Student Evaluation of Teaching

Highline College Common Questions

Prepared by Sue Frantz on behalf of the Highline College Tenure Review Committee

September, 2020

Version 2.0

For evaluating faculty performance, there is no one best way. Since each measure has its own shortcomings, the best recommendation is to use multiple measures (Boysen et al., 2015; Kite, 2012). Highline College’s Tenure Review Committee (TRC) does just that. With a combination of peer observations, self-reflection, and student evaluations of teaching—and discussion of each with the tenure working committee or the post-tenure committee—our process is robust. The TRC has encouraged strengthening of peer observations through, for example, the use of the Peer Led Assessment of Instructional Design (PLAID) rubric and the strengthening of the self-reflection by asking that each of the tenure criteria be addressed. Now it is time to give serious consideration to our student evaluations of teaching (SETs).

There is no such thing as a questionnaire that will yield unbiased results. All questions we answer are answered through our own personal lens, even seemingly objective questions. For example, “Does your instructor start class on time?” is a straight-forward, objective question. To answer the question, however, the student has to rely on their memory to answer it, presuming that they were not writing down the start time of every class session as the quarter progressed. Students who like the instructor may have an easier time remembering the times when class *did* start on time. Student who do not like the instructor may have an easier time remembering the times when class *did not* start on time. Why does the student like or dislike a particular instructor? Unfortunately, many factors can influence a student’s opinion of an instructor other than instructor effectiveness, such as the instructor’s gender (Bavishi et al., 2010; Bennett, 1982; Boring, 2017; MacNell et al., 2014; Mitchell & Martin, 2018), ethnicity (Bavishi et al., 2010), and physical attractiveness (Ambady & Rosenthal, 1993).

To further complicate matters, there is evidence that SETs do not positively correlate with student learning (Uttl et al., 2017). In fact, they may negatively correlate (Carpenter et al., 2020)—on average, the better ratings instructors get, the poorer students do in future courses. In terms of the quality of instruction, students are unreliable narrators. To help us evaluate the quality of instruction, we must rely on other tools: peer observation and self-evaluation that specifically addresses the tenure criteria.

# What value do course evaluations have?

SETs can tell us how much a student enjoyed a course. There is evidence that “even if that enjoyment is not correlated with learning, [it] could encourage them to continue on to advanced coursework in that domain” (Oppenheimer & Hargis, 2020, p. 171). It is not too far of a leap to say that if students do not enjoy their courses, they are less likely to persist—in those courses or even in college.

SETs give students the opportunity to provide feedback, and that alone helps students feel empowered (Lind et al., 1990). “Given the large literature showing that the opportunity to provide feedback drives perceptions of procedural justice, and the fact that procedural justice drives a host of positive educational outcomes, there is good reason to think that offering students a chance to complete course evaluations can itself lead to beneficial consequences” (Oppenheimer & Hargis, 2020, p. 172).

Qualitative responses may be less subject to bias than quantitative ratings and mid-term evaluations may be less subject to memory biases (Carpenter et al., 2020). Neither approach will guarantee unbiased SETs, but they may mitigate the impact of bias.

Research evidence tells us—at the core—students want a few basic things from their instructors. Students want instructors who care about them and are supportive, students want instructors who know what they are doing, and students want instructors who have strong communication skills (Keeley et al., 2006). Although, in terms of what students want, we need to keep in mind cultural differences. For example, Chinese students rate instructor approachability as less important than Japanese or American students (Liu et al., 2015).

# Rationale for the common SET questions

Each of the proposed SET questions (see below) are related to student perceptions of instructor support, instructor competence, and instructor communication, each with a corresponding non-exhaustive list of examples. The content of each is based on the heavily-researched Teacher Behavior Checklist (Keeley et al., 2006). While each question asks for a quantitative assessment, students are invited to write more in an open-ended textbox.

Most Highline College students come from a collectivist cultural background where interdependence is valued over independence, such as members of the working class, women, and those who grew up in Africa or Asia (Markus & Conner, 2013). While higher education tends to be framed in terms of independence (e.g., “become an independent thinker”) students who value interdependence (e.g., “I’m in college to be able to better help my family”) may feel like their goals aren’t valued (Stephens et al., 2012). This cultural mismatch makes it harder for students with a collectivist cultural background to persist at difficult tasks. Using interdependent language reduces that cultural mismatch and increases the persistence of students from collectivist backgrounds. The switch to interdependent language has no impact on students who lean individualistic (Stephens et al., 2012). For this reason, in these SET questions we chose to use collectivist language (e.g., “our instructor encourages us”) instead of individualistic language (e.g., “my instructor encourages me”).

While these questions are designed as an end-of-term evaluation, the use of formative evaluations are encouraged.

All SETs will be evaluated by the TRC with care (e.g., awareness of potential bias, patterns in the SETs will be considered more heavily than outliers) and in the context of other information (e.g., peer observations, self-reflection).

Instructors, departments, and divisions are welcome to use additional SET questions as they would like. While probationers may include student responses to these other questions, the only required questions for the TRC are the following.

# The questions

1. Our instructor is respectful, approachable, and personable (for example, answers our questions, responds to our emails, welcomes our contributions to class, encourages us to ask questions, listens to our ideas) [*Always Frequently Sometimes Rarely Never*]

TEXT BOX for comments

1. Our instructor encourages and cares for students (for example, knows our names, provides constructive feedback, explains how to do things better, encourages us to ask for help, helps those who need it, makes changes to the class as needed, meets us outside office hours) [*Always Frequently Sometimes Rarely Never*]

TEXT BOX for comments

1. Our instructor is knowledgeable about the subject matter (for example, gives clear answers to questions, explains things in different ways; doesn’t lecture straight from the book or read slides, uses relevant examples) [*Always Frequently Sometimes Rarely Never*]  
     
   TEXT BOX for comments

1. Our instructor has clear expectations and grading is fair (for example, we know when assignments are due, we know how we are doing in the class, test questions are relevant, grading criteria are clear [*Always Frequently Sometimes Rarely Never*]

TEXT BOX for comments

# References

Ambady, N., & Rosenthal, R. (1993). Half a minute: Predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness. *Journal of Personality and Social Psychology*, *64*(3), 431–441. https://doi.org/10.1037/0022-3514.64.3.431

Bavishi, A., Madera, J. M., & Hebl, M. R. (2010). The effect of professor ethnicity and gender on student evaluations: Judged before met. *Journal of Diversity in Higher Education*, *3*(4), 245–256. https://doi.org/10.1037/a0020763

Bennett, S. K. (1982). Student perceptions of and expectations for male and female instructors: Evidence relating to the question of gender bias in teaching evaluation. *Journal of Educational Psychology*, *74*(2), 170. https://doi.org/10.1037//0022-0663.74.2.170

Boring, A. (2017). Gender biases in student evaluations of teaching. *Journal of Public Economics*, *145*, 27–41. https://doi.org/10.1016/j.jpubeco.2016.11.006

Boysen, G. A., Richmond, A. S., & Gurung, R. A. R. (2015). Model teaching criteria for psychology: Initial documentation of teachers’ self-reported competency. *Scholarship of Teaching and Learning in Psychology*, *1*(1), 48–59. https://doi.org/10.1037/stl0000023

Carpenter, S. K., Witherby, A. E., & Tauber, S. K. (2020). On students’ (mis)judgments of learning and teaching effectiveness. *Journal of Applied Research in Memory and Cognition*, *9*(2), 137–151. https://doi.org/10.1016/j.jarmac.2019.12.009

Keeley, J., Smith, D., & Buskist, W. (2006). The Teacher Behaviors Checklist: Factor analysis of its utility for evaluating teaching. *Teaching of Psychology*, *33*(2), 129–131. https://doi.org/10.1207/s15328023top3302\_1

Kite, M. E. (2012). *Effective evaluation of teaching: A guide for facaulty and administrators*. http://teachpsych.org/ebooks/asle2014/index.php

Lind, E. A., Kanfer, R., & Earley, P. C. (1990). Voice, control, and procedural justice: Instrumental and noninstrumental concerns in fairness judgments. *Journal of Personality and Social Psychology*, *59*(5), 952–959. https://doi.org/10.1037/0022-3514.59.5.952

Liu, S., Keeley, J., & Buskist, W. (2015). Chinese college students perceptions of characteristics of excellent teachers. *Teaching of Psychology*, *42*(1), 83–86. https://doi.org/10.1177/0098628314562684

MacNell, L., Driscoll, A., & Hunt, A. N. (2014). What’s in a Name: Exposing Gender Bias in Student Ratings of Teaching. *Innovative Higher Education*, *40*(4), 291–303. https://doi.org/10.1007/s10755-014-9313-4

Markus, H. R., & Conner, A. (2013). *Clash! 8 cultural conflicts that make us who we are*. Hudson Street Press.

Mitchell, K. M. W., & Martin, J. (2018). Gender Bias in Student Evaluations. *PS: Political Science & Politics*, *2015*, 1–5. https://doi.org/10.1017/S104909651800001X

Oppenheimer, D. M., & Hargis, M. B. (2020). If Teaching Evaluations Don’t Measure Learning, What Do They Do? *Journal of Applied Research in Memory and Cognition*, *9*(2), 170–174. https://doi.org/10.1016/j.jarmac.2020.03.001

Stephens, N. M., Fryberg, S. A., Markus, H. R., Johnson, C. S., & Covarrubias, R. (2012). Unseen disadvantage: How American universities’ focus on independence undermines the academic performance of first-generation college students. *Journal of Personality and Social Psychology*, *102*(6), 1178–1197. https://doi.org/10.1037/a0027143

Uttl, B., White, C. A., & Gonzalez, D. W. (2017). Meta-analysis of faculty’s teaching effectiveness: Student evaluation of teaching ratings and student learning are not related. *Studies in Educational Evaluation*, *54*, 22–42. https://doi.org/10.1016/j.stueduc.2016.08.007