There are many questions about student evaluations, in general, and online evaluations specifically. These FAQs\(^1\) will attempt to answer those questions.

1. *Aren’t student ratings popularity contests?* No – if the form you use asks questions tested for validity and reliability, and was designed using psychometric procedures, then the results are both valid and reliable. They report various aspects of faculty performance with accuracy.

2. *Aren’t student rating forms just plain unreliable and invalid?* Yes and no – most forms used today were created in departmental meetings or committees and do not use questions that have been measured for validity and reliability. Therefore, yes, those forms are invalid and unreliable. Boise State University’s rating forms have undergone much scrutiny over the past year and most departments are now using questions that are both valid and reliable.

3. *Aren’t students too immature, inexperienced, and capricious to make any consistent judgments about the instructor and instruction?* No – research dating back to 1924 has all come to the same conclusion that this is an inaccurate statement. The stability of student ratings from one year to the next results in substantial correlations in the range of .087 to .089.

4. *Isn’t it true that I can buy good student ratings just by giving easy grades?* No – this aspect of student evaluations has been studied more than any other. Nearly 500 studies on this very topic have all reached the same conclusion – there is no consistent correlation between the grades a faculty member gives and the ratings he or she receives. The correlations appearing in the literature range from -.18 to .18. There is also a common belief that students with higher grades give higher ratings. Twenty-two studies have reported zero correlation and 28 studies have found a positive relationship. However, in those 28 studies, the majority of studies found a relatively weak relationships – the median correlation was about 0.14, the mean was 0.18 and the standard deviation was 0.16. When all studies are combined in a meta-analysis, the average correlation was 0.0.

---

5. *Isn’t it easier to get good ratings in higher level courses?* Yes – each level (freshmen – seniors) tend to rate courses more harshly than the level above them. Seniors tend to rate courses most positively while freshmen rate courses most harshly.

6. *Isn’t it true that the students who are required to take a course tend to rate the course more harshly than those take it as an elective?* Yes – the majority of studies have found this to be an accurate statement (though 2 studies found no correlation).

7. *Isn’t there a gender bias in student ratings? Don’t female faculty tend to get lower ratings than male faculty?* The answer to these questions appears to be no. There are far fewer studies in this area than in other aspects of ratings, and the results seem to be less consistent. Some studies have found no difference, some find that males are rated higher than females, but more studies have found that female students tend to rate female instructors higher than male instructors. Other studies found that females rate all instructors higher on some sub-categories of evaluations.

8. *Isn’t it more difficult for math and science faculty to get good ratings?* Yes – though this doesn’t mean the math and science professors are inferior instructors. Humanities courses tend to receive higher ratings than math and science, but the specific reason or reasons for this have eluded researchers.

9. *Isn’t it true that the only faculty who are really qualified to teach or evaluate their peer’s teaching are those who are actively involved in conducting research in their field?* No – though at one time it was believed that teaching and research were so closely aligned that one needn’t evaluate them separately. A few studies have found weak positive correlations between teaching and research, but an equal number of studies have found no correlation between teaching and research productivity. Additionally, studies have found no correlation between research productivity and peer’s evaluations of their teaching. Two recent studies (Hattie & Marsh, 1997; Hattie, 2002) have ended the speculation. Their highly regarded research puts the correlation between productivity and teaching effectiveness at effectively zero. It is becoming clear that research in a subject area does not translate into disciplinary pedagogy.

10. *Don’t students have to be away from the course, and possibly the college, for several years before they are able to make accurate judgments about the instructor and instruction?* No - in most circumstances. Some items have been deemed inappropriate (e.g., I will be able to use the information/skills learned in this course in my future career; My instructor was knowledgeable in this subject area; My instructor carefully designed this course, etc.) for student evaluations because no wording can produce reliable responses. Students are not in a position to be able to accurate assess these topics. Graduates who have been working for a number of years, however, are in a position to evaluate how well a course prepared them for their career. In these situations, the answer to the question would be yes. In the topics that belong on student evaluations, the answer is no.

11. *Isn’t it true that class size affects student ratings?* Apparently the answer is no. Clarification is in order. Often, when faculty teach a course in small sections and
receive high ratings for their teaching effectiveness, but then move to teaching larger sections, their ratings do drop while the instructor gains facility in teaching large sections. Once the instructor “gets the hang” of teaching large classes, the ratings seem to mirror their ratings in the smaller sections. The correlation here seems to be much more closely tied to how well prepared the instructor is and how much learning the students perceive took place in the course. Interestingly, though, most large courses are taught in lower division courses where student ratings tend to be lower anyway.

12. Does the time of the day the course is taught affect student ratings? We don’t think so. This issue has not received much attention, but in the limited research (3 studies) available, time of day does not affect the ratings of instructors.

13. Do majors in a course rate it differently than nonmajors? No – this area, too, has limited research, but four studies that examined this; there were no significant differences between student ratings and their major status.

14. Does the rank of instructor affect student ratings? Not really – instructors of higher rank tend to teach upper division courses (where we do find differences). It isn’t clear that the rank of the instructor played a role in the higher ratings.

15. Won’t response rates be lower when we move to online evaluations? Yes – response rates tend to drop in the first semesters of use. While public perception may be that pen and paper evaluations have nearly a 100% response rate, research suggests the figure is likely closer to 75-80% (due to absences, those who decline to participate, etc.)

16. Won’t lower response rates reduce my ratings? No – the evidence is clear that lower response rates do not equate with lower ratings. In fact, the research shows that ratings remain either the same or go up slightly when universities move to online course evaluations.

17. Moving to online course evals will mean that students who were unhappy with the course are more likely to respond than those who enjoyed it, right? No – the literature is clear that students who enjoyed the course are more likely to complete the survey than those who were either ambivalent or actively disliked the course.

18. Won’t lower response rates mean that the comments are shorter or less useful? No – the evidence is clear on this topic. Students tend to write much more and the information they provide in open-ended responses is more useful than comments written on paper (online responses are more directional and focused). Researchers suggest that since students aren’t rushed into completing the evaluations in the 10 minutes (or less) that faculty tend to provide with the paper versions, online responses are more thoughtful and discuss more topics than paper versions.
Suggestions for asking questions

1) Consider what you might be most willing to change about the course (altering pedagogy, adding/removing topics, etc.) and formulate questions that encourage students to provide feedback about that
2) Use quantitative measures (see handout) when you want a snapshot of a topic
3) Use qualitative measures when you want a broader response
4) Be careful not to conflate one teaching feature with another (don’t ask about amount of time homework takes if you’re actually interested in how well students are gaining a conceptual understanding)

Things to avoid when writing questions

1) Questions that ask a student to assess what other students think (students in this course felt comfortable enough to ask questions; INSTEAD, posit “I felt comfortable enough to ask questions”)
2) Questions that ask what the instructor is thinking about them (The instructor believed that the students in this course were well-prepared for each class meeting.)
3) Questions that relate to certain aspects of teaching about which the students have little or no knowledge
   a. This include instructor’s content expertise
   b. Types of teaching methods
   c. Quality of evaluation techniques.

Questions should not have (or be)

1) Absolutes (always, never)
2) Non-absolutes that are warm and fuzzy (sometimes, usually frequently, often, etc.)
3) Value-laden or inflammatory words
4) Jargon
5) Questions that overlap other questions
6) Factual (Syllabus was handed out on the first day of class)
7) Two or more behaviors or thoughts

Syracuse University has an item bank for quantitative items. It is online at
https://oira.syr.edu/Assessment/StudentRate/ItemBank.htm You may also download a pdf version of the bank there.