RESEARCH

How Faculty Recognized for Teaching Excellence Interpret and Respond to Student Ratings of Teaching

Jeff Cain, EdD, MS,a Cindy D. Stowe, PharmD,b Dina Ali, PharmD,c Frank Romanelli, PharmD, MPHa,d

a University of Kentucky College of Pharmacy, Lexington, Kentucky
b Sullivan University College of Pharmacy, Louisville, Kentucky
c University of Kentucky Healthcare, Lexington, Kentucky
d Executive Associate Editor, American Journal of Pharmaceutical Education, Arlington, Virginia

Submitted July 28, 2017; accepted November 22, 2017; published May 2019.

Objective. To determine how experienced pharmacy educators who have been recognized for teaching excellence interpret and respond to end-of-course student ratings of their teaching.

Methods. An expert sampling technique was used to identify pharmacy faculty members who had been recognized by the American Association of Colleges of Pharmacy (AACP) for teaching excellence as potential participants in the study. Sixteen of these faculty members were independently identified by two or more of the researchers as trusted candidates for the study, and 10 of these were randomly selected and invited to be interviewed via a web-conferencing platform. The interviews were transcribed and thematic analysis was used to identify overall themes.

Results. Nine pharmacy faculty members representing both practice-based (n=5) and non-practice (n=4) backgrounds participated in the interviews. The three primary themes that emerged from the interviews regarding approach to interpreting student evaluations of instruction were “Use to improve,” “Trends,” and “Value input.” The four primary themes that emerged regarding advice for new instructors were “Use to improve,” “Reflect,” “Do not take personally,” and “Themes.”

Conclusion. The faculty participants recognized for teaching excellence were consistent in their attitudes of valuing student feedback and using it as a tool for continuous quality improvement. While recognizing the limitations of student ratings of teaching, the participants used them as part of a reflective and holistic approach to teaching. These teaching experts provided valuable insight for new instructors, such as do not compromise instruction and do not take student ratings personally.

Keywords: faculty, teaching, teaching evaluations

INTRODUCTION

Student ratings of instruction have long been a primary gauge to assess teaching. Although numerous studies have supported their validity and usefulness,1-3 these types of evaluations have recently come under increased scrutiny4,5 with critics citing several shortcomings, including gender bias,6 lack of validity,7,8 and no correlation between ratings and actual learning.9 Additionally, some have suggested that student evaluations do not routinely lead to improvement in instructor teaching.10 Other authors and researchers have implied that student ratings of teaching are nothing more than customer satisfaction surveys.11

Despite the criticisms, these ratings may be the sole or primary means of feedback regarding students’ perceptions of their instructors’ teaching effectiveness. Among the many faculty anecdotes regarding these instruments is the challenge of interpreting and/or responding to non-constructive or harsh comments that sometimes accompany anonymous evaluations. In particular, overly critical remarks or personal attacks from students may result in teachers having decreased morale, diminished job satisfaction, and perceived psychological distance from students.12 When teaching efforts are met with heavy criticism, an instructor’s natural response may be feelings of incompetence or a desire to create a psychological barrier by devaluing the evaluation results. Furthermore, student evaluations may contain conflicting information, leaving faculty members perplexed by what the overall results actually mean. As indicated in a study by Wong and Moni of clinical instructors’ perceptions of student
evaluations of teaching, educators have mixed responses regarding the potential usefulness of student evaluations in promoting and advancing teaching. Because student ratings of teaching are susceptible to misinterpretation, those new to an instructional role should receive guidance on how to mentally process teaching evaluations and how to use them to improve their teaching skills. While the topic of student ratings of instruction has been widely studied, and while there is an abundance of literature pertaining to the administration and validity of student ratings of teaching, and general best practices for using them, there is very little literature on how higher education faculty members, particularly those from the clinical sciences, actually read, interpret, and respond to those evaluations. This study adds to the literature by discussing the actual approach and mindsets that “excellent” teachers use with their own evaluations. Given the renewed questioning of the validity and usefulness of these ratings, this information is important as it demonstrates current and authentic examples from pharmacy instructors. The objective of this study was to understand how experienced faculty members who had been recognized for excellence in teaching interpret and mentally respond to their own evaluations.

METHODS
This study used an expert sampling technique to identify experienced educators recognized for teaching excellence who could provide insight into how they interpret and respond to teaching evaluations. Expert sampling, which is a type of purposive/judgment sampling, involves the assembly of persons with known expertise and is used when researchers are searching for specific knowledge versus generalizable information. Because the purpose of the study was to gain insight from skilled teachers, the researchers used existing American Association of Colleges of Pharmacy (AACP) databases of teaching awards to draw the sample. The initial list of 354 potential participants was formed by combining the AACP Robert K. Chalmers Distinguished Pharmacy Educator Award recipients, AACP Distinguished Teaching Scholar Award recipients, and the 2015 and 2016 Teachers of the Year recipients as reported to AACP by member institutions. From that list, three researchers individually and independently identified faculty members who they believed would provide valuable feedback. Those individual faculty members selected by multiple researchers formed the final list of 12 potential participants from which the sample was selected. The list was sorted into practice-based and non-practice-based faculty members to ensure that perspectives across disciplines were represented. Non-practice-based included faculty members from basic sciences and social and administrative sciences who did not teach clinical courses. After randomization, the first five practice-based and first five non-practice-based faculty members were contacted via email (obtained via respective institution websites) and invited to participate in the study.

A standard interview template that was developed by the investigators contained demographic questions, broad primary questions pertaining to study objectives, and optional, detailed follow-up questions if specific points were not addressed in response to the primary questions. Protocol design was based on recommendations of Turner and Jacob and Furgeron for qualitative interview research. Interview questions for the protocol itself were informed from a combination of primary literature and investigator experience with issues surrounding student ratings of instruction. In particular, research and commentary by Lindahl, Beran, and Ryan formed the basis for the specific questions to be asked. Broad primary questions focused on the overall process that instructors use with their own evaluations, including the logistics, internal mental processing, and effects on teaching. An example primary question was “What is your process for reading and interpreting instructor evaluations?” Corresponding optional follow-up questions included “How do you internally deal with negative evaluations?” and “Are there particular things you look for in evaluations?” Another primary question addressing the study objectives was “What advice would you give to new instructors regarding how they read and interpret their own instructor evaluations?” The interview questions were pilot tested with a faculty member who confirmed that the meaning and intent of each question were clear. Based on the feedback, minor revisions were made to form the final instrument (available upon request).

A mutually agreeable time for an online interview was established and one member of the research team conducted and recorded all interviews using Zoom videoconferencing software (Zoom, San Jose, CA). Interviews occurred from November 2016 through January 2017. Per common protocol for this type of research, the number of interviews continued until data saturation (i.e., the point where posing the same questions to additional subjects resulted in no new information being obtained) was reached. A second member of the research team transcribed all interviews using oTranscribe (oTranscribe, London, England), a web-based transcription tool. The transcriptions were then made available to all four investigators for analysis. The University of Kentucky Institutional Review Board approved this study and each interviewee provided verbal consent to participate in the study.
Thematic analysis was used to identify consistent themes from the interviews. Thematic analysis is a rigorous qualitative research method used for identifying, analyzing, and reporting patterns and themes within data.26 The analysis generally followed the basic protocol precepts recommended by Braun and Clarke, which involved familiarization with data, generation of initial codes, and identification, review, and refinement of themes.26 Each of the four researchers independently read all transcripts and identified initial codes for each interview. In this study, codes refer to a basic idea or concept that can be assessed in a meaningful way.27 An example of a code is “perspective,” which in this study refers to the basic concept that students do not always have a complete perspective of what they need to know. The codes were identified for two different portions of the interview: “overall approach to interpreting and responding to student ratings of teaching” and “advice to young faculty regarding how they should interpret and respond to student ratings of teaching.” As a group and through an iterative process, the researchers discussed the independently derived codes to categorize, develop, revise, and arrive at a consensus for the set of themes to use in the final analysis of each portion of the interviews. Whereas the codes were narrow and specific, themes were broader and usually formed from the combination of multiple codes. Each theme was defined with a phrase or sentence that described the essence of the theme, but for ease of recording and analysis, each theme was named with a concise description. Each researcher then independently reread the transcripts and assigned one or more of the agreed upon set of themes to each interview. Finally, as a group, the researchers discussed each interview and collaboratively reconciled any differences in individual theme assignment to determine the final themes assigned to each interview.

RESULTS

Nine faculty experts participated in the structured interviews. These experts had a cumulative total of 232 years of experience in academic pharmacy education and represented a balance between practice (n=5) and non-practice based (n=4) faculty members (Table 1). Most of the faculty experts (n=6) had never received formal mentoring regarding the interpretation of student ratings of instruction.

Ten themes emerged from the analysis of “Approach to interpretation of and response to student ratings of instruction” (Table 2). One theme, “use to improve,” was present in all interviews. Participants indicated that using the feedback to become better was a priority for them. The main objective was to use evaluation feedback to make the course better and more meaningful for students.

“Value input,” “trends,” “themes,” and “limitations” were also consistent and occurred in the majority (n=6-8) of the transcripts. Value input from students was indicated in a variety of ways among the participants, but the essence of this theme was that we all learn through criticism and failure. “Themes” and “trends” were two other primary themes that emerged from the interviews and related to the practice of interpreting the results of student ratings. Themes referred to searching for common items or issues cited by multiple students within a course. If evaluations produced high numbers of comments pertaining to a particular aspect of teaching such as talking too fast or too much content, then those are areas in which the instructor should investigate thoroughly and potentially act upon. Similar to themes, trends involve searching for items or

---

Table 1. Background, Academia Experience, and Mentoring of Study Participants

<table>
<thead>
<tr>
<th>Faculty Expert</th>
<th>Practice-based Faculty Member?</th>
<th>Years in Higher Education, Total</th>
<th>Years in Higher Education, Public</th>
<th>Years in Higher Education, Private</th>
<th>Received Mentoring</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Yes</td>
<td>18</td>
<td>18</td>
<td>0</td>
<td>No</td>
</tr>
<tr>
<td>2</td>
<td>No</td>
<td>19</td>
<td>19</td>
<td>0</td>
<td>No</td>
</tr>
<tr>
<td>3</td>
<td>Yes</td>
<td>17</td>
<td>0</td>
<td>17</td>
<td>Informal</td>
</tr>
<tr>
<td>4</td>
<td>Yes</td>
<td>27</td>
<td>6</td>
<td>21</td>
<td>Informal</td>
</tr>
<tr>
<td>5</td>
<td>Yes</td>
<td>43</td>
<td>39</td>
<td>4</td>
<td>No</td>
</tr>
<tr>
<td>6</td>
<td>No</td>
<td>22</td>
<td>0</td>
<td>22</td>
<td>No</td>
</tr>
<tr>
<td>7</td>
<td>No</td>
<td>43</td>
<td>43</td>
<td>0</td>
<td>Formal</td>
</tr>
<tr>
<td>8</td>
<td>Yes</td>
<td>25</td>
<td>25</td>
<td>0</td>
<td>Formal</td>
</tr>
<tr>
<td>9</td>
<td>No</td>
<td>18</td>
<td>18</td>
<td>0</td>
<td>Formal</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>232</td>
<td>162</td>
<td>68</td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td>25.8 (SD=9.7)</td>
<td>18.7 (SD=14.5)</td>
<td>7.1 (SD=9.3)</td>
<td></td>
</tr>
</tbody>
</table>
issues that consistently appear on student ratings semester after semester. Identifying and examining those common issues allow faculty to gauge the success of course and instructional alterations.

Understanding the limitations was the fourth common theme across all interviews. Participants acknowledged the criticisms and limitations of student ratings without dismissing their value. There was a general understanding a response bias may exist with responses only from those who really love or hate the instructor. The participants recognized that those responses are still important and worth reading because they reflect what students really feel about the course, even if it does not reflect everyone’s opinion. One participant noted that there may still be elements of truth within those evaluations regardless if they are positive or negative. Within this theme, participants also recognized the need to address limitations through additional methods. Seeking peer feedback is one way to overcome the shortfalls of student perceptions, which can sometimes be more like a popularity contest. One participant noted that student ratings are beneficial, but not for evaluating every aspect of teaching. Student feedback is a data point regarding the quality of teaching, but the message when mentoring others is that it is only one source of information. There are good uses of student feedback, such as their experience within the class, but there are other aspects of teaching (eg, rigor and content selection) that students may not be able to evaluate accurately.

The other themes (“scholarly process,” “feedback process,” “educational rationale,” “do not compromise,” and “big picture”) occurred in three or four of the transcripts. Scholarly process referred to both the attitude and methods of taking a scholarly perspective when interpreting student ratings. In other words, examining the results with a curiosity for learning, similar to analysis of scientific studies. One participant took it as a personal challenge to study the results to determine if the course was going well, if the teaching was sufficient, and if there were other personal characteristics that affect the student experience.

The importance of the “feedback process” was revealed by several participants who approach the student rating process from an expanded viewpoint, starting with informing students how to provide constructive and useful student ratings. One participant explained this by stating that emphasizing to students that evaluations are important, that instructors look at them, and want to hear from students will evoke a better response. Even harsh critiques are very constructive because the students may become more reflective with what occurs in the classroom and more responsible with their role in shaping the learning process.

“Educational rationale” was another proactive approach to student ratings and referred to explaining the reasons why certain teaching methods are used and why particular skills and attitudes are important. One participant explained this well by stating that students do not always understand the full landscape of a course. Providing context about the course, the content covered, and how the instruction is designed may help students better understand the bigger picture. In this participant’s experience it helped increase the number of specific, helpful comments while reducing the number of non-constructive comments.

Participants also commented on how sometimes good teaching practice may not be well-received by
students in the moment, but that it is important to not compromise those practices for the sake of higher ratings. As one participant noted, there may be certain aspects of the way someone teaches that students will not like. Examples of this include making students participate in class discussion or engage with the material outside of class. Even though those strategies consistently received negative evaluations, these were important elements of the instructional design that should not be altered.

“Look at the big picture” was the final theme and encompassed an attitude of not concentrating solely on a few comments. This quote from one of the participants was characteristic of others who expressed a similar approach. “I have been at it for 22 years and I know the first semester when I got my evaluations, I was in [another instructor’s] office asking what the heck went wrong here? It’s funny, because we all do that. We concentrate on the negative, you know. So, and that’s what you sadly pay attention to... like, oh my god! How can they say that? And it might be one student, or two, and you completely neglect all of the positive things. We had a talk then. I remember very vividly, and she did say to me to take it easy. Try to see the big picture rather than just concentrating on a few comments and so on.”

Nine themes emerged from the analysis of “Advice for new instructors” (Table 3). No unanimous themes occurred across all interviews, but two themes emerged in seven or more transcripts: “use to improve” and “reflect.” Once again, “use to improve” was cited as an optimal approach to interpreting student ratings. The other theme found in the majority of interviews was “reflect,” which participants described as a specific approach to take when examining one’s respective ratings. One participant stated that by self-evaluating and self-reflecting on what you do in the classroom that you begin to look at evaluations from a different lens. You begin to see them from a perspective of trying to do the right thing for students.

Themes identified in four to six transcripts were “trends,” “themes,” “do not take personally,” and “big picture.” Participants not only look for trends and themes themselves but recommend those strategies to others. One of the other themes emerging as a recommendation is the advice to not take student ratings personally. Learning to deal with criticism, developing a “thick skin,” and realizing that this is a student’s perception at that point in time were common sentiments expressed by the participants. The “big picture” theme reflected the idea that multiple issues and factors can affect student ratings and specific comments. Participants indicated that it is important to consider everything that is happening in the teaching and learning environment that may prompt certain ratings and comments.

The three themes least commonly identified were “use a mentor,” “transparency,” and “do not compromise.” Using a mentor was cited as an important way for new instructors to learn how to approach and interpret student ratings. Transparency referred to the process of establishing up front with the students what and how a course will be taught, but also, why. Finally, “do not compromise” was strong advice meant to ensure that appropriate rigor is maintained in teaching and that instructors should not sacrifice quality for fear of poor ratings. “There are a lot of things we do for students because it’s the right thing to do for students in the way you teach them and it doesn’t mean they have to like it. That was a hard thing for me to realize as I was getting more experience as an instructor. I do know what’s best for them in some ways, and they may not think...”

Table 3. Prevalence of Final Themes for Advice to New Instructors by Study Participants

<table>
<thead>
<tr>
<th>Theme Name</th>
<th>Description of Theme</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use to Improve</td>
<td>Use student ratings of instruction to improve</td>
</tr>
<tr>
<td>Reflect</td>
<td>Self-reflect on the ratings and how you teach</td>
</tr>
<tr>
<td>Do Not Take Personally</td>
<td>Move beyond taking ratings and comments personally</td>
</tr>
<tr>
<td>Themes</td>
<td>Look for themes within the ratings</td>
</tr>
<tr>
<td>Big Picture</td>
<td>Recognize the different reasons for why students rate the way they do</td>
</tr>
<tr>
<td>Trends</td>
<td>Look for trends over time with the ratings</td>
</tr>
<tr>
<td>Transparency</td>
<td>Be transparent with students on how and why you teach the way you do</td>
</tr>
<tr>
<td>Use a Mentor</td>
<td>Seek advice on how to interpret and respond to student ratings</td>
</tr>
<tr>
<td>Do Not Compromise</td>
<td>Do not compromise teaching in order to achieve higher student ratings of instruction</td>
</tr>
</tbody>
</table>

*Based on investigator consensus of nine expert faculty interviews
it’s what’s best for them, but that doesn’t mean it’s not what’s best for them. And I am going to hear about it on these evaluations and that is okay. You have to have enough confidence to know that what you are doing is the right thing, regardless of what their comments are.”

DISCUSSION

The topic of student ratings of instruction in higher education is nuanced, complex, and controversial, and can be discussed from a multitude of perspectives. This investigation examined an area related to instructor evaluations that has not been well studied, that is, how faculty members who have been recognized as outstanding educators by various groups, specifically use quantitative and qualitative student assessments to inform their teaching. More precisely, we endeavored to elicit their philosophical approaches to student ratings of instruction. The majority of instructors included in this study had received at least one national teaching recognition. By understanding the approaches used by these accomplished educators and by describing any patterns involved in their use of quantitative and qualitative data, we hope to inform the academy and further contribute to best practices in the area of teaching evaluation.

As previously mentioned, the preponderance of published literature related to the utility, impact, and validity of instructor evaluations has focused on student-associated factors. Investigators have long attempted to determine correlates that potentially demonstrate the likelihood that students will or will not evaluate a course or instructor in a favorable manner. Additionally, investigators have often sought to describe patterns related to student ratings and actual learning outcomes. In all instances, the data is replete with controversy and mixed results. The only consistency is that academic institutions of all kinds use and will likely continue to use student ratings of instruction as at least one indicator of overall teaching effectiveness.

Participants in this study were very consistent in their approaches to student ratings of instruction. While student ratings are commonly criticized for their potential shortcomings and limitations, respondents here were unanimous in stating that they valued and used evaluation data to improve their teaching. Regardless of the feedback received, it is incumbent on faculty members as educators to take it and learn from it. Even though the teaching literature states this as best practice, there are still many faculty members who dismiss the ratings as invalid, and by extension, they do not consider them useful. While some administrators may use student ratings as a summative assessment of teaching effectiveness, participants in this study appear to view them more as formative assessments. This approach of using student ratings to improve their teaching concurs with advice from other teaching literature regarding best uses of evaluations by students. Respondents also commonly stated that they valued the input garnered from student ratings. Almost every study participant reported that when reviewing evaluations, they tend to focus on themes and trends that emerge from qualitative data. This can be translated to include both themes that emerge in a given discrete evaluation period as well as trends across multiple time periods and courses. The concept of approaching student ratings as a single data point that should not be interpreted simplistically but rather holistically is not a new concept. Many authors have advised this approach, and these noted educators continue to echo that concept. The presence of various confounding factors may influence the applicability of data garnered from instructor evaluations; therefore, any single data set should be considered with an overabundance of caution. The use of multiple data sets that may or may not reveal themes and trends offers a higher level of validity by removing potential confounding issues. Themes and trends that repeatedly emerge over time may describe a myriad of either favorable and/or unfavorable aspects of a given instructor or course and may include such things as timeliness, professionalism, quality of audiovisuals, etc.

In agreement with recent literature regarding student ratings, respondents did acknowledge that the instructor evaluation process does have limitations that should be considered. A common limitation that has been cited in the literature includes the potential for students to assume a “customer mindset” or entitlement paradigm wherein there is an expectation for the acquisition of credit based solely on effort rather than actual achievement. This paradigm can create a no-win situation for the instructor and must be taken into consideration when interpreting results. Other limitations may involve semantics associated with the surveying procedure. Factors such as survey timing, instrument length, and survey fatigue can all potentially effect the disposition of an evaluation. Many colleges and schools of pharmacy are currently undergoing curricular reform, which may substantially alter existing curricula in terms of content and/or delivery. Implementation of such dramatic and disruptive change may influence student temperament as it relates to evaluation. Thus, it will be important for experienced instructors to play a role in helping others use and interpret student evaluations in a well-reasoned manner.

Several respondents commented on the need to approach student ratings in a scholarly fashion, which may include clearly articulating teaching philosophy and classroom semantics to students. As expressed by the
respondents, students are more likely to respond favorably to course assignments, delivery methods, and other logistics if they have a clear understanding of expectations and intended outcomes for such processes. A scholarly approach also involves considering the classroom as a venue for positive experimentation. Whenever possible, modifications to teaching and other interventions should be driven by sound assessments and backed by an evidence-based foundation.

A final theme emerging from the interviews was the concept of “not compromising.” While all of the participants noted that they used student ratings of instruction to improve their teaching, several also commented that it was important to not concede indiscriminately with all student feedback. In this sense there appear to be at least some core principles that are uncompromising for these decorated educators. This underscores the fact that all faculty members (even those that are well accomplished) will receive some feedback that they cannot or will not address. This feedback may involve suggestions that would compromise their personal teaching philosophy or that are simply inconsistent with their expert opinion. This is one area in which the approach of participants in this study differed from that found in previous research, which showed that one effect of student ratings is the pressure faculty members receive to become more lenient or that are simply inconsistent with their expert opinion. As might be expected from award-winning instructors, their commitment to high educational outcomes overrides the desire for higher student ratings. Ultimately, the faculty member in all cases is the judge of what material must be instructed, by what manner it will be instructed, and the rigor with which it will be assessed. Often, students at a given point in their educational career may be unable to appreciate course content and how it prospectively relates to their overall training and career. In many cases, negative ratings may not be an act of malice, but rather an indicator of the students’ novice status.

When study participants were asked what advice they would extend towards junior faculty members with regard to student ratings of instruction, many similarities were seen in terms of themes associated with their own personal approach to evaluations. Approaching evaluations as a means to improve teaching was again a common theme as was self-reflection. Self-reflection is an important component of the improvement proposition particularly for junior faculty members. Results from this study concur with previous research showing that reflective practitioners receive increased student ratings, and that those demonstrating higher levels of reflection improve their evaluations the most. Reflection allows the faculty member to introspectively contemplate what went well and what did not go well through the progression of a given learning experience and/or semester. Purposeful thought can then be given to the reasons different approaches or interventions were or were not perceived to be successful and in turn, how modifications can be implemented. Reflection that is documented also allows the faculty member to demonstrate a tangible commitment to self-improvement.

Respondents advised junior faculty members not to approach student ratings personally and perhaps, that may be the most important advice given the effects that student ratings may have on newer instructors. These more seasoned faculty members had probably encountered evaluations that elicited both positive and negative visceral responses. Student ratings have the potential to include comments that could be judged as being “cruel,” and multiple studies have indicated that student ratings decrease faculty morale. Even well-intended constructive feedback from students can sometimes be construed as malevolent by less-experienced faculty members. Participants were overall less likely to advise junior faculty members to “not compromise” as compared to their own approaches to the classroom. This was perhaps due to the common wisdom accumulated as senior faculty members, as they may be at a point in their careers at which they have already made modifications that led them to their current philosophy and successful approaches.

Some respondents reiterated the need for junior faculty members to discuss evaluation results with mentors. Senior faculty mentors may be able to assist in interpreting the data and advise on best practices. Lastly, respondents advised junior faculty members to be transparent with students in terms of their intended outcomes related to the material being taught. This concept is similar to the aforementioned approach regarding educational transparency. Students should understand what is being done in the classroom and for what reasons. Intended outcomes should be clear so that students can articulate an understanding of the instructor’s educational approach and philosophy.

In the course of data collection, none of the study participants alluded to or commented on the proposition of eliminating student ratings of instruction. Perhaps the reality of these instruments as at least one component of teaching evaluation is clear to this group. Several authors have commented on the obligatory use of student ratings and the unlikely forecasting of their elimination. Given that the tradition of student ratings is unlikely to be dismissed, it is imperative that faculty members maximize use of any available data. Faculty members and administrators alike should also strive to improve existing evaluation practices in order to garner the most robust
data possible. Relying on the collective wisdom of recognized faculty members may prove to be an essential part of this ongoing process.

The following limitations should be considered when interpreting study results. This study focused only on how faculty members recognized for teaching excellence interpret and respond to student ratings of instructors. Excellence in teaching was determined through the use of available AACP teaching award databases. The awards themselves have their own unique criteria, therefore faculty members recognized in these awards may have met different criteria for excellent teaching. In addition, the results may or may not differ from those that would have been obtained from faculty members who were not perceived by students or peers as excellent instructors. While some may interpret the relatively small number (n = 9) of faculty members interviewed for this study as a limitation, the sampling protocol used was appropriate for qualitative research that was intended to reveal, not generalize.44 Also, as protocol dictated, the number of interviews continued until no new information was being gained from additional subjects.25

CONCLUSION
Because student ratings of instruction continue to be used at least in part to assess teaching, faculty members must understand how to read, interpret, and respond to both the quantitative and qualitative results they receive. This is particularly important for graduate students, residents, and junior faculty members who are new to the role of instructor. While each individual’s approach to reading, interpreting, and responding to student ratings of teaching may have differed slightly, they all believed that the ratings play an important role despite their limitations. In particular, instructors should continually use student ratings to reflect upon trends and themes for the purpose of improving instruction. Just as important, but perhaps more difficult to accomplish, instructors should not sacrifice academic rigor or compromise quality in order to achieve better student ratings. Finally, student ratings should be viewed from a continuous quality improvement perspective and not as an evaluation of an instructor’s personal or academic worth. New instructors are encouraged to find a faculty mentor(s) who can help them view student ratings objectively, improve teaching accordingly, and maintain both a personal and professional perspective when interpreting and responding to student ratings.

ACKNOWLEDGMENTS
The authors thank the study participants who took time to discuss their thoughts, attitudes, and approaches toward student evaluations of teaching.

REFERENCES