

Note: This motion has been revised slightly since the draft dated March 3, 2025 presented at the March Senate meeting, to incorporate feedback provided by the College P&B Executive Committee. The chairs were very positive about the motion ("First and most importantly, most chairs agree that the proposed changes would be a big improvement!"), but recommended small changes in wording to questions #5 and #9 in the college-wide questionnaires. The Senate Committee on Teaching Excellence and Evaluation accepts these recommendations and they are incorporated here, with the deleted language in strikethrough.

Motion to Revise Queens College's Student Course Evaluations March 31, 2025

Whereas Queens College has not updated its student evaluations of teaching in seventeen years and new evidence has emerged since then on teaching evaluation best practices, and

Whereas the Senate Committee on Teaching Excellence and Evaluation has voted unanimously to support recommendations developed by a working group guided by evidence on best practices and having full-time faculty members from every college division as well as student representation,

Be it resolved that the Academic Senate approves the adoption of:

- A. The teaching evaluation questions in Appendix 1 for all courses other than asynchronous online ones
- B. The teaching evaluation questions in Appendix 2 for all asynchronous online courses
- C. The additional three questions in Appendix 3 to be added to the course evaluations for all writing-intensive ("W") courses
- D. Giving each instructor the option of adding up to three additional course evaluation questions selected from an "Add-On Question Bank" maintained by the Office of Institutional Effectiveness, initially including the questions in Appendix 4.
- E. Within the first two weeks of the semester following the approval of this motion, OIE will invite all department chairs and interdisciplinary program directors to submit up to three additional questions to be added to the Add-On Question Bank.
- F. Within the academic year following the implementation of changes, the Senate Committee on Teaching Excellence and Evaluation shall review the new course evaluation data and process to assess their effectiveness, identify unforeseen problems, and recommend further revisions it judges advisable.

Justification

This motion aims to provide more useful and less biased course evaluation data to instructors, chairs, and faculty committees making promotion decisions, and to students making enrollment choices. Extensive research on student evaluations of teaching published since the college last updated its course evaluation questionnaire in 2008 suggests that a reconsideration and revision is overdue. In particular, the literature recommends abandoning "overall evaluation of the instructor" and "overall evaluation of the course" questions, responses to which do not correlate with student learning and are prone to counter-productive and unfair biases. Most other colleges and universities that have revised their SETs in the past decade have dropped these questions and made significant other improvements along the lines of those proposed here. We believe the revised questionnaire will make the college's faculty evaluation and

promotion process more equitable by replacing the questions shown to yield more biased results.

The recommended changes proposed in this motion seek to improve course evaluations by:

- 1. Including in the college-wide questionnaire questions soliciting feedback relevant to all or most disciplines, without exceeding 14 questions.*
- 2. Asking students about objectively observable effective teaching practices or other predictors of student learning.*
- 3. Minimize biases stemming from implicit prejudices and other feelings unrelated to student learning.*
- 4. Asking students questions likely to yield actionable feedback.*
- 5. Asking students questions that they will find easily interpretable.*
- 6. Asking students about course or teaching qualities that cannot be observed in course syllabi.*
- 7. Asking students at least one question assessing a culture of inclusion*
- 8. Asking open-ended questions inviting students to mention course highlights or problems, which a limited number of closed-ended ratings questions cannot cover.*
- 9. Enabling instructors to solicit students' feedback on learning objectives, course requirements, and teaching methods that are specific to individual courses or departments.*

The motion does not contemplate any changes to the current process and policies on reporting SET data (e.g., statistics released only after the grade deadline, written comments would continue to be made available only to instructors, etc.). However, the Committee on Teaching Excellence and Evaluation is considering additional steps to provide guidance to students and faculty on the conduct and use of SETs, including steps to increase response rates.

A more lengthy and detailed justification for this motion, along with background on the process leading to it, is provided in Appendix 6. In addition, this proposal has been approved without dissent by the Senate Committee on Teaching Excellence and Evaluation and an ad hoc faculty-student working group that met monthly in Fall 2024, including students Esther Yee (Math) and Allison Bandura (Psychology) and full-time faculty representatives from each of the college's divisions and having expertise in instruction, assessment, or survey research: Sara Alvarez (English), Anisha Clarke (Math), Anastasiya Lipnevich (Education), Soniya Munshi (Urban Studies and CETLL Director), Kristine Rosales (Sociology), and Peter Liberman (Political Science). The ad hoc working group also includes Lizandra Friedland (Philosophy and OIE) and consulted with Dean Savage (Emeritus Professor of Sociology and former Senate chair), Rebekah Chow (Associate Provost of Institutional Effectiveness), and Nathalia Holtzman (Biology and Associate Provost for Innovation and Student Success).

Appendix 1. Proposed college-wide course evaluation questionnaire for all in-person courses

Selected response items (except as noted, response options are: strongly disagree, somewhat disagree, neither agree nor disagree, somewhat agree, strongly agree, Don't know/Not applicable)

1. The course is organized such that it is easy to understand how topics, assignments, and activities fit together.
2. The instructor followed the course syllabus or clearly explained any changes.
3. The graded assessments (exams, papers, projects, etc.) directly reflect course topics and assignments (lectures, readings, exercises, etc.).
4. I received guidance on how to do well on graded assessments (such as course grading criteria, study guides, rubrics, etc.)
5. I received regular feedback about how to best meet my instructor's expectations (such as comments, exam reviews, informal feedback inside/outside of class).
6. I had opportunities to be actively engaged in this class (through interactive lectures, exercises, discussions, group work, peer reviews, hands on experiences, collaborative projects or other interactive classroom activities).
7. I felt welcome in this class.
8. There are multiple open channels for communication (such as office hours, email, or discussion forums) where students can ask questions and seek help.
9. How would you rate the level of challenge in this course (e.g., topics or materials, assignments, activities, etc.)? [response options: Much too difficult; A bit too difficult; About right; A bit too easy; Much too easy]
10. In an average week, how many hours per week did you spend on this course (or section), including: attending class, doing homework, attending rehearsals, doing readings, reviewing notes, writing papers, attending study groups, doing lab work (unless the lab is a separate section), and any other course related work? [SCALE: 0-20 HOURS PER WEEK]
11. What aspects of the course enhanced your learning the most? [open-ended]
12. What aspects of the course could be improved to better support students to succeed? [select all that apply; open-ended option]
 - a. clarity of instructions for assignments
 - b. helpfulness of feedback on my work
 - c. time spent on clearing up points of confusion
 - d. accessibility of course materials
 - e. relevance of course materials
 - f. frequency of updates on my performance
 - g. timeliness of responses to student inquiries
 - h. timeliness of notices of deadlines
 - i. Other (please specify)
13. Is there anything else that you would like to share about this course? [open-ended]

Appendix 2. Proposed online asynchronous college-wide course evaluation questionnaire

This version substitutes questions 6 and 7 on accessibility and variety of course materials (marked with “*” below) for the questions on active learning and inclusion in the college-wide questionnaire.

Selected response items (except as noted, response options are: strongly disagree, somewhat disagree, neither agree nor disagree, somewhat agree, strongly agree, DK/NA)

1. The course is organized such that it is easy to understand how topics, assignments, and activities fit together.
2. The instructor followed the course syllabus, or clearly explained any changes.
3. The graded assessments (exams, papers, projects, etc.) directly reflect course topics and assignments (lectures, readings, exercises, etc.)
4. I received guidance on how to do well on graded assessments (such as course grading criteria, study guides, rubrics, etc.)
5. I received regular feedback about how to best meet my instructor’s expectations (such as comments, exam reviews, informal feedback inside/outside of class).
6. Course materials are easy to find and navigate.*
7. Course materials are available in a variety of formats (e.g., lecture, text, graphics, video, etc.).*
8. There are multiple open channels for communication (such as office hours, email, or discussion forums) where students can ask questions and seek help.
9. How would you rate the level of challenge in this course ~~How challenging was this course for you~~ (e.g., topics or materials, assignments, activities, etc.)? [response options: Much too difficult; A bit too difficult; About right; A bit too easy; Much too easy]
10. In an average week, how many hours per week did you spend on this course (or section), including: attending class, doing homework, attending rehearsals, doing readings, reviewing notes, writing papers, attending study groups, doing lab work (unless the lab is a separate section), and any other course related work? [SCALE: 0-20 HOURS PER WEEK]
11. What aspects of the course enhanced your learning the most? [open-ended]
12. What aspects of the course could be improved to better support students to succeed? [select all that apply; open-ended option]
 - a. clarity of instructions for assignments
 - b. helpfulness of feedback on my work
 - c. time spent on clearing up points of confusion
 - d. accessibility of course materials
 - e. relevance of course materials
 - f. frequency of updates on my performance
 - g. timeliness of responses to student inquiries
 - h. timeliness of notices of deadlines
 - i. Other (please specify)
13. Is there anything else that you would like to share about this course? [open-ended]

Appendix 3. Questions to be added to all writing intensive (W) courses

Selected response items (strongly disagree, somewhat disagree, neither agree nor disagree, somewhat agree, strongly agree, Don't know/Not applicable):

1. The instructor devoted class time to writing instruction (such as revision strategies, disciplinary concerns, proper attribution, the writing process, research, rhetorical strategies, or writing in online environments).
2. Writing assignments were broken up into stages (e.g., prewriting, outlining, drafting, etc.).
3. I had opportunities to revise and improve my writing.

Appendix 4. Add-On Question Bank—Initial Set (2/4/25)

This set of questions includes all the desired course-evaluations questions solicited from chairs and program directors in an October 2024 survey. Unless otherwise noted, response options are: strongly disagree, somewhat disagree, neither agree nor disagree, somewhat agree, strongly agree, Don't know/Not applicable)

STUDENT EFFORT AND INTEREST	REASON FOR TAKING	Which best describes the reason why you took this course? [response options below] - To satisfy a requirement for my major or minor - To satisfy a General Education (Pathways) requirement - To explore a new topic or skill - To learn more about a subject that interests me - I needed a course that fits my schedule
	STUDENT EFFORT	How would you rate the level of effort you put into this course? [1=Very low, 5=Very high]
	SELF-REPORTED PERFORMANCE	How much progress did you make towards the learning goals of this course? [response options below] - No progress - Little progress - Some progress - Significant progress - Complete progress
STUDENT FEEDBACK FOR NEW COURSES	SUCCESS CHALLENGES	What aspects of the course did you find most challenging, and why? [open-ended]
	NEEDED INFORMATION	Is there anything you wish you had known about the course at the beginning of the term? [open-ended]
EQUITABLE AND INCLUSIVE PRACTICES	UNIVERSAL DESIGN* (included for OA Courses)	Course materials are available in a variety of formats (e.g., lecture, text, graphics, video, etc.).
	EQUITY	I had the same opportunities to do well in this class as my peers.
	FAIRNESS	The methods for evaluating my work were fair.
	ACCESSIBILITY* (included for OA Courses)	Course materials are easy to find and navigate.
	PEER-BASED LEARNING	This course required regular interaction with my classmates (such as group work, discussions, peer reviews, or collaborative projects).
	CULTURAL RESPONSIVENESS	The instructor related the course content to my interests, identities, or culture.
ABILITY TO PLAN AHEAD	ADVANCE DEADLINES	Deadlines for major graded assessments (exams, papers, projects, etc.) were stated in advance.
	PERFORMANCE UPDATES	I was regularly informed about how I was performing in this class (such as test or assignment scores, evaluations, current course averages, etc.).
DEEP LEARNING	PROGRESSION	I feel prepared for more advanced topics in this subject because of my experience in this course.

	CAPSTONE EXPERIENCE	I worked on a final project, paper, presentation, assignment, or exam that allowed me to synthesize what I learned in a comprehensive way.
PERSONAL AND PROFESSIONAL DEVELOPMENT	ETHICAL UNDERSTANDING	I explored ethical issues or societal challenges in ways that deepened my understanding of my own moral perspectives.
	CROSS-CULTURAL UNDERSTANDING	This course encouraged empathetic understandings of cross-cultural differences.
	SELF-KNOWLEDGE	I reflected on my own personal attitudes, values, beliefs, or life goals.
	MENTORSHIP	I met regularly with the instructor to discuss my progress, challenges, and goals.
	PERSONAL RELEVANCE	What I learned in this course is relevant to my academic or career goals.
	SELF-EFFICACY	I can see myself as a successful professional in this field.
	REAL-WORLD APPLICATION	This course provided opportunities to connect theory to practice in real-world situations.
	COMMUNICATION	I had opportunities to develop my ability to communicate effectively.
	TEAMWORK	I had opportunities to work collaboratively with others towards a shared goal.
	LEADERSHIP	I led a discussion, task, or team-based project that helped me develop my leadership skills.
CLOSE READING	ACTIVE READING	Reading assignments included marking our texts with notes or observations.
	LITERARY ANALYSIS	Discussion of texts included the unpacking of language, tone, or structure.
	SLOW READING	Assignments included reading a text slowly, or multiple times, to uncover layers of meaning.
HUMANITIES AND SOCIAL SCIENCES	CRITICAL ANALYSIS	Assignments encouraged the critical analysis of texts, events, or concepts.
	CONTEXT	Assignments helped me understand the cultural or historical context of course topics.
	DIVERSE PERSPECTIVES	The course integrated diverse perspectives and viewpoints into discussions or assignments.
	DISCUSSION AND DEBATE	The instructor facilitated active participation in discussions or debates.
	RELEVANCE OF READINGS	The reading assignments helped deepen my understanding of the subject matter.
	CIVIC LEARNING	I gained a deeper understanding of social interdependence, civic purpose, or responsible citizenship.
	PRIMARY RESEARCH	I had hands-on opportunities to collect, analyze, and interpret data.
	PERSONAL RESEARCH QUESTIONS	I developed and pursued my own research questions, gaining a better understanding of my own goals as a researcher or scholar.
ARTISTIC DEVELOPMENT	ARTISTIC TECHNIQUE-THEORY	This course had a good balance between technique/skill-building and creative exploration.
	ARTISTIC GROWTH	The course encourages creativity or experimentation that supports artistic growth.

	ARTISTIC RANGE	I had opportunities to explore a range of styles, techniques, or traditions.
	ARTISTIC RISK-TAKING	My experiences in this course encouraged me to push the boundaries of my artistic practice, or to challenge existing norms within the field.
	CAREER READINESS in ARTS	I learned about industry practices that will help me prepare for an artistic career (e.g., preparing a portfolio, preparing for a live performance or showing, etc.)
	ARTISTIC SHOWCASING	The course provided opportunities to present or showcase my work (e.g., exhibitions, performances, recitals)
CRITICAL THINKING	INQUIRY-BASED LEARNING	This course explored open-ended problems, issues, or questions without straightforward answers.
	WEIGHING EVIDENCE	I had opportunities to distinguish between strong and weak evidence.
	ARGUMENT ANALYSIS	Assignments included identifying the reasons and conclusions of arguments.
	ARGUMENT EVALUATION	I had opportunities to distinguish between good reasoning and poor reasoning.
	VALID REASONING	I developed my ability to justify my own conclusions with reason and evidence.
	ACTIVE REFLECTION	I had opportunities to actively challenge my own ideas and assumptions.
	MODELED ANALYSIS	The instructor modeled the critical thinking process by presenting analyses step by step.
	INDEPENDENT THINKING	This course encouraged free, independent thinking.
STEM and SKILL-BASED LEARNING	STUDENT READINESS	My prior learning experiences prepared me to meet the requirements in this course.
	COMPLEX TOPICS	The instructor explained challenging concepts or methods by breaking them down into more manageable parts.
	PROBLEM SOLVING	The instructor demonstrated how to systematically approach challenging problems step by step.
	APPLICATIONS OPPORTUNITIES	The course provided ample opportunities to practice and develop needed skills.
	SCAFFOLDING	Assignments build on one another to develop skills incrementally over time.
	INQUIRY-BASED LEARNING	I had opportunities to explore scientific questions through experimentation.
LANGUAGE LEARNING	FLUENCY	The course emphasized practice in communication to foster language fluency.
	CULTURAL NORMS	I learned about important cultural norms alongside new vocabular and grammar.
	LOSS AND GAINS	I had opportunities to consider what is lost and gained in translations among languages.
	IMMERSIVE EXPOSURE	I had opportunities to become immersed in a new language.
EDUCATION	INSTRUCTIONAL STRATEGIES	I had opportunities to apply evidence-based instructional strategies.
	CLASSROOM MANAGEMENT	I had opportunities to apply classroom management techniques.
	MICROTEACHING	I had opportunities to practice delivering lessons.
	TEACHING EFFICACY	I gained greater confidence in my ability to guide students to success.

Appendix 5. 2008 QC course evaluation questionnaire

1. Did you receive a detailed syllabus during the first week of class?
2. How difficult is the course?
3. On average, how much time did you spend per week working on this class outside of the regularly scheduled class time? (Number of hours)
4. Reading assignments were valuable.
5. How many pages of formal writing (essays, lab reports, narratives, term papers, etc.) were assigned?
6. What is your overall evaluation of the course, as distinct from the instructor?
7. The instructor presents the content in a clear and lucid manner.
8. The instructor interacts well with students.
9. The instructor provides useful feedback (e.g., comments on written work and exams, informal feedback inside/outside of class).
10. The instructor returns assignments/exams in a timely fashion.
11. The instructor is available outside of class.
12. What is your overall evaluation of the instructor, as distinct from the course?
13. What did you like most about this class?
14. Please explain why you would or would not recommend this instructor to a friend.
15. What, if anything, could the instructor have done to improve this class?

Appendix 6. Justification and Background for the February 20, 2025 Motion to Revise Queens College's Student Course Evaluations

The extensive research published on student evaluations of teaching (SETs) since the college's system was last updated in 2008 suggests that reconsideration is overdue. Revising the college-wide SET questionnaire and introducing customization would provide more useful and less biased data to instructors, chairs, and committees making promotion decisions, and to students making enrollment choices.

This report provides a detailed justification for these recommended changes after first explaining its origins and authorship.

Background on this report

This report reflects the combined efforts of the Senate Committee on Teaching Excellence and Evaluation and an ad hoc faculty-student working group on course evaluations formed in September 2024. Both groups have approved the motion provided above without dissent. Although responsibility for reviewing and proposing such changes lies with the Committee on Teaching Excellence and Evaluation, it has only four faculty members and lacks representation from all four of the college's divisions. The ad hoc working group read literature on SET best practices, reviewed SETs at other institutions, surveyed department and program chairs about the current SET items they used and changes they would like to see made (response rate=50%), and met monthly over the Fall 2024 semester to discuss criteria for SET design and recommend any desirable changes for the college's SET. The recommendations provided here were drafted by Peter Liberman (Political Science and Chair of the Committee on Teaching Excellence and Evaluation), with revisions suggested by the ad hoc working group and other members of the Committee. This motion and justification were approved by the Committee and reviewed by its ex-officio members Rebekah Chow (Associate Provost of Institutional Effectiveness), Nathalia Holtzman (Biology and Associate Provost for Innovation and Student Success) at the Committee's February 14th 2025 meeting.

The motion and this report thus represent the recommendations of the following faculty members and students of the Committee on Teaching Excellence and Evaluation (indicated by *) and the members of the Ad Hoc Working Group:

- Mohamed Aljami, Student*
- Sara Alvarez, English
- Allison Bandura, Student
- Anisha Clarke, Math
- Lizandra Friedland, Associate Director of Survey Research and Assessment, Office of Institutional Effectiveness
- Omar Ibrahim, Student
- Peter Liberman, Political Science*
- Anastasiya Lipnevich, Education

- Soniya Munshi, Urban Studies and Director of the Center for Excellence in Teaching, Learning & Leadership (CETLL)
- Kristine Rosales, Sociology
- Esther Yee, Student
- Ex officio members of the Committee on Teaching Excellence and Evaluation:
 - Rebekah Chow, Associate Provost of Institutional Effectiveness
 - Nathalia Holtzman Associate Provost for Innovation and Student Success
 - Soniya Munshi, Director of CETLL

In preparing these recommendations and report, we consulted with and benefited greatly from discussions with Dean Savage (Emeritus Professor of Sociology and former chair of the Senate), who strongly supports the recommendations. We also consulted with Amy Wan (English and Special Assistant to the Provost on Writing), who proposed the “W” SET questions, the College’s Inclusive Excellence Committee, and—as noted, by means of a survey—department and program chairs.

Justification for the recommended changes to QC’s course evaluations

A considerable body of research on SETs has appeared since QC’s SET was last updated, in 2008.¹ From this literature, several best practices emerge for improving course evaluations:

1. Including in the college-wide questionnaire questions soliciting feedback relevant to all or most disciplines, without exceeding 14 questions.
2. Asking students about objectively observable effective teaching practices or other predictors of student learning.
3. Minimize biases stemming from implicit prejudices and other feelings unrelated to student learning.
4. Asking students questions likely to yield actionable feedback.
5. Asking students questions that they will find easily interpretable.
6. Asking students about course or teaching qualities that cannot be observed in course syllabi.
7. Asking students at least one question assessing a culture of inclusion
8. Asking students open-ended questions inviting students to mention course highlights or problems that a limited number of closed-ended ratings questions cannot cover.
9. Provide students with the response options for each selected-response question, rather than combining questions with repeated response options into question/response matrices.

¹ Recent reviews include Medina, M. S., Smith, W. T., Kolluru, S., Sheaffer, E. A., & DiVall, M. (2019). A Review of Strategies for Designing, Administering, and Using Student Ratings of Instruction. *Am J Pharm Educ*, 83(5), 7177. doi:10.5688/ajpe7177; Spooren, P., Brockx, B., & Mortelmans, D. (2013). On the Validity of Student Evaluation of Teaching. *Review of Educational Research*, 83(4), 598-642. doi:10.3102/0034654313496870

10. Enabling instructors to solicit students' feedback on the learning objectives, course requirements, modalities, and teaching methods specific to individual courses or departments.

Below we explain in more depth the reasons for each of these criteria that guided the proposed motion.

College-wide SET questions should solicit information relevant to all or most disciplines, without exceeding 14 questions total

There are limits to how much time colleges can expect students to spend completing SETs. Greater length reduces response rates and, relatedly, instructors' willingness to have students complete SETs during class time (the best way to increase completions). The best practice according to a variety of sources is a questionnaire with 10-20 rating items and at least one open-ended written response items.²

Given the greater time needed to write responses to open-ended items and items asking students to select from a menu of very different options, two rating questions should be dropped for the addition of each of these more time-consuming questions to maintain the desired length and response rates.

Given the advantages of including up to six customized ratings questions (three added by each instructor for their own courses plus three added for writing-intensive courses), and three more time-consuming questions (two open-ended and one "select all that apply"), the college-wide questionnaire should include *at most* 12 course rating items. Most college-wide SETs we have collected range from 8-12 questions total. Thus, it would be inadvisable to include more than 12-14 college-wide questions total, including the 3 more time-consuming, open-ended ones.

Given these constraints, the college-wide questions should minimize redundancy and maximize breadth of coverage of observable practices and predictors of effective teaching. If the length of time the college could convince students to spend on each course evaluation was much greater, we could ask multiple questions about each desirable feature of a course or instructor and obtain more fine-grained and reliable measures of student perceptions. But this is simply not feasible. Thus, it makes sense to include in the set of college-wide items questions that address course structure (i.e., features that help students know what to expect in a course), student engagement (i.e., features known to encourage students to be engaged in learning), student support (i.e., features known to help students succeed), and open-ended student feedback.

Given a limited number of questions that can be included, most or all the college-wide questions should apply widely to all or most disciplines. That would make even more sense if the college allowed individual instructors to ask up to 3 additional questions of

² Medina et al (2019).

students taking their courses. Of course, individual departments remain free to determine for themselves which college-wide SET items are most relevant to their own curricula. A department that judges a particular college-wide item to be inapplicable would be free to ignore it in annual evaluations, promotion decisions, etc.

SET questions should ask students about objectively observable effective teaching practices and predictors of student learning

Research reviews on SETs agree that such questions are preferable to traditional questions asking students to broadly “evaluate” courses and instructors. A primary reason is that *the responses to such questions yields do not correlate with student learning*. The most rigorous studies of this question, based on experiments that investigate student learning from different instructors teaching identical courses, have yielded widely varying results. But the most rigorous meta-analyses (i.e., estimates of average effects of cumulative research findings) have found that students do not learn more from more highly rated instructors.³

Several explanations have been given for this. First, students have difficulty assessing what constitutes effective teaching, and from the inherent difficulty of assessing one’s own learning and level of mastery in a subject. Second, students tend to rate easy courses and instructors more highly than challenging ones but tend to learn less in the former. Third, biases related to gender, age, race, and ethnicity affect responses to broad ratings questions, as discussed below.

Thus, instead of asking broad, ambiguous evaluative questions about the course and the instructor, SET questions should ask students about objectively observable effective teaching practices and predictors of student learning. Moreover, although there is a diversity of opinions about what best supports student learning, there are several practices that common sense dictates are important for student learning, but are not always implemented fully by all instructors, such as:

1. Providing students with guidance on how to do well on graded assessments
2. Aligning graded assessments with course topics and assignments
3. Sticking to the syllabus or explaining any changes
4. Being available to students in office hours and via email or phone
5. Encouraging an inclusive and welcoming class culture
6. Being well-prepared for class
7. Organizing class content and communicating that organization to students

Other, less self-evident practices and factors that support learning have been amply demonstrated in educational research. One of the most universal findings at all levels of education is the importance of frequent testing (“formative assessment”) and feedback on assignments, both of which help students assess their mastery and learn from their

³ 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. doi:10.1016/j.stueduc.2016.08.007

mistakes.⁴ Formative assessment and feedback are particularly beneficial for less privileged students.⁵

There is also abundant research demonstrating that students across the social sciences, humanities, and STEM learn more from classes that include active learning than from those providing just lecture.⁶ Moreover, active learning also disproportionately benefits less-advantaged students, narrowing achievement gaps between privileged and underrepresented students.⁷

The rigor and time spent on a course also strongly predict learning. The largest and most rigorous study to date on undergraduates' development of critical thinking skills found that time spent studying and taking courses that assigned both >40 pages of reading/week and >20 pages of writing/semester were strong predictors of improvement, controlling for individual student background and institution type.⁸ There is likely to be some error in student recollections of their time use, but these errors cancel each other out in overall class averages (otherwise, time spent would not be so highly correlated with student learning). Student-reported time-spent data thus provides useful information for instructors and chairs—especially when considered relative to

⁴ A. A. Lipnevich & J. K. Smith (2018). *The Cambridge handbook of instructional feedback*. New York: Cambridge University Press. Sotola, L. K., & Crede, M. (2020). Regarding Class Quizzes: a Meta-analytic Synthesis of Studies on the Relationship Between Frequent Low-Stakes Testing and Class Performance. *Educational Psychology Review*, 33(2), 407-426. doi:10.1007/s10648-020-09563-9

⁵ Eddy, S. L., & Hogan, K. A. (2014). Getting under the hood: how and for whom does increasing course structure work? *CBE—Life Sciences Education*, 13(3), 453-468. doi:10.1187/cbe.14-03-0050. Paul, G., & Verhulst, S. (2010). Improving the reading comprehension skills of minority adults from educationally disadvantaged backgrounds. *Journal of Adolescent & Adult Literacy*, 54(2), 131-140; Pennebaker, J. W., Gosling, S. D., & Ferrell, J. D. (2013). Daily online testing in large classes: boosting college performance while reducing achievement gaps. *PloS one*, 8(11), e79774. doi:10.1371/journal.pone.0079774

⁶ The effect size is about half of a standard deviation, a large effect, corresponding to the difference between the 50th and 69th percentile in student learning, and failure rates are 50% higher in traditional lecture courses. See Anastassis Kozanitis and Lucian Nenciovici. "Effect of Active Learning Versus Traditional Lecturing on the Learning Achievement of College Students in Humanities and Social Sciences: A Meta-Analysis." *Higher Education* 86, no. 6 (2022): 1377-94; S Freeman et al. (2014) Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the Natl Academy of Sciences (PNAS)* 111:8410–8415.

⁷ E. J. Theobald, M. J. Hill, E. Tran, S. Agrawal, E. N. Arroyo, S. Behling, N. Chambwe, et al. "Active Learning Narrows Achievement Gaps for Underrepresented Students in Undergraduate Science, Technology, Engineering, and Math." *Proceedings of the Natl Academy of Sciences (PNAS)* (Mar 24 2020): 6476-83.

⁸ Arum, R., & Roksa, J. (2011). *Academically adrift: Limited learning on college campuses*. Chicago: University of Chicago Press.

departmental and college means—on whether a course should be made more or less rigorous.⁹

SET questions should minimize biases stemming from implicit prejudices and other feelings unrelated to student learning

Student responses to open-ended course and instructor rating questions have been widely found to be biased by instructors' gender, age, ethnicity, race, accent, volume of homework assigned, and grading scale.¹⁰ Feelings unrelated to student learning appear to shape responses to other questions as well, including open-ended ones. Such biases are very problematic if allowed to result in unfair discrimination in a college's faculty promotion decisions, even indirectly by influencing student course selection (if students choose courses based on biased SET data). To the extent that course easiness/difficulty bias affects enrollments (and indirectly, departmental scheduling and curricular decisions) and promotion, it can indirectly undermine the college's success in promoting student learning.¹¹

Bias is likely to be the strongest in broad evaluative questions, because the complexity and difficulty of such evaluation—especially for those who are not experts in teaching quality or methods and who lack awareness of how much they have learned—leads students to draw on intuitive feelings. However, bias can affect other types of questions as well, including open-ended ones.¹² It is probably impossible to eliminate completely all biases in teaching evaluations, leading some experts to recommend abandoning SETs altogether.¹³

However, questions about straightforwardly observable teaching practices and other predictors of learning in theory should be less prone to bias, and we have not found research showing that they are. Thus, it makes sense to try such questions and then conduct research on accumulated data to evaluate biases in them. In addition, even if

⁹ Although this is measure of student behavior, not of perceived teaching practices, it is actionable because (within limits) instructors can shape the amount of time student spend learning by better motivating students (and relatedly encouraging a growth mindset), by adjusting out-of-class assignments, by altering assessment practices, and by preventing cheating or short-cuts on assessments.

¹⁰ Kreitzer, R. J., & Sweet-Cushman, J. (2021). Evaluating Student Evaluations of Teaching: a Review of Measurement and Equity Bias in SETs and Recommendations for Ethical Reform. *Journal of Academic Ethics*, 20(1), 73-84. doi:10.1007/s10805-021-09400-w; Peterson, D. A. M., Biederman, L. A., Andersen, D., Ditonto, T. M., & Roe, K. (2019). Mitigating gender bias in student evaluations of teaching. *PloS one*, 14(5), e0216241. doi:10.1371/journal.pone.0216241;

¹¹ Stroebe, W. (2023). If Student Evaluations of Teaching Are Invalid, Why Are They Still Being Used? Comments on Uttl (2023). *Human Arenas*, 7(2), 438-445. doi:10.1007/s42087-023-00385-z

¹² <https://www.nytimes.com/roomfordebate/2015/12/16/is-it-fair-to-rate-professors-online/gender-bias-exists-in-professor-evaluations>

¹³ Uttl, B. (2023). Student Evaluation of Teaching (SET): Why the Emperor Has No Clothes and What We Should Do About It. *Human Arenas*, 7(2), 403-437. doi:10.1007/s42087-023-00361-7

biases cannot be eliminated, their negative effects can be mitigated by ensuring that department chairs and promotion committees are aware of them. That is why most SETs include items on course difficulty and time demands, which enable chairs and promotion committees to fairly weigh the typically lower ratings of more demanding instructors.

SET questions should be actionable

Actionable questions provide information enabling instructors to learn how they can improve their teaching and alerting department chairs to significant virtues and deficiencies in teaching quality. They focus on aspects of the course or teaching methods, allowing for targeted changes. Asking students if “the feedback on assignments was timely and constructive” is actionable because it provides information to instructors about the timeliness and constructiveness of the feedback they provide. If students provide consistently low ratings, this informs instructors that they should provide more prompt and/or detailed feedback on student assignments or seek out professional development opportunities to improve their feedback methods. This item also gives chairs information about how well faculty are following a practice important for student learning.

Non-actionable questions, on the other hand, are often too vague or general to provide useful insights. For example, receiving a low rating on the question “What is your overall evaluation of the instructor, as distinct from the course?” does not provide an instructor with specific enough information to improve their teaching.

SET questions should be interpretable by students

Asking interpretable straightforward questions is essential for collecting useful responses in a survey of any kind. Easily interpretable SET questions are straightforward and specific, making it easier for students to provide meaningful feedback. They avoid ambiguity and ensure that students understand exactly what is being asked, leading to more accurate and useful responses. On the other hand, questions that are vague, ambiguous, or complex can confuse students and result in unclear or unhelpful feedback. Examples of overly vague questions are ““What is your overall evaluation of the instructor, as distinct from the course?” and ““What did you think of the course?” An example of an overly complex question would be “Was the course content relevant and interesting, and did it help you achieve your learning goals?”

SET questions should ask students about course or teaching qualities that cannot also be observed in course syllabi

Given the constraints on SET length noted above, SET questions should solicit student feedback on course qualities that are not observable by other means. Conveying to students a course's learning objectives, the topics covered and their organization, major assignments, deadlines, and grading and other course policies, and the organization of

a course's topics all promote student learning. But this information also ought to be stated clearly in the course syllabus, which can be reviewed by department chairs (and the instructors themselves).

The SET questionnaire should include at least one item assessing a culture of inclusion

Queens College is committed to fostering an environment of inclusive excellence. Our success as an institution relies on appreciating, involving, and advancing inclusion and equity among our students, staff, and faculty. This commitment is especially significant to Queens College due to the rich diversity of our student body.

This mission is one of the main priorities established in the College's 2021-2026 Strategic Plan:

QC strives to reflect the great diversity that exists in the Borough of Queens in New York City. QC has a higher proportion of student ethnic diversity than many colleges, currently enrolling a student population that speaks 79 languages and hails from 140 countries. For its increased enrollment of Latinx and Asian students in recent years, QC is designated an Hispanic Serving Institution (HSI) and Asian American Native American Pacific Islander Serving Institution (AANAPISI). Despite this diverse profile, students, faculty, and staff have spoken out about the need for systemic change to tackle lack of representation, racism, and other biases that impact our community. This goal calls upon the entire campus community to participate in culture renovation, an effort to make QC a place where everyone feels welcome, understood, supported, and protected.¹⁴

To achieve this goal, the Plan specifically mandates that steps be taken to “incorporate DEI as an element of assessment frameworks.” That is part of a broader strategy to “build DEI into our campus-wide policies, processes, and interactions. Foundational steps in culture renovation, including intentional and sustained education, are needed to foster an environment where everyone has the awareness, knowledge, and communication skills to participate fully as members of the QC community.”¹⁵

Our recommended changes to the college's SET incorporates inclusivity into student assessment of teaching by adding a question to the in-person SET questionnaire about how welcome they felt in the class and by dropping questions that are particularly prone to implicit biases, with potentially discriminatory consequences for instructors.

SET questions should include open-ended questions inviting students to mention course highlights or problems that a limited number of closed-ended ratings questions cannot cover

¹⁴ Queens College Strategic Plan for 2021–2026, Office of the President, Queens College, 2001, p. 16. (https://www.qc.cuny.edu/ab/wp-content/uploads/sites/43/2022/04/Queens_College_Strategic_Plan_2021_2026.pdf).

¹⁵ Ibid.

Including open-ended, written-response questions in student evaluations of teaching allows students to provide more detailed and nuanced feedback than ratings questions, which limit responses to predefined options. Open-ended questions enable students to elaborate on their experiences, providing richer insights into what worked well and what could be improved. Students can offer specific examples and anecdotes that highlight strengths or areas for improvement and can reveal issues that may not have been addressed by ratings questions. This specificity can help instructors understand the context of the feedback and make more targeted adjustments to their teaching methods and can elicit suggestions that the college or instructor might not have thought to ask students about. Furthermore, when students are asked for their detailed feedback, they are more likely to feel that it is valued, and thus more likely to engage with the evaluation process.

Questions asking students to select “all that apply” from a set of heterogeneous, dichotomous response options can be useful for eliciting students’ feedback on areas of teaching strengths and weaknesses that are not covered in the ratings questions and that students might not think of absent prompting. Such questions take more time to complete than simple ratings questions, but not necessarily as much as open-ended ones, and in some cases could yield more useful information than the latter. Open-ended options also can be usefully added to the response options, to elicit student feedback on ways a course could be improved that are not included in the (necessarily limited) response options.

Course evaluations should allow colleges and instructors to solicit students’ feedback on the learning objectives, course requirements, modalities, and teaching methods specific to individual courses or departments

Many colleges and universities have embraced customization to varying degrees, a step facilitated by the near universal shift from paper to online SETs. Customization allows asking about learning objectives, course requirements, or teaching methods and modalities that are specific to individual departments, teaching styles, or types of courses.

There are some immediately obvious differences between disciplines and modalities that would benefit from customized SETs. Asynchronous online courses rely much more heavily on digitally accessed resources than in-person courses and lack the real-time interpersonal interactions of in-person ones. Thus, for all SETs for asynchronous online courses, it would be advisable to replace two items in the college-wide core questionnaire that are mainly relevant to classroom learning with items more relevant to asynchronous online learning (see Appendix 2).

Moreover, there are unique features of writing-intensive (“W”) courses, lab courses, and performance courses that would benefit from student feedback elicited by questions specific to each type of learning experience. Thus, the three additional items (suggested by Amy Wan, Professor of English and Special Assistant to the Provost on

Writing, and provided in Appendix 3) should be included on the SET questionnaire for all “W” courses.

Individual departments and instructors often have their own teaching methods and student learning objectives and may want to collect students’ feedback or perceptions of progress towards some of those. Customizability of SET questionnaires would empower faculty and encourage reflective teaching practices by providing instructors with the flexibility to solicit targeted feedback on specific aspects of their courses—such as close reading, critical reasoning, skill-based learning, or artistic development. Data collected from instructor-added questions should be provided only to the instructor of record for a course.

Customization can be easily achieved using the current Anthology SET platform by inviting individual faculty to add up to three additional questions to the SET questionnaire for each of their courses. These questions would be selected from an “Add-On Question Bank” maintained by OIE and Anthology. Appendix 4 provides items already added to the item bank, many of which were suggested by department chairs and interdisciplinary program directors in response to a Fall 2024 survey of their SET practices and preferences. The items proposed by the chairs and program directors who completed this survey on use of course evaluation data have already been incorporated into the bank.

In addition, before the optional item bank is finalized, all departments will be invited to propose up to three additional questions to be added to the bank. Departments’ proposed questions could be designed to reflect program- or course-specific learning objectives, though adoption of these items will be optional. Departments are encouraged to engage all their faculty in this process. The OIE will provide consultation and feedback to departments to ensure that all questions in the bank are valid, reliable, and aligned with best practices, and to minimize repetition.

Conclusion

The proposed changes to QC’s SETs will greatly increase the utility and the fairness of the data they provide. However, even improved SET data will remain limited as a measure of effective teaching and should be considered as only one of multiple factors in departmental and college evaluation of its instructors’ teaching effectiveness.¹⁶ The college should develop and provide resources to students on the importance of doing SETs and doing them fairly, and information for instructors, chairs, deans, and faculty serving on personnel and promotion committees on appropriate use of SET data. Finally, the Committee on Teaching Excellence and Evaluation, OIE, and CETLL

¹⁶ American Sociological Association, *Statement on Student Evaluations of Teaching*, Sept 2019; retrieved from https://www.asanet.org/wp-content/uploads/asa_statement_on_student_evaluations_of_teaching_feb132020.pdf. Other important sources of evidence could include peer observations, teaching practices inventories, teaching philosophy statements, responses to student and peer feedback, and participation professional development opportunities.

together should review the new SET data and process two semesters following the implementation of changes to look for any unforeseen problems and opportunities for further improvement.