Skills vs Rapport: Do Students in STEM and Humanities Courses Value Different Characteristics in Their Professors?

Nicole Sanders

Skills vs Rapport: Do Students in STEM and Humanities Courses Value Different Characteristics in Their Professors?

Student evaluation of teaching (SET) is vital to higher education, affecting both teachers and students. For students, filling out SETs is an easy and anonymous way to give truthful feedback to instructors. Although students may not directly benefit from this feedback, their opinions may benefit future students' educational experience. The teachers themselves are able to learn what works and what does not from the SETs, thus giving them an opportunity to adjust accordingly. This process supports the improvement of teaching by providing a chance for honest communication that benefits students and instructors alike. Outside of the classroom, the results of SETs may influence promotion and salary decisions (McKeachie, 1997). They may also be used during the audit of higher education institutions to determine accreditation (Johnson, 2000). SETs are used in almost every institution of higher education throughout the world (Zabaleta, 2007), and as a result, there are many different methods for how SETs are collected and analyzed (Spooren et al., 2013). In addition, SETs used at universities and colleges vary widely in what aspects of teaching are evaluated.

To address this variability, standardized teaching evaluations have been developed including the Students' Evaluation of Education Quality (SEEQ; Marsh, 1982; Marsh et al., 2009), Students' Evaluation of Teaching Effectiveness Rating Scale (SETERS; Toland & De Ayala, 2005), and the Teaching Proficiency Item Pool (Barnes et al., 2008). One popular standardized measure of teaching effectiveness is the Teacher Behavior Checklist (TBC) (Keeley et al., 2010; Keeley et al., 2006). The TBC is especially powerful because it provides behavioral anchors for each of its items; for every characteristic, the TBC provides behaviors that professors can adopt if they want to improve their score (Keeley et al., 2006). Students rate teachers on a 5 point Likert scale, from always exhibiting a behavior to never doing it, on behaviors like Accessible (Posts office hours, gives out phone number and e-mail information) (Keeley et al., 2006). Furthermore, previous research has shown links to enhanced student satisfaction (Darawong & Sandmaung, 2019), a way to develop faculty teaching and learning strategies (McGovern & Miller, 2008), and optimal learning conditions for students and faculty (Lammers & Smith, 2008).

The 28 items in the TBC are split into two subscales: the Caring and Supportive Subscale and Professional Competency and Communication Skills Subscale. The former includes items like rapport (makes class laugh through jokes and funny stories, initiates and maintains class discussions, knows student names, and interacts with students before and after class) while the latter includes knowledgeable about subject matter (easily answers students' questions, does not read straight from the book or notes, and uses clear and understandable examples). Using the TBC, Addison et al. (2015) created two categories for measuring teaching effectiveness by turning the Caring and Supportive Subscale into a rapport category and the Professional Competency and Communication Skills Subscale into a skills category.

The skills category measures if the instructor is clear, enthusiastic, organized, knowledgeable, and engaging, while the rapport category looks at if the instructor is personable, approachable/accessible, respectful, humorous/fun, and helpful (Addison et al., 2015). Addison et al. (2015) examined whether the attributes (from the skills and rapport categories) that are important to students differ based on the class they are evaluating. Specifically, Addison et al., (2015) compared the comments from student evaluations of introductory psychology classes and psychology statistics classes. After analysis, the study found that for students in both classes, a teacher's level of skill and rapport were equally important. There is reason to believe, however, that this may not be the case if students in Science, Technology, Engineering, and Mathematics (STEM) and humanities courses were compared due to differences between students in these disciplines.

For example, personality differences can predict choice of academic discipline (Berring et al., 2017). Specifically, Berring et al. found students that chose to study the humanities had higher scores on intrinsic value orientation. In contrast, students that pursued STEM were found to have higher scores on extrinsic value orientation (Berring et al., 2017). People who have an intrinsic value orientation are motivated by the goal of benefiting psychological well-being through activities like close relationships, while people who have an extrinsic value orientation are motivated by external signs of worth like money (Berring et al., 2017). Humanities students with intrinsic value orientations also prioritized values like community contribution and self-growth (Davey & Lalande, 2004; Ros et al., 1999), which are similar to the attributes in the rapport category (personable, helpful, etc.) developed by Addison et al. (2015).

On the other hand, STEM students focused on values like status and prestige (Kasser & Ahuvia, 2002; Vansteenkiste et al. 2006); because these students have a greater concern for maintaining a positive public image (Chantara et al., 2014), they are more likely to value skills category attributes like being organized and knowledgeable that will help them be viewed positively. Additionally, humanities students scored higher on extraversion (e.g. more sociable, outgoing, etc.), and studies have shown extraverted individuals seek out opportunities that emphasize social interaction (Barrick et al., 2003). In other words, they would prefer professors that are approachable and fun, two attributes from the rapport category. Given personality differences between STEM and humanities majors, and the possibility that personality might relate to preferred characteristics in instructors, the present study explored whether students in

STEM and the humanities commented on different instructor characteristics in their teaching evaluations.

Using data from online teaching evaluations (RateMyProfessor.com), I predict that students in STEM courses will mention skills-related words more frequently and those in humanities courses will mention rapport-related words more often in their written comments. Additionally, I predict that professors who were described with a high-frequency of both skills and rapport words will have a higher overall rating.

Method

RMP Sampling

Rate My Professor (RMP) is an online database where students can rate their professors. In addition to offering a section for written comments, RMP allows students to provide ratings of the ease of the professor's course, their clarity, and their helpfulness (Freng & Webber, 2009). These ratings are completed using a 1 (low) to 5 (high) scale.

As RMP receives evaluations from over 6000 institutions worldwide, I developed the sample using several stages. First, I used data previously obtained by Freng and Webber (2009) which included evaluations from the largest public university in each U.S. state. Then, I chose three departments each from the STEM and humanities disciplines. The determination of what departments were considered STEM or humanities was based on guidelines from the federal government. The Congressional Research Service produced the "Science, Technology, Engineering, and Mathematics (STEM) Education: An Overview" in 2018; it provided several definitions used throughout the government for STEM, and three departments that were mentioned in these definitions were chemistry, physics, and mathematics. As such, I chose to include these three in the study. The Department of Defense Education Activity (DoDEA) defined the humanities as “the integrated study of history, literature, language, philosophy, the visual arts, theater, dance, and music according to culture, place, and time” (DoDEA, para 1). A random word selector was used to choose three departments from this sampling frame; the departments selected were: music, history, and philosophy.

Participants

RMP does not collect demographic data on students rating professors. The evaluations used in this study came from 34 public universities from Freng and Webber (2009); only 34 public universities were used because they were the only schools whose data had been properly cleaned in order to be run through LIWC. From this data, a total number of 35,513 student ratings and comments were collected – 20,740 from STEM courses and 14,773 from humanities courses. In addition, the number of ratings ranged from 511 (University of Montana) to 16,435 (Arizona State University).

Procedure

I analyzed comments from the 34 universities (Freng, 2021) that were written on teachers from the chemistry, physics, mathematics, music, history, and philosophy departments using the Linguistic Inquiry and Word Count (LIWC, 2001) software. This software, created by Pennebaker et al. (2007), uses 80 different linguistic categories (word count, personal pronouns, Present Focus, etc.) to analyze word use within a provided text (Robinson et al., 2013). The frequency with which people use words within these categories can relate to personality and real-world outcome measures (Robinson et al., 2013). For this study, I used LIWC to analyze two categories: skills-related words (clear, enthusiastic, organized, knowledgeable, engaging, etc.), and rapport-related words (personable, approachable/accessible, respectful, humorous/fun, helpful, etc.) (Addison et al., 2015). In order to prepare data for LIWC, it must be cleaned; this means that the evaluations were spell-checked and corrected.

I created dictionaries to measure skills and rapport. A dictionary is the collection of words that define a category (Tausczik & Pennebaker, 2010), so I produced these using the 10 words listed above. To supplement these words, I used Merriam-Webster's online thesaurus to find synonyms for each of the 10 words. The final skills dictionary consisted of the following words: clear, straightforward, transparent, enthusiastic, eager, excited, organized, neat, detailed, knowledgeable, educated, skilled, engaging, interesting, and entertaining. The rapport dictionary included the words: personable, cheerful, enjoyable, approachable, accessible, attainable, obtainable, respectful, polite, courteous, humorous, fun, funny, amusing, helpful, friendly, and accommodating. Once the skills and rapport dictionaries were finalized, LIWC then measured the frequency of relevant words in students' written comments. I also measured the overall quality rating of a professor by taking the average of RMP ratings of helpfulness and quality.

Results

I only selected students' written comments if they were 10 or more words in length to avoid analyzing comments that were less reflective (e.g., I love professor A; Horrible teacher). Then, I recoded STEM courses (chemistry, physics, and mathematics) into category 1; humanities courses (music, history, and philosophy) were coded into category 2.

Rapport and Skills

Because both t-tests violated the assumption of equal variances, I used equal variances not assumed values. Contrary to prediction, students evaluating instructors in STEM classes wrote more about rapport (M = 1.33, SD = 2.25) than students evaluating teachers in humanities courses (M = 1.17, SD = 2.17), t (32477.39) = 6.66, $p < .001$, d = .07 (see Table 1 for descriptive statistics and correlations). Contrary to prediction, students evaluating instructors in humanities classes wrote more about skills (M = 1.05, SD = 2.05) than students evaluating teachers in STEM courses (M = 0.49, SD = 1.42), t (24551.58) = -28.9, $p < .001$, d = .32

Overall Teaching Quality

Next, I used multiple regression to determine if rapport and skills comments predicted overall teaching quality for STEM courses and humanities courses. The overall multiple regression model predicting overall quality for STEM classes was significant, F (2, 20737) = 1087.69, $p < .001$, $R^2 = .095$. In the model, both rapport (β = .25, SE = .004, $p < .001$) and skills (β = .15, SE = .007, $p < .001$) predicted overall quality ratings. In other words, among students in STEM classes, mentioning rapport and skills positively predicted teaching quality. For humanities classes, the overall multiple regression model predicting overall quality was significant, F (2, 14770) = 655.99, $p < .001$, R^2= .082. In the model, both rapport (β = .22, SE = .005, $p < .001$) and skills (β = .17, SE = .005, $p < .001$) predicted overall quality ratings. Similarly to STEM courses, this means that among students in humanities classes, mentioning rapport and skills positively predicted teaching quality. Although students in STEM commented more on rapport and less on skills than students in humanities classes, commenting on both rapport and skills of the instructor/course predicted higher overall quality ratings of instructors.

Table 1. Means, Standard Deviations, and Zero-Order Correlations

Variable	***M***	***SD***	**1**	**2**
1. Overall Quality	3.63	1.42		
2. Total Rapport	1.32	2.31	.25*	
3. Total Skills	.86	1.86	.19*	.09*

*$p < .01$

Discussion

Past research has tried to establish what characteristics in a professor are important to students, and it has shown that students value traits like enthusiasm, organization, respect, and being accommodating in their professors (Addison et al., 2015). Previous studies, however, have not examined whether these values differ based on the type of class the professor is teaching, specifically a STEM or humanities course. I did not find the difference that I predicted, and this may be explained by the similarity between the words in the two dictionaries. For example, enthusiastic, engaging, and entertaining were in the skills dictionary and enjoyable and funny were in the rapport dictionary. According to Merriam Webster, these words are synonyms. This could have prevented me from actually measuring what I wanted to; specifically, the skills dictionary does not address what one typically thinks of as skills – hard skills like being able to apply formulas to real-world problems or create substances using chemistry. This means that the skills dictionary did not pick up on the type of skills that STEM students want to learn and instead measured teacher skills. Despite this, I found that no matter the type of class, students consistently value the same professor characteristics.

There are, however, some limitations to this study. First, the sample only came from large, public universities. This would make it difficult to generalize results to private or liberal arts colleges as these students and professors may differ from their peers at public institutions. A second limitation is that RMP only has evaluations from students motivated enough to complete them. It is still disputed whether there is a difference in ratings between students who are motivated to go to a site like RMP and those who complete conventional teaching evaluations (Freng & Webber, 2009). Furthermore, because RMP does not collect demographic data on students that fill out evaluations, it is unknown the level of representativeness of these students (Silva et al., 2008). This is especially limiting in this study because the major of the students leaving the evaluations is undetermined, and the studies that look at differences between these students are based on the major of the student and not just one class they are taking.

Moving forward, better dictionaries should be developed that could more accurately measure skills using words like innovation, creativity, and research. Even if this is pursued, it seems likely that differences between student values in STEM and humanities courses will not be significant. Students who mentioned skills and rapport-related words at higher rates, gave teachers higher overall quality ratings, so it seems that in general, regardless of what class they are in, students may value the same characteristics. More interesting avenues in this line of research would be to examine what motivates students to use services like RMP or to see if student evaluations actually influence professor behaviors and teaching methods. Research has shown that professors do pay attention to RMP evaluations (Boswell, 2016), but the influence on their actually teaching behaviors is not well established yet.

References

Addison, W. E., Stowell, J. R., & Reab, M. D. (2015). Attributes of introductory psychology and statistics teachers: Findings from comments on RateMyProfessors.com. *Scholarship of Teaching and Learning in Psychology, 1*(3), 229–234. https://doi.org/10.1037/stl0000034

Barnes, D., Engelland, B., Matherne, C., Martin, W., Orgeron, C., Ring, J., et al. (2008). Developing a psychometrically sound measure of collegiate teaching proficiency. *College Student Journal, 42*, 199-213.

Barrick, M. R., Mount, M. K., & Gupta, R. (2003). "Meta-Analysis of the Relationship between the Five-Factor Model of Personality and Holland's Occupational Types." *Personnel Psychology, 56* (1), 45–74.

Berring, L., Kumari, S., & Ahuja, S. (2017). Impact of personality traits and personal values on curriculum choice of young adults. *Journal of Beliefs & Values, 39*(3), 263–278. https://doi.org/10.1080/13617672.2017.1293930

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

Chantara, S., Ravinder, K., & Kaewkuekool, S. (2014). "Relationship between Lifestyle Values and Achievement Goal Orientation among Vocational Students in Thailand." *Journal of Beliefs & Values, 35* (1), 36–47.

Communications, D. D. E. A. (n.d.). *Humanities Standards*. DoDEA. https://www.dodea.edu/Curriculum/FineArts/Humanities/standards.cfm.

Darawong, C., & Sandmaung, M. (2019). Service quality enhancing student satisfaction in international programs of higher education institutions: a local student perspective.

Journal of Marketing for Higher Education, 29(2), 268–283. https://doi.org/10.1080/08841241.2019.1647483

Davey, H., & Lalande, V. (2004). "Gender Differences, Values, and Occupational Choice." http://www.natcon.org/archive/natcon/papers/natcon_papers_2004_Davey_Lalande.pdf.

Education Policy, & Granovskiy, B., Science, Technology, Engineering, and Mathematics (STEM) Education: an overview (2018). Washington D.C.; Congressional Research Service.

Freng, S., & Webber, D. (2009). Turning up the Heat on Online Teaching Evaluations: Does "Hotness" Matter? *Teaching of Psychology, 36*(3), 189–193. https://doi.org/10.1080/00986280902959739

Johnson, R. (2000). The authority of the student evaluation questionnaire. *Teaching in Higher Education*, *5*, 419–434. doi:10.1080/713699176

Kasser, T., & Ahuvia, A. (2002). "Materialistic Values and Wellbeing in Business Students." *European Journal of Social Psychology, 32*, 137–146.

Keeley, J., Furr, R. M., & Buskist, W. (2010). Differentiating psychology students' perceptions of teachers using the Teacher Behavior Checklist. *Teaching of Psychology, 37*, 16–20. doi:10.1080/00986280903426282

Keeley, J., Smith, D., & Buskist, W. (2006). The Teacher Behaviors Checklist: Factor analysis of its utility for evaluating teaching. *Teaching of Psychology, 33*, 84–90. doi:10.1207/s15328023top3302_1

Lammers, W. J., & Smith, S. M. (2008). Learning Factors in the University Classroom: Faculty and Student Perspectives. *Teaching of Psychology, 35*(2), 61–70. https://doi.org/10.1080/00986280802004586

Marsh, H. W. (1982). SEEQ: A reliable, valid and useful instrument for collecting students' evaluations of university teaching. *British Journal of Educational Psychology, 52*, 77–95. doi:10.1111/j.2044-8279.1982.tb02505.x

Marsh, H. W., Muthèn, B., Asparouhov, T., Lüdtke, O., Robitzsch, A., Morin, A. J. S., & Trautwein, U. (2009). Exploratory structural equation modeling, integrating CFA and EFA: Application to students' evaluations of university teaching. *Structural Equation Modeling, 16*, 439–476. doi:10.1080/10705510903008220

McGovern, T. V., & Miller, S. L. (2008). Integrating Teacher Behaviors with Character Strengths and Virtues for Faculty Development. *Teaching of Psychology, 35*(4), 278–285. https://doi.org/10.1080/00986280802374609

McKeachie, W. J. (1997). Student ratings: The validity of use. *American Psychologist*, *52*, 1218–1225.

Pennebaker, J. W., Booth, R. J., & Francis, M. E. (2007). Linguistic Inquiry and Word Count: LIWC [Computer software]. Austin, TX: LIWC.net.

Robinson, R. L., Navea, R., & Ickes, W. (2013). Predicting Final Course Performance From Students' Written Self-Introductions. *Journal of Language and Social Psychology, 32*(4), 469–479. https://doi.org/10.1177/0261927x13476869

Ros, M., Schwartz, S. H., & Surkiss. 1999. "Basic Individual Values, Work Values, and the Meaning of Work." *Applied Psychology, 48*(1), 49–71.

Silva, K. M., Silva, F. J., Quinn, M. A., Draper, J. N., Cover, K. R., & Munoff, A. A. (2008). Rate My Professor: Online Evaluations of Psychology Instructors. *Teaching of Psychology, 35*(2), 71–80. https://doi.org/10.1080/00986280801978434

Spooren, P., Brockx, B., & Mortelmans, D. (2013). On the Validity of Student Evaluation of Teaching. *Review of Educational Research, 83*(4), 598–642. https://doi.org/10.3102/0034654313496870

Tausczik, Y. R., & Pennebaker, J. W. (2010). The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. *Journal of Language and Social Psychology, 29*(1), 24–54. https://doi.org/10.1177/0261927x09351676

Toland, M., & De Ayala, R. J. (2005). A multilevel factor analysis of students' evalu-ations of teaching. *Educational and Psychological Measurement, 65*, 272–296. doi:10.1177/001316440426866

Vansteenkiste, M., Duriez, B., Simons, J., & Soenens, B. (2006). "Materialistic Values and Well-Being among Business Students: Further Evidence of Their Detrimental Effect." *Journal of Applied Social Psychology, 36*(12), 2892–2908.

Zabaleta, F. (2007). The use and misuse of student evaluation of teaching. *Teaching in Higher Education*, *12*, 55–76. doi:10.1080/13562510601102131