Assessment of Teaching Effectiveness in U.S. Dental Schools and the Value of Triangulation

Leila Jahangiri, D.M.D., M.M.Sc.; Thomas W. Mucciolo, B.B.A.; Mijin Choi, D.D.S., M.S.; Andrew I. Spielman, D.M.D., Ph.D.

Abstract: The routine evaluation of teaching effectiveness is important in improving faculty, departmental, and institutional efforts. There are three main categories of assessments: those performed by students, peers, and self. Although each category is independently valid, a collection of data from all three categories leads to a more comprehensive outcome and a creation of a triangulation model. The purpose of this study was to identify commonly used methods of assessing teaching effectiveness and to suggest the use of a triangulation model, which has been advocated in the literature on performance assessment as an optimal approach for evaluating teaching effectiveness. A twelve-question survey was sent to all U.S. dental schools to identify evaluation methods as well as to find evidence of triangulation. Thirty-nine out of fifty-seven schools responded. The majority of the schools used student evaluations (81 percent) and peer reviews (78 percent). A minority of schools reported using self-evaluations (31 percent). Less than one in five dental schools reported using all three strategies to achieve triangulation (19 percent). The three most commonly used evaluation methods ("performed routinely") were all in the student evaluation category. Less than half of the schools routinely evaluated clinical teaching effectiveness by any means (42 percent). In conclusion, dental schools should implement a triangulation process, in which evaluation data are obtained from students, peers, and self to provide a comprehensive and composite assessment of teaching effectiveness.

Dr. Jahangiri is Assistant Professor and Chair, Department of Prosthodontics, New York University College of Dentistry; Mr. Mucciolo is Adjunct Assistant Professor, Department of Prosthodontics, New York University College of Dentistry; Dr. Choi is Clinical Assistant Professor, Department of Prosthodontics, New York University College of Dentistry; and Dr. Spielman is Professor of Basic Sciences and Associate Dean for Academic Affairs, New York University College of Dentistry. Direct correspondence and requests for reprints to Dr. Leila Jahangiri, Chair, Department of Prosthodontics, New York University College of Dentistry, 423 East 23rd Street, 16 North, New York, NY 10010-4086; 212-998-9279 phone; 212-995-4686 fax; lj14@nyu.edu.

Key words: teaching effectiveness, teaching assessment, triangulation, teacher evaluation

Submitted for publication 11/30/07; accepted 3/14/08

Teaching effectiveness can be defined as the extent to which the teaching activity fulfills its intended purpose, function, and goal. Although goals can vary across institutions and may be specific or global, evaluating a teacher's effectiveness is proof of the academic institution's commitment to continual improvement. This study reviews survey findings to determine the most commonly used methods for assessing teaching effectiveness in U.S. dental schools. An evaluation process is a review of faculty performance; when appropriately implemented, the process results in objective decisions regarding promotions and recommendations.[1] The outcomes from assessments of teaching effectiveness can be used for formative and summative purposes. Formative assessments are used to develop or improve persons or programs. Summative assessments are used to determine "final" achievement and to provide accountability, as the results are more directly related to the efficacy of teaching.[1] Summative evaluations focus on the outcome of teaching rather than the process.[1,2] A comprehensive method of assessing teaching effectiveness should include various types of evaluation strategies,[1,3] specifically those of the formative nature.

A wide-ranging collection of data or evidence from multiple and diverse sources is described in the education literature as triangulation.[1,4-12] This composite assessment methodology is thought to be more valuable because it compensates for the inadequacies that may be related to any one assessment method.[1,9,13-16] According to Berk,[1] using a triangulation methodology provides "a more accurate, reliable, and comprehensive method of teaching effectiveness than just one source." Many other experts on performance assessment in the academic

arena recognize that the measurement of teaching effectiveness is a complex process, and they strongly endorse using multiple sources of evidence,[6,7,12,16-23] in an attempt to balance out some of the shortcomings of each corner of the "triangle." The three most frequently used data sources that comprise the triangle are student ratings, peer observations, and self-assessments. In summary, teaching effectiveness can be assessed using diverse evaluation strategies, the combination of which is referred to as triangulation.[1,4,6,7,11,12] For a triangulation method to be most effective, data from different sources should ideally overlap to some extent.[4,6] For example, similar questions could be asked of students, peers, and self. The extent to which triangulation models are used to evaluate teaching effectiveness in dental education has not been determined. The purpose of this study was to identify commonly used methods of assessing teaching effectiveness in U.S. dental schools and to propose a model, based on the concept of triangulation, which includes evaluations from students, peers, and self.

Methods of Measuring Teaching Effectiveness

Student Assessment

Teaching invariably involves interaction with students; therefore, a huge emphasis has been given to student evaluations, ratings, feedback, and performance.[24-42] The validity of student evaluations has been a subject of much debate (Table 1). A meta-analysis of forty-one research studies indicated that student evaluations/ratings should be an essential component of evaluating teaching effectiveness.[43] A large number of students can respond to a survey on teaching effectiveness, making the results reliable.[15,16,27,44-47] Moreover, the respondents are those who have observed the teachers on a number of occasions[9,48] and have been personally affected by the teaching,[48] giving this type of evaluation a high face validity.[25,33,34,44,46,49]

Studies also support the use of retrospective "student" evaluations, such as exit surveys and alumni ratings, as additional methods of assessing past teaching effectiveness.[1] Presumably, there is less of a correlation with exams or assignments and more of a relationship to the value of the education, the individual's future goals, and the real world. However, these ratings are not necessarily any more reliable than current student evaluations.[25,50-52]

Some have questioned whether students have the wisdom and the experience to make qualified observations and evaluations about teaching quality.[25,53-56] In addition, it has been suggested that students' evaluations may be overly influenced by their perceptions of faculty members' leniency or severity in assigning grades (e.g., lenient, "easy" graders receive higher ratings than strict graders), which may undermine the value of students' course and teacher evaluations and induce skepticism among the faculty in the value of these ratings.[1,7,15,25,35,38,46,57-60] In fact, there is conflicting evidence and opinion in the faculty assessment literature concerning the effect that students' expectations for high grades have on their evaluation of teachers.[61-63] Some studies totally support the leniency hypothesis,[64-69] while others found no evidence that leniency affects ratings.[70,71] There appears to be little or no support in the literature that evaluations are affected by course content-related elements such as the course workload or difficulty[15,35,72] or even the student's grade point average.[15,73] Yet, there is some evidence that students' assessment of teaching and teachers may be somewhat influenced by students' prior interest in the content or subject.[15,74,75]

A number of studies find little or no evidence that gender differences (instructor or student) play a role in determining evaluation outcomes.[15,25,46,76-78] Some studies note that the characteristics of the instructor, such as personality[25,79] and popularity,[25] may have an effect on the evaluations, although the "popularity" myth has been debunked, for the most part.[15,80,81] There are even reports that ratings are influenced by the presence of the instructor in the room while evaluations are being completed.[82,83] Some studies recommend that a third party administer and collect course/teacher evaluation forms to protect student anonymity and reduce the "psychological presence" of the instructor upon students' decision making about ratings.[16,20]

Evidence suggests that past or present students should not be asked to comment on the adequacy or timeliness of a course or curriculum. Nor should they be required to judge the appropriateness of related assignments or exams.[1] Faculty (peers) are considered more suitable for such evaluations.[15] While evaluations are completed more accurately when students are given adequate time and instruction,[84] the content

Table 1. Literature findings of factors influencing student evaluations

Student-Related Issue	Agree	Disagree	Open to Debate
Student bias (class size, grading by faculty)	Ory and Ryan,[58] Chacko,[64] Koshland,[65] Nimmer and Stone,[66] Powell,[67] Snyder and Clair,[68] Vasta and Sarmiento[69]	McKeachie,[34] Greenwald and Gillmore,[57] Abrami et al.,[61] Marsh,[70] Theall and Franklin[71]	Marsh,[15] Aleamoni and Hexner[46]
Student limited self-insight in measuring teaching effectiveness (immaturity, lack of experience, and unpredictability)		Albanese,[53] Hativa and Raviv,[54] Palchik et al.,[55] Ross and Bruce[56]	
Content of rating form drives what is taught	Ory and Ryan[58]		
Instructor presence in room	Feldman,[82] Pulich[83]		
Instructor's personality			Aleamoni,[25] Erdle et al.[79]
Instructor's popularity		Marsh,[15] Feldman,[80] Tang[81]	Aleamoni[25]
Gender of instructor and/or student		Aleamoni and Hexner,[46] Amin,[76] Dukes and Victoria[77]	Marsh,[15] Aleamoni,[25] Feldman[78]
Student prior interest in content	Marsh and Cooper,[74] Ory[75]	Marsh[15]	
Student expectations of high grades	Abrami et al.[62]	Centra[20]	
Student GPA		Marsh,[15] Theall and Franklin[73]	
Course workload and/or difficulty		Theall et al.[72]	Marsh[15]
Student learning and achievement related to ratings	Cohen,[43] Aleamoni and Hexner,[46] Centra,[86] McKeachie[87]	Doyle and Crichton[47]	Marsh[85]
Adequate time, notice, and instruction given to students	Svinicki[84]		
Number of student evaluators; experience of evaluators with teachers (many hours observing)			Hoyt and Pallett[48]
High face validity (personally affected by teacher)	Hoyt and Pallett[48]		
Students have multiple opportunities to experience day-to-day teaching	Berk et al.[9]		
Reliability of ratings	Marsh,[15] Braskamp and Ory,[16] Centra,[20] Cohen,[43] Aleomoni,[44] Cashin,[45] Doyle and Crichton,[47] Marsh,[70] Centra,[86] McKeachie,[87] Seldin,[91] Seldin,[92] Ramsden,[93] Marsh and Dunkin,[94] Koon and Murray[95]		Hoyt and Pallett[48]
Student feedback improves teaching effectiveness		Hoyt and Pallett,[48] Cohen,[88] Brinko[89]	Ory and Ryan[58]
Alumni ratings more reliable than current student ratings		Aleamoni,[25] Marsh,[50] McKeachie et al.,[51] Overall and Marsh[52]	

of the rating form itself may drive educators to modify what is taught.[58] Whether the extent of learning, as indicated by test performance, is directly related to ratings[43,46,85-87] and whether student feedback actually improves teaching[48,52,58,88-90] are both debated in the literature. In spite of the controversy, the current belief is that student ratings and evaluations are highly valuable as determinants of teaching effectiveness and that the ratings provide a feedback mechanism for continuous improvement.[24,35] It appears that the majority of researchers and experts in academic performance assessment believe in the reliability and validity of student evaluations as a means of evaluating teaching.[15,43,70,86,87,91-96]

Peer Assessment

Peer review by faculty colleagues is considered another valuable assessment method for measuring teaching effectiveness[5,9,47,60,97,98] (Table 2). Peer review can be either formative or summative.[2] Formative review is used to enhance good teaching practices, develop faculty, and share resources. Peer evaluation has a high validity for judging course goals and objectives, course content, and appropriateness of tests or assignments, something that students are not able to effectively assess.[15] Formative peer reviews are viewed more favorably than summative reviews by both faculty who participate in colleague assessment programs and by experts in performance assessment.[1,13,20,99,100] Consequently, most assessment authorities recommend that peer evaluation should be formative—namely, a mentoring process that

Table 2. Literature findings of factors influencing peer evaluations

Peer-Related Issue	Agree	Disagree	Open to Debate
Friendship bias	Berk et al.[9]	Cederblom and Lounsbury,[98] Cosh[103]	Love,[60] Cederblom and Lounsbury[98]
Ratings are very generous, skewed to the upper scale	Berk et al.,[9] Root,[101] Braskamp et al.[102]		Doyle and Crichton[47]
Rating scales do not measure effective teaching characteristics	Berk et al.[9]		Braskamp and Ory[16]
Informal collegial relationship	Shortland,[104] Rowland[105]		
Value of the feedback			Cederblom and Lounsbury,[98] Shortland,[104] Martin and Double[107]
Promotes self-knowledge and personal development through observation of others (reflective approach)	McKeachie,[34] Cosh,[103] Shortland,[104] Martin and Double,[107] Blackwell and McLean[108]		
Qualifications or competence of observers	Berk et al.,[9] Root,[101] Cosh[103]	Martin and Double[107]	Goldstein,[5] Sullivan and Skanes[106]
No evidence of instructor self-improvement	Cosh[103]	Berk et al.,[9] Overall and Marsh,[52] Marsh and Roche,[90] Shortland,[104] Martin and Double,[107] Cohen and McKeachie[113]	
Bad feelings, defensiveness, interpersonal conflict	Cosh,[103] Bridges[109]	Shortland[104]	Goldstein[5]
Lack of confidence in the outcomes or reliability of results	Braskamp et al.[102]	Cosh,[103] Lonsdale[110]	Berk et al.,[9] Cederblom and Lounsbury[98]
Fear and resistance	Jarzabkowski and Bone,[8] Martin and Double[107]	Berk et al.,[9] Shortland,[104] Lonsdale[110]	Cederblom and Lounsbury[98]
Personal subjectivity	Berk et al.[9]	Cederblom and Lounsbury[98]	
Limited number of observers	Berk et al.[9]		
Observation is time-consuming	DeZure[114]	Berk et al.[9]	
Observation of a single class is not representative of an entire course	Braskamp et al.,[102] Centra[123]	Berk et al.[9]	
Reputation, tenure, or faculty status influences observer bias ("halo" effect)	Berk et al.[9]	Cederblom and Lounsbury[98]	Hoyt and Pallett,[48] Sullivan and Skanes[106]
Gender, race, sexual orientation bias	Berk et al.[9]		
Observers exchange positive ratings for one another (back scratching)	Berk et al.[9]		
Function or use of peer input or feedback	Martin and Double[107]	Cederblom and Lounsbury[98]	
Previous peer ratings		Cederblom and Lounsbury[98]	
Relevancy of teaching portfolio	Hoyt and Pallett,[48] Seldin[119]		
Ratings related to learning or achievement		Doyle and Crichton[47]	

focuses on providing feedback to a colleague for purposes of enhancing teaching, but the content of the feedback should not be used as a data source for formal performance evaluations or promotion and tenure decisions.

It has been suggested that friendship bias, reputation, tenure or faculty status, informal collegial relationship, and personal subjectivity can result in the skewing of faculty peer evaluations.[9,47,48,60,98,101-106] Therefore, faculty tend to be more accepting of peer evaluations when the outcomes are used for feedback purposes and collaborative reflection rather than as a basis for formal performance assessment, which may have wage implications for promotion.[34,98,103,107,108] Studies have suggested that peer evaluation can create bad feelings, defensiveness, and interpersonal conflict among faculty.[5,103,104,109] This can also result in fear and resistance from those being evaluated.[8,9,98,104,107,110] The limited number of observers who are involved in the faculty peer review process has been described as a disadvantage,[9] although recent use of videos of teaching in action allows for more observers with a greater scheduling flexibility for review.[111,112] Others question the qualifications or competence of peers to function as reviewers of teaching effectiveness.[5,9,101,103,106,107] There is also debate as to whether the rating scales actually measure effective teaching characteristics[9,16] or result in instructor improvement.[9,102-104,107,110,113]

Although there are reports that peer evaluations are biased, unreliable, unfair, and limited in scope and may lead to overly positive reviews if used for formal (summative) performance evaluations,[60,103,104,107,114,115] peer reviews are necessary to augment other assessment methods, such as student ratings or self-evaluations.[1,5,47,98,104,107] Where summative peer evaluations are performed, they should be done with the same rigor as a peer review process for assessment of research and scholarly activity, and they may result in teaching awards.[3] For this purpose, many academic institutions encourage the use of teaching portfolios to formalize the process.[8,9,111,116-119] In summary, peer evaluations, especially formative ones, are valuable when conducted in conjunction with other types of evaluations to counter the potential bias of the peer evaluator.

Self-Assessment

Self-evaluation or self-assessment is another method used in evaluating teaching effectiveness (Table 3).[1,56,112,116,117,120] Self-assessment involves a faculty self-report that describes teaching, scholarship, service, and practice. Although this self-report is not a true assessment of teaching, it has merits mainly by enforcing a reflective approach for self-development.[15,34,103,104,107,108,116,121] One process used for self-evaluation, specifically for formative purposes, is the review of videos of teaching in action.[1,112] Lee and Wu[112] have suggested five potential benefits of the use of videos: 1) easy access to video content, 2) comprehensive view of teaching styles, 3) shared expertise, 4) venue for constructive feedback, and 5) greater utilization of peer-to-peer learning.

However, some investigators contend that the reliability of self-assessment is low because individuals have been shown to provide overly positive self-appraisals in comparison to other data sources[47,56,122,123] and because assessments may be inaccurate, overstated, or self-serving, especially when it comes to summative decisions.[14,16,56,116,120,124]

Table 3. Literature findings of factors influencing self-evaluations

Self-Related Issue	Agree	Disagree	Open to Debate
Gain awareness of different teaching styles through observation (reflective approach) for self-development	Cosh,[103] Beck et al.,[116] Tucker et al.[121]		Marsh[15]
Subjective view of ability (inflated view of skill) or leniency effect	Doyle and Crichton,[47] Kulik and Kulik[122]	Ross and Bruce,[56] Centra[123]	
Validity and fairness of using own portfolios or self-reports	Braskamp and Ory[16]	Seldin[14]	Beck et al.[116]
Efficacy—reliance on own skill	Ross and Bruce,[56] Ross[124]		
Influence from peers (peer input and feedback)	Ross and Bruce,[56] Cameron and Pierce[115]		
Ratings related to learning or achievement		Doyle and Crichton[47]	

The real value of self-reflection is in the opportunity to demonstrate one's own perception of teaching effectiveness.[15] An underexplored question in the literature on assessment of teaching effectiveness is: does teacher self-assessment support or contradict student and peer evaluations? Self-evaluation is a common form of assessment among universities and higher education, and has a greater validity when used in conjunction with other teaching effectiveness strategies, such as peer and student evaluations.[1]

Triangulation Model

Amidst all of the debate regarding methods of assessing teaching effectiveness (Tables 1, 2, and 3), there is recognition of independent value in using student, peer, and self-evaluations as methods of assessing teaching effectiveness. However, when these methods are combined, a triangulation model is achieved, resulting in a more comprehensive assessment outcome.[9,14-16]

Berk has identified twelve different strategies to measure teaching effectiveness.[1] These include elements from students, peers, and self. Based on a modification of these strategies, we developed a survey for applications specific to dental education in order to identify the incidence of commonly used methods of assessing teaching effectiveness in U.S. dental schools and to find evidence of use of triangulation.

Methods

A twelve-question multiple-choice survey was created with four options per question (yes; if yes, performed routinely; no; not sure), plus a comments section (Table 4). The survey was sent electronically via the American Dental Education Association

Table 4. Survey questions and distribution of responses on methods of assessment of teaching effectiveness in U.S. dental schools, by number and percent of total responses

Responses to survey (n=36)

Question	Yes		If yes, performed routinely		No		Not sure	
	Responses	Percent	Responses	Percent	Responses	Percent	Responses	Percent
1. Student written evaluations or ratings	33	91.67%	29	80.56%	2	5.56%	1	2.78%
2. Student interviews, verbal evaluations	20	55.56%	13	36.11%	16	44.44%	1	2.78%
3. Peer review of written documents, such as course syllabi, course goals and objectives, lectures, outlines	23	63.89%	19	52.78%	12	33.33%	1	2.78%
4. Peer observation of lectures, seminars, and/or videos	18	50.00%	6	16.67%	15	41.67%	2	5.56%
5. Peer ratings	9	25.00%	2	5.56%	25	69.44%	2	5.56%
6. Faculty self-evaluation in a written form	18	50.00%	11	30.56%	17	47.22%	1	2.78%
7. Self-evaluation by reviewing videos of teaching in action	2	5.56%	0	0.00%	33	91.67%	1	2.78%
8. Exit survey of students	32	88.89%	27	75.00%	3	8.33%	1	2.78%
9. Teaching awards by institution	32	88.89%	23	63.89%	2	5.56%	1	2.78%
10. Learning outcome measures: exams, national boards, etc.	31	86.11%	26	72.22%	4	11.11%	1	2.78%
11. Teaching portfolios	19	52.78%	8	22.22%	16	44.44%	2	5.56%
12. Does your dental school have a formal mechanism to evaluate clinical teaching effectiveness?	21	58.33%	15	41.67%	13	36.11%	2	5.56%

Open-Ended Written Comments (received from seven schools)

"Clinical faculty are evaluated by their department chair annually. Faculty complete a self-report detailing their activities. The chair considers student evaluation and input from their Team Coordinator in his/her evaluation."

"Love to see what you get for feedback."

"Student evaluation of courses and faculty using an Audience Response System."

"Provost requests an annual report of teaching load and activities for evaluation."

"Courses are evaluated by students. A variety of outcomes measures are used (several noted in survey) to evaluate courses, curriculum, and the institution. However, faculty teaching effectiveness is not assessed."

(ADEA) Academic Affairs listserv to academic deans of all U.S. dental schools in July and again in August of 2007. Each school was allowed only one submission. Out of fifty-seven U.S. dental schools, thirty-nine submitted survey responses. Three submissions were eliminated because, in each case, the schools had answered only one of the twelve questions. Results of the remaining thirty-six surveys were tabulated and analyzed to determine commonly used methods of teaching effectiveness for an actual response rate of 63 percent.

Additionally, the following process was used to determine the percentage of schools that employ a triangulation model (Table 5). Where a school had a response of "yes, performed routinely" in each category of student (question 1, 2, or 8), peer (question 3, 4, 5, 9, 10, or 11), and self (question 6 or 7) review, the school was identified as having achieved triangulation (Table 6). Question 12, pertaining to a formal mechanism of assessment of clinical teaching effectiveness, was excluded in the definition of triangulation because it was not specific to any single defined category (student, peer, or self).

Table 5. Questions assigned to categories for determining triangulation

Evaluation Category	Related Question
Student	• Student written evaluations or ratings • Student interviews, verbal evaluations • Exit survey of students • Learning outcome measures: exams, national boards, etc.
Peer	• Peer review of written documents such as course syllabi, course goals and objectives, lectures, outlines • Peer observation of lectures, seminars, and/or videos • Peer ratings • Teaching awards by institution • Teaching portfolios
Self	• Faculty self-evaluation in a written form • Self-evaluation by reviewing videos of teaching in action

Table 6. Number of schools and related percentages for those responding "Yes, performed routinely" to each evaluation category: student, peer, self

Evaluation Category	Responded "Yes, performed routinely" Number of Schools (n=36)	Percent
Student	29	80.56%
Peer	28	77.78%
Self	11	30.56%
Triangulation (Student and Peer and Self)	7	19.44%

Results

The survey results are shown in Table 4, including the written comments that were provided by seven of the thirty-six schools. For all questions, there was a higher response to an evaluation method simply being used, in contrast to its being "performed routinely." Of the eleven strategies (questions 1-11), the most commonly used methods "performed routinely" were student written evaluations or ratings (81 percent); exit survey of students (75 percent); and learning outcome measures, exams, and national boards (72 percent). The three least routinely used assessment methods were self-evaluation by reviewing videos of teaching in action (0 percent); peer ratings (6 percent); and peer observation of lectures, seminars, and/or videos (17 percent). Less than half (42 percent) of the schools routinely performed clinical teaching effectiveness evaluations.

To determine the extent of triangulation, the survey questions were assigned to one of the three evaluation categories (Table 5). Review of the survey results (Table 6) illustrates that, of the three evaluation categories (student, peer, and self), the most commonly used method ("performed routinely") was student evaluation. Nearly 20 percent of the responding schools reported using all three categories of student, peer, and self, and thus potentially had implemented triangulation in the performance assessment process.

Discussion

Routine performance of evaluations is important for the development of the faculty, the department, and the institution. To optimize the value of the assessment outcomes, such evaluations should include elements from student, peer, and self-evalu-

ation categories. Approximately 20 percent of the responding schools reported using multiple assessment sources. However, it cannot be determined from these data whether the evaluations were purposefully combined into a single assessment (i.e., triangulation). Moreover, surveys addressed to associate deans for academic affairs have the potential limitation of not yielding the most current appraisal of teaching effectiveness strategies utilized in all departments since department chairs play a key role in the assessment of teaching effectiveness for their faculty. However, the intent of this survey was to explore institution-wide policies, which fall within the responsibilities of the associate deans for academic affairs.

The literature related to assessment of teaching effectiveness recommends that schools embrace a triangulation policy for assessment of teaching effectiveness. This formal mechanism should be used in the performance assessment process for developmental (formative) and evaluative (summative) purposes by all departments. Triangulation should be done on a routine basis because if the evaluation method is sporadic, intermittent, or limited by department, the effectiveness of the assessment is compromised. Table 5 can be used as a guide in selecting particular evaluation methods from each of the categories. However, within each method, institutions have much leeway in developing their own assessment components to meet their specific goals. For example, the specifics of exit surveys, student written evaluations or ratings, or components of teaching portfolios are within the control of each institution. There is currently no evidence to suggest that certain types of evaluation methods from each category of student, peer, and self can be combined in a particular way to achieve an optimal level of triangulation. Instead, it is the combination of methods, choosing at least one from each category, that results in the triangulation model that is advocated.[1,4-12]

Student evaluations are valuable when related to course organization and instructor delivery,[1,15,25,44,46,47] but not when students are asked to evaluate course content, relevancy, or appropriateness of goals and objectives. Many dental schools, however, value student views regarding curriculum. The ADEA Senior Survey is primarily devoted to asking students their opinions about curricular content, and this evaluation tool is widely used in dental schools to justify curriculum decisions. Although the ADEA survey is valuable in providing a scope of different curricula nationally, its interpretation for all curriculum changes should be restricted because of student limitations in judging the appropriateness of content while they are still in school and prior to experiencing the realities of actual dental practice in the community.

Peer reviews are more appropriate than student assessment for content evaluation.[1,15] However, sustaining a viable peer review system may be time-consuming and resource-intensive.[97] Many faculty members may be reluctant to embrace another evaluation process, considering that they are overwhelmed with existing job responsibilities and tenure requirements.

Another challenge to a peer review process is related to the current culture and environment in dental schools. Faculty are likely to hesitate in critiquing one another for fear of creating bad feelings that can damage future working relationships.[103,109] In reality, however, the promotion and tenure system at most major universities now expect submission of peer appraisals of various aspects of academic performance—more commonly on the research side, but increasingly for teaching. Dental schools will need to determine a way to successfully implement peer reviews of teaching for summative purposes. For example, when selecting individuals to perform summative peer reviews, if faculty are allowed to identify their own reviewers, then "buddy bias" is a significant issue.[9,60,98] If department chairs designate peer reviewers, there is less potential for friendship bias, but faculty may fear that chairs are out to "get them" by purposely choosing reviewers who have interpersonal conflicts, biases, or other issues with the reviewed faculty member.[5,103,109] These realities need to be considered when using peer evaluations for summative purposes.

Peer assessments, together with self-evaluations,[103,116,121] have the greatest value in focusing faculty on a process of routine self-reflection and determining or redirecting faculty efforts. One method of evaluation—review of video of teaching in action—warrants a discussion. The review of videos is a well-described mechanism for self- or peer evaluation.[1,111,112] Based on the findings from this survey, it does not appear that any U.S. dental schools utilize "videos of teaching in action" as an evaluation mechanism. Video review may be more timely now than ever before. Students today are not limited to note-taking or traditional transcription. There is an increasing use of technology for recording voice and video of lectures and other teaching activities. These digital recordings can be stored,

reviewed, shared, analyzed, and used as tools for training and development. The fact that students are using this technology for study or review purposes makes it even more critical that the faculty evaluate these videos for teaching effectiveness, prior to or in conjunction with the student review.

While not directly related to the main issue in this study, the fact that only 42 percent of the responding schools reported that they routinely evaluated clinical teaching effectiveness is disconcerting. This is a low percentage considering that more than half of the students' total time in dental school is in the clinic/patient care environment and this phase of the curriculum certainly is the students' most powerful learning experience. It is in the best interest of dental schools to focus on evaluating clinical teaching effectiveness with the same level of commitment that is given to evaluating classroom teaching effectiveness.[21-23]

The process of evaluating teaching effectiveness can only have value if there is a forum for formalized feedback of these outcomes to create action plans for continuous improvement. In this manner, the students, the faculty, and the institution will greatly benefit from these evaluations.

Future studies are indicated to determine an optimal combination of evaluation methods for each category of student, peer, and self in achieving triangulation.

Limitations of the Study

The survey sent to the U.S. dental schools was used to identify the incidence of commonly used methods of assessing teaching effectiveness and to find evidence of triangulation. The survey was intended to collect data on methods used, from which a triangulation model could be proposed. However, none of the questions in the survey directly or indirectly asked respondents whether the school used a triangulation method. At the time of survey construction, it was perceived that the term "triangulation" was not commonly used or understood among dental faculty. Although triangulation is well defined in the education literature, its use in the dental literature is uncommon. For those schools that showed potential evidence of triangulation, there is no way to determine whether their collection of data from multiple sources was consciously done to implement a triangulation strategy.

Conclusion

The results of this study are as follows:

1. It was determined that a triangulation method for assessing teaching effectiveness was potentially performed by only 19 percent of the responding schools, although there was no evidence derived from this survey to determine if the one in five schools that reported multiple data sources have consciously determined to implement a triangulation process.
2. The three most commonly used evaluation methods, identified as "performed routinely," were all from the student evaluation category (written evaluations, exit surveys, and exam results).
3. The self-evaluation category was found to be the least utilized area for assessment (31 percent) as compared to the peer (78 percent) and student (81 percent) categories.
4. Fewer than half of the responding schools routinely evaluated clinical teaching effectiveness (42 percent).
5. To improve the assessment of teaching effectiveness, the use of a triangulation model is suggested, whereby evaluation methods from all three categories of students, peers, and self are performed routinely.

REFERENCES

1. Berk RA. Survey of 12 strategies to measure teaching effectiveness. Int J Teach Learn Higher Educ 2005;17(1):48–62.
2. Knapper C. Broadening our approach to teaching evaluation. New Dir Teach Learn 2001;88(4):3–9.
3. French DP. What currency should we use? J College Sci Teach 2006;35(4):60.
4. Jick TD. Mixing qualitative and quantitative methods: triangulation in action. Administrative Sci Q 1979;24(4): 602–11.
5. Goldstein J. Making sense of distributed leadership: the case of peer assistance and review. Educ Eval Policy Analysis 2004;26(2):173–97.
6. Golafshani N. Understanding reliability and validity in qualitative research. Qual Rep 2003;8(4):597–607.
7. Felder R, Brent R. How to evaluate teaching. Chem Eng Educ 2004;38(3):202–4.
8. Jarzabkowski P, Bone Z. A "how-to" guide and checklist for peer appraisal of teaching. Innovat Educ Teach Int 1998;35(2):177–82.
9. Berk RA, Naumann PL, Appling SE. Beyond student ratings: peer observation of classroom and clinical teaching. Int J Nurs Educ Scholarship 2004;1(1):1–26.
10. Belgrave LL, Zablotsky D, Guadagno MA. How do we talk to each other? Writing qualitative research for quantitative readers. Qual Health Res 2002;12(10):1427–39.

11. Drapeau M. Subjectivity in research: why not? But…. Qual Rep 2002;7(3):1–14.
12. Bowen GA. Preparing a qualitative research-based dissertation: lessons learned. Qual Rep 2005;10(2):208–22.
13. Arreola RA. Developing a comprehensive faculty evaluation system: a handbook for college faculty and administrators on designing and operating a comprehensive faculty evaluation system. Bolton, MA: Anker Publishing, 2000.
14. Seldin P. Self-evaluation: what works? what doesn't? In: Seldin PA, ed. Changing practices in evaluating teaching. Bolton, MA: Anker Publishing, 1999:97–113.
15. Marsh HW. Students' evaluations of university teaching: dimensionality, reliability, validity, potential biases, and utility. J Educ Psychol 1984;76(5):707–54.
16. Braskamp LA, Ory JC. Assessing faculty work: enhancing individual and institutional performance. San Francisco: Jossey-Bass Higher and Adult Education Series, 1994.
17. Soderberg LO. A credible model: evaluating classroom teaching in higher education. Instr Evaluation 1986;8: 13–27.
18. Romberg E. Description of peer evaluation within a comprehensive evaluation program in a dental school. Instr Evaluation 1985;8(1):10–6.
19. Keig L, Waggoner MD. Collaborative peer review: the role of faculty in improving college teaching. ASHE-ERIC Higher Educ Rep 1994(2):18–22.
20. Centra JA. Reflective faculty evaluation: enhancing teaching and determining faculty effectiveness. San Francisco: Jossey-Bass Publishers, 1993.
21. Chambers D, Geissberger M, Leknius C. Association amongst factors thought to be important by instructors in dental education and perceived effectiveness of these instructors by students. Eur J Dent Educ 2004;8(1):18–23.
22. Henzi D, Davis E, Jasinevicius R, Hendricson W. North American dental students' perspectives about their clinical education. J Dent Educ 2006;70(4):361–77.
23. Henzi D, Davis E, Jasinevicius R, Hendricson W. In the students' own words: what are the strengths and weaknesses of the dental school curriculum? J Dent Educ 2007;71(5):632–45.
24. Abrami PC, D'Apollonia S, Cohen PA. Validity of student ratings of instruction: what we know and what we do not. J Educ Psychol 1990;82(2):219–31.
25. Aleamoni LM. Student rating myths versus research facts from 1924 to 1998. J Personnel Eval Educ 1999; 13(2):153–66.
26. Apodaca P, Grad H. The dimensionality of student ratings of teaching: integration of uni- and multidimensional models. Stud Higher Educ 2005;30(6):723–48.
27. d'Apollonia S, Abrami PC. Navigating student ratings of instruction. Am Psychol 1997;52(11):1198–208.
28. Davidovitch N, Soen D. Using students' assessments to improve instructors' quality of teaching. J Further Higher Educ 2006;30(4):351–76.
29. Eiszler CF. College students' evaluations of teaching and grade inflation. Res Higher Educ 2002;43(4):483–501.
30. Greenwald AG. Validity concerns and usefulness of student ratings of instruction. Am Psychol 1997;52(11):1182–6.
31. Harrison PD, Ryan JM, Moore PS. College students' self-insight and common implicit theories in rating of teaching effectiveness. J Educ Psychol 1996;88(4):775–82.
32. Howard GS, Conway CG, Maxwell SE. Construct validity of measures of college teaching effectiveness. J Educ Psychol 1985;77(2):187–96.
33. Marsh HW, Roche LA. Making students' evaluations of teaching effectiveness effective: the critical issues of validity, bias, and utility. Am Psychol 1997;52(11):1187–97.
34. McKeachie WJ. Student ratings: the validity of use. Am Psychol 1997;52(11):1218–25.
35. Theall M, Franklin J. Looking for bias in all the wrong places: a search for truth or a witch hunt in student ratings of instruction? New Dir Institutional Res 2001;27(5): 45–56.
36. Millea M, Grimes PW. Grade expectations and student evaluation of teaching. College Student J 2002;36(4): 582–90.
37. Onwuegbuzie AJ, Witcher AE, Collins KMT, Filer JD, Weidmaier CD, Moore CW. Students' perceptions of characteristics of effective college teachers: a validity study of a teaching evaluation form using a mixed-methods analysis. Am Educ Res J 1997;44(1):113.
38. Sojka J, Gupta A, Deeter-Schmelz D. Student and faculty perceptions of student evaluations of teaching: a study of similarities and differences. College Teach 2002;50(2): 44–9.
39. Sproule R. The underdetermination of instructor performance by data from the student evaluation of teaching. Economics Educ Rev 2002;21(3):287–94.
40. Toland MD, De Ayala R. A multilevel factor analysis of students' evaluations of teaching. Educ Psychol Meas 2005;65(2):272–96.
41. Victoroff KZ, Hogan S. Students' perceptions of effective learning experiences in dental school: a qualitative study using a critical incident technique. J Dent Educ 2006; 70(2):124–32.
42. Wachtel HK. Student evaluation of college teaching effectiveness: a brief review. Assess Eval Higher Educ 1998; 23(2):191–212.
43. Cohen PA. Student ratings of instruction and student achievement: a meta-analysis of multisection validity studies. Rev Educ Res 1981;51(3):281–309.
44. Aleamoni LM. Student ratings of instruction. In: Millman J, ed. Handbook of teacher evaluation. Beverly Hills, CA: Sage Publications, 1981.
45. Cashin WE. Student ratings of teaching: the research revisited. IDEA Paper No. 32. Manhattan: Kansas State University, Center for Faculty Evaluation and Development, 1995:1–10.
46. Aleamoni LM, Hexner PZ. A review of the research on student evaluation and a report on the effect of different sets of instructions on student course and instructor evaluation. Instr Sci 1980;9(1):67–84.
47. Doyle KO, Crichton LI. Student, peer, and self-evaluations of college instructors. J Educ Psychol 1978;70(5): 815–26.
48. Hoyt DP, Pallett WH. Appraising teaching effectiveness: beyond student ratings. Idea Paper No. 36. Manhattan: Kansas State University, Center for Faculty Evaluation and Development, 1999:1–8.

49. Aleamoni LM. The usefulness of student evaluations in improving college teaching. Instr Sci 1978;7(1):95–105.
50. Marsh HW. The validity of students' evaluations: classroom evaluations of instructors independently nominated as best and worst teachers by graduating seniors. Am Educ Res J 1977;14(4):441–7.
51. McKeachie WJ, Lin Y-G, Mendelson CN. A small study assessing teacher effectiveness: does learning last? Contemp Educ Psychol 1978;3(4):352–7.
52. Overall JU, Marsh HW. Midterm feedback from students: its relationship to instructional improvement and students' cognitive and affective outcomes. J Educ Psychol 1979;71(6):856–65.
53. Albanese MA. The validity of lecturer ratings by students and trained observers. Acad Med 1991;66(1):26–8.
54. Hativa N, Raviv A. University instructors' ratings profiles: stability over time and disciplinary differences. Res Higher Educ 1996;37(3):341–65.
55. Palchik NS, Burdi AR, Hess GE, Dielman TE. Student assessment of teaching effectiveness in a multi-instructor course for multidisciplinary health professional students. Eval Health Professions 1988;11(1):55–73.
56. Ross JA, Bruce CD. Teacher self-assessment: a mechanism for facilitating professional growth. Teach Teacher Educ 2007;23(2):146–59.
57. Greenwald AG, Gillmore GM. Grading leniency is a removable contaminant of student ratings. Am Psychol 1997;52(11):1209–17.
58. Ory JC, Ryan K. How do student ratings measure up to a new validity framework? New Dir Institutional Res 2001;27(5):27–44.
59. Shevlin M, Banyard P, Davies M, Griffiths M. The validity of student evaluation of teaching in higher education: love me, love my lectures? Assess Eval High Educ 2000;25(4):397–405.
60. Love KG. Comparison of peer assessment methods: reliability, validity, friendship bias, and user reaction. J Appl Psychol 1981;66(4):451–7.
61. Abrami PC, Dickens WJ, Perry RP, Leventhal L. Do teacher standards for assigning grades affect student evaluations of teaching? J Educ Psychol 1980;72(1):107–18.
62. Abrami PC, Perry RP, Leventhal L. The relationship between student personality characteristics, teacher ratings, and student achievement. J Educ Psychol 1982;74(1):111–25.
63. Centra JA. Use of the teaching portfolio and student evaluations for summative evaluation. Paper presented at Annual Meeting of American Educational Research Association, Atlanta, GA, April 12–16, 1993.
64. Chacko TI. Student ratings of instruction: a function of grading standards. Educ Res Q 1983;8(2):19–25.
65. Koshland DE Jr. Teaching and research. Science 1991;251(4991):249.
66. Nimmer JG, Stone EF. Effects of grading practices and time of rating on student ratings of faculty performance and student learning. Res Higher Educ 1991;32(2):195–215.
67. Powell R. Faculty rating scale validity: the selling of a myth. College Engl 1978;39(5):616–29.
68. Snyder CR, Clair M. Effects of expected and obtained grades on teacher evaluation and attribution of performance. J Educ Psychol 1976;68:75–82.
69. Vasta R, Sarmiento RF. Liberal grading improves evaluations but not performance. J Educ Psychol 1979;71(2):207–11.
70. Marsh HW. Students' evaluations of university teaching: research findings, methodological issues, and directions for future research. Int J Educ Res 1987;11(3):253–388.
71. Theall M, Franklin J. Using student ratings for teaching improvement. New Dir Teach Learn 1991;48(4):83–96.
72. Theall M, Abrami PC, Mets LA. The student ratings debate: are they valid? how can we best use them? San Francisco: Jossey-Bass, 2001.
73. Theall M, Franklin J. Student ratings in the context of complex evaluation systems. New Dir Teach Learn 1990;43(3):17–34.
74. Marsh HW, Cooper TL. Prior subject interest, students' evaluations, and instructional effectiveness. Multivariate Behav Res 1981;16(1):83–104.
75. Ory JC. The influence of students' affective entry on instructor and course evaluations. Rev Higher Educ 1980;4(1):13–24.
76. Amin ME. Gender as a discriminating factor in the evaluation of teaching. Assess Eval High Educ 1994;19(2):135–43.
77. Dukes RL, Victoria G. The effects of gender, status, and effective teaching on the evaluation of college instruction. Teach Sociol 1989;17(4):447–57.
78. Feldman KA. College students' views of male and female college teachers: Part II—evidence from students' evaluations of their classroom teachers. Res Higher Educ 1993;34(2):151–211.
79. Erdle S, Murray HG, Rushton J. Personality, classroom behavior, and student ratings of college teaching effectiveness: a path analysis. J Educ Psychol 1985;77(4):394–407.
80. Feldman KA. The association between student ratings of specific instructional dimensions and student achievement: refining and extending the synthesis of data from multisection validity studies. Res Higher Educ 1989;30(6):583–645.
81. Tang TLP. Teaching evaluation at a public institution of higher education: factors related to the overall teaching effectiveness. Public Personnel Management 1997;26:379–89.
82. Feldman KA. The significance of circumstances for college students' ratings of their teachers and courses. Res Higher Educ 1979;10(2):149–72.
83. Pulich MA. Better use of student evaluations for teaching effectiveness. Improving College University Teach 1984;32:91–4.
84. Svinicki MD. Encouraging your students to give feedback. New Dir Teach Learn 2001;87(3):17–24.
85. Marsh HW. Validity of students' evaluations of college teaching: a multitrait-multimethod analysis. J Educ Psychol 1982;74(2):264–79.
86. Centra JA. Student ratings of instruction and their relationship to student learning. Am Educ Res J 1977;14(1):17–24.
87. McKeachie WJ. Research on college teaching: the historical background. J Educ Psychol 1990;82(2):189–200.
88. Cohen PA. Effectiveness of student-rating feedback for improving college instruction: a meta-analysis of findings. Res Higher Educ 1980;13(4):321–41.

89. Brinko KT. The interactions of teaching improvement. New Dir Teach Learn 1991;48(4):39–49.
90. Marsh HW, Roche L. The use of students' evaluations and an individually structured intervention to enhance university teaching effectiveness. Am Educ Res J 1993;30(1):217–51.
91. Seldin P. The use and abuse of student ratings of professors. Chron Higher Educ 1993;39(3):40.
92. Seldin P. Changing practices in faculty evaluation. San Francisco: Jossey-Bass, 1984.
93. Ramsden P. A performance indicator of teaching quality in higher education: the course experience questionnaire. Stud Higher Educ 1991;16(2):129–50.
94. Marsh HW, Dunkin MJ. Students'evaluations of university teaching: a multidimensional perspective. Higher Educ: Handbook Theory Res 1992;8:143–233.
95. Koon J, Murray HG. Using multiple outcomes to validate student ratings of overall teacher effectiveness. J Higher Educ 1995;66(1):61–81.
96. Centra J. Reflective faculty evaluation. San Francisco: Jossey-Bass, 1993.
97. Irby D. Evaluating instructional scholarship in medicine. J Am Podiatr Med Assoc 1993;83(6):332–7.
98. Cederblom D, Lounsbury JW. An investigation of user acceptance of peer evaluations. Personnel Psychol 1980;33(3):567–79.
99. Aleamoni LM. Components of the instructional setting. Instr Evaluation 1982;7:11–6.
100. Keig LW, Waggoner MD. Peer review of teaching: improving college instruction through formative assessment. J Excellence College Teach 1995;6(3):51–83.
101. Root LS. Faculty evaluation: reliability of peer assessments of research, teaching, and service. Res Higher Educ 1987;26(1):71–84.
102. Braskamp LA, Brandenburg DC, Ory JC. Evaluating teaching effectiveness: a practical guide. Beverly Hills, CA: Sage Publications, 1984.
103. Cosh J. Peer observation in higher education: a reflective approach. Innovat Educ Teach Int 1998;35(2):171–6.
104. Shortland S. Peer observation: a tool for staff development or compliance? J Further Higher Educ 2004;28(2):219–28.
105. Rowland S. The enquiring university teacher. London: Society for Research into Higher Education, Ltd., 2000.
106. Sullivan AM, Skanes GR. Validity of student evaluation of teaching and the characteristics of successful instructors. J Educ Psychol 1974;66(4):584–90.
107. Martin GA, Double JM. Developing higher education teaching skills through peer observation and collaborative reflection. Innovat Educ Teach Int 1998;35(2):161–70.
108. Blackwell R, McLean M. Peer observation of teaching and staff development. Higher Educ Q 1996;50(2):156.
109. Bridges EM. The incompetent teacher: the challenge and the response. Philadelphia: Falmer Press, 1986.
110. Lonsdale AJ. Academic staff appraisals in Australian higher education. Part 1: principles and guidelines. Canberra: Australian Government Pub. Service, 1988.
111. Milman NB. Web-based digital teaching portfolios: fostering reflection and technology competence in preservice teacher education students. J Technology Teacher Educ 2005;13(3):373.
112. Lee GC, Wu CC. Enhancing the teaching experience of pre-service teachers through the use of videos in web-based computer-mediated communication (CMC). Innovat Educ Teach Int 2006;43(4):369–80.
113. Cohen PA, McKeachie WJ. The role of colleagues in the evaluation of college teaching. Improving College University Teach 1980;28(4):147–54.
114. DeZure D. Evaluating teaching through peer classroom observation. In: Seldin PA, ed. Changing practices in evaluating teaching. Bolton, MA: Anker Pub. Co., 1999:70–96.
115. Cameron J, Pierce WD. Reinforcement, reward, and intrinsic motivation: a meta-analysis. Rev Educ Res 1994;64(3):363–423.
116. Beck RJ, Livne NL, Bear SL. Teachers' self-assessment of the effects of formative and summative electronic portfolios on professional development. Eur J Teach Educ 2005;28(3):221–44.
117. Leggett M, Bunker A. Teaching portfolios and university culture. J Further Higher Educ 2006;30(3):269-82.
118. De Rijdt C, Tiquet E, Dochy F, Devolder M. Teaching portfolios in higher education and their effects: an explorative study. Teach Teacher Educ 2006;22(8):1084-93.
119. Seldin P. The teaching portfolio: a practical guide to improved performance and promotion/tenure decisions. 2nd ed. Bolton, MA: Anker Pub. Co., 1997.
120. Centra JA. Self-ratings of college teachers: a comparison with student ratings. J Educ Measurement 1973;10(4):287-95.
121. Tucker PD, Stronge JH, Gareis CR. Handbook on teacher portfolios for evaluation and professional development. Larchmont, NY: Eye on Education, 2002.
122. Kulik JA, Kulik CLC. Student ratings of instruction. Teach Psychol 1974;1(2):51–7.
123. Centra JA. Self-ratings of college teachers: a comparison with student ratings. J Educ Measurement 1973;10(4):287–95.
124. Ross JA. Teacher efficacy and the effects of coaching on student achievement. Can J Higher Educ 1992;17(1): 51–65.