship with the professor" ( $F = 27.29$ ), "I will be able to earn an A in this class" ( $F = 31.71$ ), and "I will learn a lot in this class" ( $F = 68.78$ ). Means and standard errors are presented in Table 2.

**Table 2** Estimated marginal means and standard errors for evaluation valence effects

	Mostly negative		Mostly positive	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Attend each class session	8.06	.11	7.39	.11
Take class notes	8.02	.12	8.21	.12
Study for the course	7.37	.14	7.99	.14
Invest effort into the course	7.36	.14	8.12	.15
Ask the professor a question during class	6.34	.18	6.67	.18
Answer the professor's question during class	5.81	.19	6.55	.19
Ask the professor for help	6.78	.17	7.54	.17
Form a good relationship with the professor	5.68	.18	7.03	.18
Earn A	6.57	.15	7.73	.15
Learn	5.90	.16	7.76	.16

Covariate: College Self-Efficacy = 103.17

## 5 Discussion

Little is known about the effects of RMP evaluations relative to those of UASET in students. Therefore, this study investigated differences in use and consideration of these evaluation sources (RMP versus UASET). It also investigated the effects of evaluation source, evaluation valence (mostly positive or mostly negative), and their interaction on students' decision making, self-efficacy for a course, and expectations about academic outcomes.

### 5.1 Behaviors and explicit beliefs

With regard to Research Question 1, participants reported that they complete or plan to complete UASET more often than RMP. Additionally, they reported that if given access to UASET, they would read it more frequently and consider it more seriously in their course decision making than RMP. Counter to previous research (Hayes and Prus 2014), these findings suggest that students may perceive UASET to be more valid than RMP. Moreover, they support the argument that students turn to RMP to inform their course selection because they lack access to UASET.

### 5.2 Source and valence effects

In addressing Research Question 2, source's influence on decision making was not supported, but results did indicate a valence effect for intention to take the course. Specifically, students reported higher intention to take a course when presented with mostly positive comments regarding said course. While the results regarding considering the comments seriously in decision making did not meet criteria for statistical significance, the difference observed was consistent with expectations, with students

expecting to take comments more seriously when they were of a mostly positive valence. No evidence emerged for an interaction between source and valence in influencing course decision making. The results relevant to Research Question 3 again supported differences based on valence but not source in students' course- and interpersonal-related self-efficacy while controlling for college self-efficacy. Regarding Research Question 4, valence, but not source, influenced expected academic outcomes for the professor's course. Not surprisingly, those exposed to mostly positive evaluations expected to earn a higher grade and learn more from the professor.

The effect of valence on participants' self-efficacy for course behavior and interpersonal behavior with the professor has significant implications for student learning. Self-efficacy is predictive of willingness to engage in a task as well as persistence at the task when obstacles occur (Bandura 1977). The current results indicate students may enter courses with fewer self-defeating beliefs or higher determination when exposed to previous comments praising elements of those courses. We explore the implications of this in greater detail below.

The discrepancy between participants' explicit attitudes toward RMP and UASET and their response to RMP and UASET content is interesting. Participants reported that they view, complete, and consider more seriously UASET compared to RMP. However, evaluation source did not significantly affect the dependent variables; participants' explicit discrimination between the evaluation sources did not align with their implicit response to them. Although interesting, this inconsistency is not highly unusual given that individuals' explicit and implicit attitudes at times conflict. Moreover, individuals are often unaware of the disagreement between their explicit and implicit attitudes as well as the effect of their implicit attitudes on behavior (e.g., Karpen et al. 2012). Although participants reported that they consider UASET more seriously than RMP, RMP had just as significant of an influence on their self-efficacy and expectations as UASET. This suggests that students' use of RMP should be taken seriously.

### 5.3 Implications

Students' self-efficacy and expectations about their collegiate abilities are related to a number of variables important for student learning such as motivation (Turner et al. 2009) and persistence (Wright et al. 2012) as well as student academic performance (Chemers et al. 2001). Therefore, RMP valence's ability to influence participants' self-efficacy has significant implications for students' learning and academic achievement. For students who are exposed to positive RMP evaluations, these results suggest clear benefits: increased confidence for their ability to study for and invest effort into the course as well as their ability to form a good, beneficial relationship with the professor. These self-efficacious beliefs would likely translate into increased course effort and greater interpersonal engagement with the professor, both factors predictive of student learning and academic success (Ingraham et al. 2018; Komarraju et al. 2010; Svanum and Aigner 2011; Svanum and Bigatti 2006).

The implications of negative RMP evaluation exposure, however, are more concerning. RMP evaluations serve as vicarious experiences with a particular professor;

these experiences will have their greatest effect on students' self-efficacy when they are unfamiliar with the professor in question. Students vicariously exposed to others' perceived failures with a particular professor may become less efficacious for that course; in turn, these beliefs may contribute to academically self-defeating behaviors such as disengagement from coursework and reluctance to communicate with faculty. These effects are especially troublesome in this context: on RMP, evaluations are anonymous, providing readers with limited ability to appraise their credibility. Those with a sincere interest in learning may be exposed to and influenced by the evaluations of students who are irritated that they did not receive an "easy A." In contrast, students gathering opinions via traditional word-of-mouth are better positioned to assess the source's credibility to determine if it must be taken with a grain or a pinch of salt.

Although students who turn to RMP cannot evaluate the credibility of each individual evaluation, they may still be able to critically examine its content. Students who choose to utilize the website may benefit from an analytical, systematic approach to evaluating RMP evaluations; we detail this approach below.

#### 5.4 Limitations and future directions

This study, while informative, possessed several noteworthy limitations. For example, the investigation used convenience sampling, and therefore, the sample's composition may limit generalizability of results. Over half of the participants identified as first-year or freshmen students. Due to their limited college experience, these students have had fewer opportunities to see their efforts lead to successful college performance (e.g., mastery experiences). Because they have less information about their ability to perform in a collegiate setting, exposure to others' experiences via RMP may exert greater influence on their self-efficacy than among their peers with more college credits. With greater experience linking personal performance to academic success, self-efficacy in sophomores, juniors, and seniors, may be less vulnerable to the effects of RMP evaluations.

Racial and ethnic make-up of the sample also presents a limitation. The majority of participants identified as White and non-Hispanic; while the demographics of this sample are representative of this university's student population, results from this sample may not generalize well to more ethnically and racially diverse universities. Because many aspects of the collegiate experience differ between ethnic minority and White students (Anderson and Smith 2005), additional research to determine if these results replicate in more diverse student populations is merited. Moreover, geography may limit generalizability of the results. RMP is a tool utilized by students internationally; however, this sample was recruited from one American university. The effect of RMP on decision making, self-efficacy, and expectations in this sample may not be reflective of the dynamics of RMP use or use of similar professor rating websites (e.g., RateMyTeachers.com Australia, Canada, Ireland, New Zealand, and United Kingdom) outside the United States because attitudes toward such sites likely vary between cultural groups. For example, French courts have banned the use

of specific professors' names on teaching and professor-rating websites (Gross 2008; Masnick 2008).

Absence of information about professor gender in the evaluation stimuli is a further limitation of this study. While the literature about gender's relationship with UASET and RMP ratings is mixed (e.g., some studies finding no relationship while others find bias in favor of men or women), the extant research indicates that under some circumstances, gender matters in students' evaluations of professors (e.g., Basow 1995; Punyanunt-Carter and Carter 2015; Joye and Wilson 2015; Rosen 2017; Sohr-Preston et al. 2016; Theyson 2015). If professor gender influences the evaluations students provide, it stands to reason that professor gender may also influence how students respond to evaluations. This may be particularly relevant to students' self-efficacy for the student-faculty relationship given that students may have different interpersonal dynamic expectations for women professors (El-Alayli et al. 2018).

## 5.5 Recommendations

Although RMP evaluations are subject to several biases, many students still utilize the website to inform their course decision making. Students looking to RMP for help with course decisions may be able to glean useful information from the site by utilizing a systematic approach similar to that suggested for faculty members who visit RMP (e.g., Sohr-Preston et al. 2016). First, it seems advisable to dismiss comments unrelated to the subject matter or professor's teaching ability; for example, "she dresses fancy and hip" and "I'd rather clean portable toilets than take his class." Next, a productive approach would include dismissing statements such as "she's literally the best professor ever" because such comments provide no concrete information. Next, one should seek comments related to qualities of effective teaching (e.g., Yermack and Forsyth 2016). For example, effective professors are typically prepared for class. An abundance of comments such as "he was always prepared and provided us with objectives for the lecture" about a particular professor would be consistent with this quality.

## 6 Conclusion

While many studies have examined relationships between the RMP variables of easiness, clarity, and quality as well as sources of bias in RMP evaluations, fewer have experimentally investigated the potential impact of these evaluations on students. Although RMP advertises itself as a "fun" way to gather information about professors and courses, these findings add to a growing body of literature suggesting that its use has significant implications for students. Given that content of these evaluations influences students' beliefs about course effort and the student-professor relationship, students exposed to negative RMP evaluations stand to develop self-defeating thoughts and behaviors before they ever enter the classroom. In turn, these may threaten students' very goals to learn and progress toward a degree. Students

who wish to utilize RMP may benefit from a systematic approach and critical perspective of its content.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

- Anderson, K. J., & Smith, G. (2005). Students' preconceptions of professors: Benefits and barriers according to ethnicity and gender. *Hispanic Journal of Behavioral Sciences*, *27*, 184–201.
- Atzmüller, C., & Steiner, P. M. (2010). Experimental vignette studies in survey research. *Methodology*, *6*, 128–138. <https://doi.org/10.1027/1614-2241/a000014>.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*, 191–215. <https://doi.org/10.1037/0033-295X.84.2.191>.
- Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist*, *37*, 122–147. <https://doi.org/10.1037/0003-066X.37.2.122>.
- Bandura, A. (1989). Human agency in social cognitive theory. *American Psychologist*, *44*, 1175–1184. <https://doi.org/10.1037/0003-066X.44.9.1175>.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.
- Basow, S. A. (1995). Student evaluations of college professors: Gender matters. *Journal of Educational Psychology*, *87*, 656–665. <https://doi.org/10.1037/0022-0663.87.4.656>.
- Bleske-Rechek, A., & Fritsch, A. (2011). Student consensus on Ratemyprofessors.com. *Practical Assessment, Research & Evaluation*, *16*(1), 18.
- Boehmer, D. M., & Wood, W. C. (2017). Student vs. faculty perspectives on quality instruction: Gender bias, “hotness”, and “easiness” in evaluating teaching. *Journal of Education for Business*, *92*, 173–178. <https://doi.org/10.1080/08832323.2017.1313189>.
- Boswell, S. S. (2016). Ratemyprofessors is hogwash (but I care): Effects of Ratemyprofessors and university-administered teaching evaluations on professors. *Computers in Human Behavior*, *56*, 155–162. <https://doi.org/10.1016/j.chb.2015.11.045>.
- Brown, M. J., Baillie, M., & Fraser, S. (2009). Rating Ratemyprofessors.com: A comparison of online and official student evaluations of teaching. *College Teaching*, *57*, 89–92. <https://doi.org/10.3200/CTCH.57.2.89-92>.
- Brown, C. L., & Kosovich, S. M. (2015). The impact of professor reputation and course attributes on student course selection. *Research in Higher Education*, *56*, 496–509. <https://doi.org/10.1007/s1162-014-9356-5>.
- Buck, G. H. (2004). Vox populi? *The Alberta Journal of Educational Research*, *50*(3), 217–220.
- Carnevale, D. (2006). University of Saskatchewan fires tenured professor accused of maligning colleagues on Ratemyprofessors website. *The Chronicle of Higher Education*, *52*(27), A28.
- Caron, P. (2006). *Tenured professor fired for maligning colleagues on RateMyProfessors.com*. [http://taxprof.typepad.com/taxprof\\_blog/2006/02/tenured\\_prof.html](http://taxprof.typepad.com/taxprof_blog/2006/02/tenured_prof.html).
- Chemers, M. M., Hu, L.-T., & Garcia, B. F. (2001). Academic self-efficacy and first year college student performance and adjustment. *Journal of Educational Psychology*, *93*, 55–64. <https://doi.org/10.1037/0022-0663.93.1.55>.
- Clayson, D. E. (2014). What does ratemyprofessors.com actually rate? *Assessment & Evaluation in Higher Education*, *39*, 678–698. <https://doi.org/10.1080/02602938.2013.861384>.
- Cohen, J. (1988). *Statistical power for the behavioral sciences*. New Jersey: Lawrence Erlbaum Associates.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155–159. <https://doi.org/10.1037/0033-2909.112.1.155>.
- Coladarci, T., & Kornfeld, I. (2007). RateMyProfessor.com versus formal in-class student evaluations of teaching. *Practical Assessment, Research, & Evaluation*, *12*(1), 6.

- Davison, E., & Price, J. (2009). How do we rate? An evaluation of online student evaluations. *Assessment & Evaluation in Higher Education*, 34, 65. <https://doi.org/10.1080/02602930801895695>.
- Edwards, A., & Edwards, C. (2013). Computer-mediated word-of-mouth communication: The influence of mixed reviews on student perceptions of instructors and courses. *Communication Education*, 62(4), 412–424. <https://doi.org/10.1080/03634523.2013.800217>.
- Edwards, C., Edwards, A., Qing, Q., & Wahl, S. T. (2007). The influence of computer-mediated word-of-mouth communication on student perceptions of instructors and attitudes toward learning course content. *Communication Education*, 56(3), 255–277. <https://doi.org/10.1080/03634520701236866>.
- Edwards, A., Edwards, C., Shaver, C., & Oaks, M. (2009). Computer-mediated word-of-mouth communication on Ratemyprofessors.com: Expectancy effects on student cognitive and behavioral learning. *Journal of Computer-Mediated Communication*, 14(2), 368–392. <https://doi.org/10.1111/1/j.1083-6101.2009.01445.x>.
- El-Alayli, A., Hansen-Brown, A. A., & Ceynar, M. (2018). Dancing backward in high heels: Female professors experience more work demands and special favor requests, particularly from academically entitled students. *Sex Roles*. <https://doi.org/10.1007/s11199-017-0872-6>.
- Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavioral Research Methods, Instruments, & Computers*, 28, 1–11. <https://doi.org/10.3758/BF03203630>.
- Faul, F., Erdfelder, E., Lang, A., & Buchner, A. (2007). G\*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175–191. <https://doi.org/10.3758/BF03193146>.
- Felton, J., Mitchell, J., & Stinson, M. (2004). Web-based student evaluations of professors: the relations between perceived quality, easiness and sexiness. *Assessment & Evaluation in Higher Education*, 29, 91–108. <https://doi.org/10.1080/0260293032000158180>.
- Field, J. R., Bergiel, B. J., & Viosca, R. C. (2008). An exploratory study of student use and perceived accuracy of faculty evaluations on Ratemyprofessors.com. *Competition Forum*, 6, 378–383.
- Flaherty, C. (2018). Bye, bye chili pepper. <https://www.insidehighered.com/news/2018/07/02/rate-my-professors-ditches-its-chili-pepper-hotness-quotient>.
- Freng, S., & Webber, D. (2009). Turning up the heat on online teaching evaluations: Does “hotness” matter? *Teaching of Psychology*, 36, 189–193. <https://doi.org/10.1080/00986280902959739>.
- George, D., Dixon, S., Stansal, E., Gelb, S. L., & Pheri, T. (2008). Time diary and questionnaire assessment of factors associated with academic and personal success among university undergraduates. *Journal of American College Health*, 56(6), 706–715. <https://doi.org/10.3200/JACH.56.6.706-715>.
- Good, M., & Adams, G. R. (2008). Linking academic social environments, ego-identity formation, ego virtues, and academic success. *Adolescence*, 43(170), 221–236.
- Gotlieb, J. (2011). Classroom entertainment, upgrades, and student evaluations of business professors. *Journal of the Academy of Business Education*, 13, 118–136.
- Gross, P. (2008). French courts say “non” to “rate my teacher”. <https://www.techradar.com/news/internet/web/french-courts-say-non-to-rate-my-teacher-257018>.
- Gump, S. E. (2005). The cost of cutting class: Attendance as a predictor of student success. *College Teaching*, 53, 21–26. <https://doi.org/10.3200/CTCH.53.1.21-26>.
- Hartman, K. B., Hunt, J. B., & Childers, C. Y. (2013). Effects of eWOM valence: examining consumer choice using evaluations of teaching. *Journal of Behavioral Studies in Business*, 6, 1–12.
- Hayes, M. W., & Prus, J. (2014). Student use of quantitative and qualitative information on Ratemyprofessors.com for course selection. *College Student Journal*, 48, 675–688.
- Hayles, N. K., & Gessler, N. (2005). Rating RateMyProfessors.com. *English Studies in Canada*, 31(4), 6–9. <https://doi.org/10.1353/esc.2007.0045>.
- Hinz, A. (2011). Attitudes of German teachers and students toward public online ratings of teacher quality. *Electronic Journal of Research in Educational Psychology*, 9(2), 745–764.
- Ingraham, K. C., Davidson, S. J., & Yonge, O. (2018). Student-faculty relationships and its impact on academic outcomes. *Nurse Education Today*, 71, 17–21. <https://doi.org/10.1016/j.nedt.2018.08.021>.
- Jaschik, S. (2006). *RateMyProfessors: hidden camera edition*. <https://www.insidehighered.com/news/2006/11/02/rmp>.
- Johnson, R. R., & Crews, A. D. (2013). My professor is hot! Correlates of Ratemyprofessors.com ratings for criminal justice and criminology faculty members. *American Journal of Criminal Justice*, 38(4), 639–656. <https://doi.org/10.1007/s12103-012-9186-y>.
- Joye, S. W., & Wilson, J. H. (2015). Professor age and gender affect student perceptions and grades. *Journal of the Scholarship of Teaching and Learning*, 15(4), 126–138.

- Karpen, S. C., Jia, L., & Rydell, R. J. (2012). Discrepancies between implicit and explicit attitude measures as an indicator of attitude strength. *European Journal of Social Psychology, 42*, 24–29. <https://doi.org/10.1002/ejsp.849>.
- Katrevich, A. V., & Aruguete, M. S. (2017). Recognizing challenges and predicting success in first-generation university students. *Journal of STEM Education: Innovations and Research, 18*(2), 40–44.
- Kindred, J., & Mohammed, S. N. (2005). He will crush you like an academic ninja!: Exploring teacher ratings on RateMyProfessors.com. *Journal of Computer-Mediated Communication, 10*(3), 9. <https://doi.org/10.1111/j.1083-6101.2005.tb00257.x>.
- Komaraju, M., Musulkin, S., & Bhattacharya, G. (2010). Role of student-faculty interactions in developing college students' academic self-concept, motivation, and achievement. *Journal of College Student Development, 51*, 332–342. <https://doi.org/10.1353/csd.0.0137>.
- Kowai-Bell, N., Guadagno, R. E., Little, T., & Ballew, J. L. (2012). Professors are people too: The impact of informal evaluations of professors on students and professors. *Social Psychology of Education, 15*, 337–351. <https://doi.org/10.1007/s11218-012-9181-7>.
- Kowai-Bell, N., Guadagno, R. E., Little, T., Preiss, N., & Hensley, R. (2011). Rate my expectations: How online evaluations of professors impact students' perceived control. *Computers in Human Behavior, 27*, 1862–1867. <https://doi.org/10.1016/j.chb.2011.04.009>.
- Kwak, D. W., Sherwood, C., & Tang, K. K. (2018). Class attendance and learning outcome. *Empirical Economics, 1*, 1. <https://doi.org/10.1007/s00181-018-1434-7>.
- Landry, N. M., Kurkul, K. E., & Poirier, C. R. (2010). Online faculty rating sites: Examining how students perceive and use RateMyProfessors.com. *Psi Chi Journal of Undergraduate Research, 15*, 139–142.
- Lawson, R. A., & Stephenson, E. F. (2005). Easiness, Attractiveness, and Faculty Evaluations: Evidence from RateMyProfessors.com. *Atlantic Economic Journal, 33*(4), 485–486. <https://doi.org/10.1007/s11293-005-2873-z>.
- Legg, A. M., & Wilson, J. H. (2012). RateMyProfessors.com offers bias evaluations. *Assessment & Evaluation in Higher Education, 37*, 89–97. <https://doi.org/10.1080/02602938.2010.507299>.
- Lewandowski, G., Higgins, E., & Nardon, N. N. (2012). Just a harmless website?: An experimental examination of RateMyProfessors.com's effect on student evaluations. *Assessment & Evaluation in Higher Education, 37*, 987–1002. <https://doi.org/10.1080/02602938.2011.594497>.
- Liu, J., Hu, J., & Furutan, O. (2013). The influence of student perceived professors' "hotness" on expertise, motivation, learning outcomes, and course satisfaction. *Journal of Education for Business, 88*, 94–100. <https://doi.org/10.1080/08832323.2011.652695>.
- Masnick, M. (2008). *France decides that expressing an opinion about your teachers should be illegal*. <https://www.techdirt.com/articles/20080304/005526425.shtml>.
- Montell, G. (2006). *The art of the bogus rating*. <https://www.chronicle.com/article/The-Art-of-the-Bogus-Rating/46887>.
- Murray, K. B., & Zdravkovic, S. (2016). Does MTV really do a good job of evaluating professors? An empirical test of the internet site RateMyProfessors.com. *Journal of Education for Business, 91*, 138–147. <https://doi.org/10.1080/08832323.2016.1140115>.
- Otto, J., Sanford, D. A., & Ross, D. N. (2008). Does RateMyProfessor.com really rate my professor? *Assessment & Evaluation in Higher Education, 33*, 355–368. <https://doi.org/10.1080/02602930701293405>.
- Pajares, F. (1997). Current directions in self-efficacy research. In M. Maehr & P. R. Pintrich (Eds.), *Advances in motivation and achievement* (pp. 1–49). Greenwich: JAI Press.
- Punyanunt-Carter, N., & Carter, S. L. (2015). Students' gender bias in teaching evaluations. *Higher Learning Research Communications, 5*(3), 28–37. <https://doi.org/10.18870/hlrc.v5i3.234>.
- Qualtrics. (2019). *Qualtrics Research Core*. <https://www.qualtrics.com/research-core/>.
- RateMyProfessors.com. (2020a). *About RateMyProfessors.com*. <http://www.ratemyprofessors.com/About.jsp>.
- RateMyProfessors.com. (2020b). *Can you explain the rating scale?* <https://help.ratemyprofessors.com/article/31-rating-scale>.
- Reed, D. K., Rimel, H., & Hallett, A. (2016). Note-taking interventions for college students: A synthesis and meta-analysis of the literature. *Journal of Research on Educational Effectiveness, 9*, 307–333. <https://doi.org/10.1080/19345747.2015.1105894>.
- Reid, L. D. (2010). The role of perceived race and gender in the evaluation of college teaching on RateMyProfessors.com. *Journal of Diversity in Higher Education, 3*, 137–152. <https://doi.org/10.1037/a0019865>.
- Ritter, K. (2008). E-evaluating learning: Rate My Professors and public rhetorics of pedagogy. *Rhetoric Review, 3*, 259–280. <https://doi.org/10.1080/07350190802126177>.

- Rosen, A. S. (2017). Correlations, trends and potential biases among publicly accessible web-based student evaluations of teaching: a large-scale study of RateMyProfessors.com data. *Assessment & Evaluation in Higher Education*, 43, 31–44. <https://doi.org/10.1080/02602938.2016.1276155>.
- Schudde, L. (2019). Short- and long-term impacts of engagement experiences with faculty and peers at community colleges. *Review of Higher Education*, 42, 385–426. <https://doi.org/10.1353/rhe.2019.0001>.
- Simonton, T. (2018). *No more chili pepper: RateMyProfessors ditches 'hotness' ratings*. <https://www.chronicle.com/article/No-More-Chili-Pepper-/243809>.
- Sohr-Preston, S. L., Boswell, S. S., McCaleb, K., & Robertson, D. (2016). Professor gender, age, and “hotness” in influencing college students’ generation and interpretation of professor ratings. *Higher Learning Research Communications*. <https://doi.org/10.18870/hlrc.v6i3.328>.
- Solberg, V. S., O’Brien, K., Villarreal, P., Kennel, R., & Davis, B. (1993). Self-efficacy of Hispanic college students: Validation of the College Self-Efficacy instrument. *Hispanic Journal of Behavioral Sciences*, 15, 80–95. <https://doi.org/10.1177/07399863930151004>.
- Solberg, V. S., & Villarreal, P. (1997). Examination of self-efficacy, social support, and stress as predictors of psychological and physical distress among Hispanic college students. *Hispanic Journal of Behavioral Sciences*, 19, 182–201. <https://doi.org/10.1177/07399863970192006>.
- Stonebraker, R. J., & Stone, G. S. (2015). Too old to teach? The effect of age on college and university professors. *Research in Higher Education*, 56, 793–812. <https://doi.org/10.1007/s11162-015-9374-y>.
- Subtirelu, N. C. (2015). “She does have an accent but...”: Race and language ideology in students’ evaluations of mathematics instructors on RateMyProfessors.com. *Language in Society*, 44, 35–62. <https://doi.org/10.1017/S0047404514000736>.
- Svanum, S., & Aigner, C. (2011). The influences of course effort, mastery and performance goals, grade expectancies, and earned course grades on student ratings and course satisfaction. *British Journal of Educational Psychology*, 81, 667–679. <https://doi.org/10.1111/j.2044-8279.2010.02011.x>.
- Svanum, S., & Bigatti, S. M. (2006). The influences of course effort and outside activities on grades in a college course. *Journal of College Student Development*, 47, 564–576. <https://doi.org/10.1353/csd.2006.0063>.
- Theyson, K. C. (2015). Hot or not: The role of instructor quality and gender on the formation of positive illusions among students using RateMyProfessors.com. *Practical Research and Assessment*, 20(4). <http://pareonline.net/getvn.asp?v=20&n=4>
- Timmerman, T. (2008). On the validity of RateMyProfessors.com. *Journal of Education for Business*, 84(1), 55–61. <https://doi.org/10.3200/JOEB.84.1.55-61>.
- Tinto, V. (1999). Taking retention seriously: Rethinking the first year of college. *NACADA Journal*, 19(2), 5–9. <https://doi.org/10.12930/0271-9517-19.2.5>.
- Turner, E. A., Chandler, M., & Heffer, R. W. (2009). The influence of parenting styles, academic motivation, and self-efficacy on academic performance and college students. *Journal of College Student Development*, 50, 337–346. <https://doi.org/10.1353/csd.0.0073>.
- Wright, S. L., Jenkins-Guarnieri, M. A., & Murdock, J. L. (2012). Career development among first-year college students: College self-efficacy, student persistence, and academic success. *Journal of Career Development*, 40, 292–310. <https://doi.org/10.1177/0894845312455509>.
- Ratemyprofessors.com (n.d.). *Ratings categories*. <http://www.ratemyprofessors.com/Categories.jsp>.
- Yermack, J., & Forsyth, D. R. (2016). Students’ implicit theories of university professors. *Scholarship of Teaching and Learning in Psychology*, 2, 169–178.
- Zajacova, A., Lynch, S. M., & Espenshade, T. J. (2005). Self-efficacy, stress, and academic success in college. *Research in Higher Education*, 46(6), 677–706. <https://doi.org/10.1007/s11162-004-4139-z>.

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Stefanie S. Boswell** is a Professor of Psychology at University of the Incarnate Word in San Antonio, TX. Her research interests include students’ and professors’ attitudes toward higher education, informal student evaluations of teaching, self-efficacy, and teaching and learning.

**Sara L. Sohr-Preston** is an Associate Professor of Psychology at Southeastern Louisiana University. Her research interests include parenting, family building (pregnancy, infertility, miscarriage, and adoption), infancy and early childhood, and college student health and learning.