UC RIVERSIDE

October 3, 2023

To: Sang-Hee Lee, Chair of the Riverside Division

From: Ken Baerenklau, Associate Provost, Professor of Public Policy, and Co-Chair of the TEIC
        *Yingbo Hua, Professor of Electrical & Computer Engineering and Co-Chair of the TEIC
        Batool Abdaljawad, undergraduate student
        *Annie Ditta, Assistant Professor of Teaching - Department of Psychology, ADT member,
            and Co-Chair of the SET question subcommittee
        Richard Edwards, Executive Director of XCITE
        Jack Eichler, Professor of Teaching - Department of Chemistry, and immediate past Chair
            of the Academy of Distinguished Teaching (ADT)
        *Ahmed Eldawy, Associate Professor of Computer Science & Engineering
        *Long Gao, Associate Professor of Operations & Supply Chain Management
        *Jacob Greenstein, Professor of Mathematics
        *Ruhi Khan, Associate Professor of Media & Cultural Studies
        *Goldberry Long, Associate Professor of Teaching - Department of Creative Writing, ADT
            member, and Co-Chair of the SET question subcommittee
        *Morris Maduro, Professor and Chair of Biology and ADT member
        Patriccia Ordonez-Kim, PhD student
        *Amit Roy Chowdhury, Professor and Chair of Robotics
        Omar Safie, Director of Evaluation and Assessment
        *Wesley Sims, Assistant Professor of Education
        Anne Sullivan, University Writing Program Lecturer and Coordinator for Teaching
            Assistant Development Program
        *Elaine Wong, Associate Professor of Management
        denotes Senate-appointed members

Cc: Cherysa Cortez, Executive Director
        Dan Jeske, Vice Provost for Academic Personnel
        Elizabeth Watkins, Provost and Executive Vice Chancellor

Re: Response to feedback on draft work products from the Joint Senate-Administrative
        Teaching Evaluation Implementation Committee (TEIC)

Dear Sang-Hee,

Thank you for your memo dated April 25, 2023 which summarizes the feedback received from
Senate committees on the TEIC draft work products, and for the opportunity to meet with the
Executive Council on June 26, 2023 to discuss the status of our revisions.

As we shared with the Executive Council in June, these were our main take-aways from the Senate committee feedback:

1. *Generally supportive but with critical comments mostly about the SET instrument and especially question wording.*
2. *Concerns about increasing workload for students, faculty, and file-preparers.*
3. *Concerns about whether 10 weeks is too short for 2 surveys.*
4. *Desire for optional questions for instructors to select from.*
5. *Concerns that terms like "useful", "appropriate", and "timely" are subjective.*
6. *Suggestion to pursue a larger pilot test before full campus roll-out.*

Meeting with the Executive Council in June provided helpful additional context and detail, and generally reinforced these key take-aways for us. Since then, we have continued working on our revisions which are now complete. Below we respond to each of the six preceding items.

1. *Generally supportive but with critical comments mostly about the SET instrument and especially question wording.*

   Our question wording subcommittee discussed the comments received in great detail, and developed recommendations which were subsequently presented and discussed at a full TEIC meeting on June 12. The subcommittee made revisions after this discussion, and then another set of revisions following the meeting with the Executive Council on June 26. Details including each issue considered, action taken, and supporting rationales are provided in appendix 1. The revised SET instrument is in appendix 2.

   Of all the work undertaken by the TEIC, the SET questions have been the most discussed and debated. This reflects how much our faculty cares about teaching and how we evaluate it. Wording choices that would go unnoticed in co-authored papers have received deep scrutiny as we have worked to build consensus. Because reasonable people can disagree about these choices, we could continue debating them ad infinitum. In order to put a much-improved SET instrument into practice, and consistent with our charge, the TEIC is ready to move from discussion to implementation. Additional changes, informed by our future experience with these new questions, can of course be made in the future and will be under the Senate's purview to decide.

2. *Concerns about increasing workload for students, faculty, and file-preparers.*

   For students, the concern was expressed that increased survey response workload could further undermine participation and response rates. The TEIC has reduced the number of questions on the instrument (detailed in appendix 1). Though the revised instrument may still look long, the questions and answers are designed to be simple and straightforward which facilitates faster completion of the survey. More than half of the questions are binary response and the text boxes are limited to one per section. The repetitive structure of the other multiple-choice answers also makes the survey easier

to complete. When 700 students piloted the previous, longer, version of the survey, completion time averaged only 7 minutes. The TEIC also believes that students will see greater value in the redesigned instrument which will tend to increase participation. If instructors remain concerned, we recommend allocating a brief amount of class time for students to complete the survey as was done with paper-based evaluations. Also, the Senate could take up a discussion of our proposal to provide an incentive for SET participation by entering students in a lottery for priority registration if they complete their SETs on time.

For faculty and staff, the concern was expressed that the increased availability of additional (optional) forms of evidence beyond the SET instrument will eventually evolve into an expectation that they be submitted in order to earn a positive review, and a corresponding increase in file submission workload. The TEIC notes that the desire to elevate the importance of submitting multiple forms of evidence (a requirement of APM-210), and the desire to make submission easier, was endorsed by the Senate Ad Hoc Committee on Evaluation of Teaching. One rationale for doing so is to provide a diverse set of tools for instructors to demonstrate evidence of teaching effectiveness, given our diversity of instructors and classroom environments, which helps to mitigate any remaining bias in the SET responses. The TEIC also believes that this workload concern derives from a possible outcome that the campus can avoid if it wants to. Although it is beyond the charge of the TEIC to address this concern more directly, we will work to educate all instructors that only two forms of evidence are required.

3. *Concerns about whether 10 weeks is too short for 2 surveys.*

   The TEIC agrees that the quarter system may be less accommodating for a two-step approach, and believes that the first "foundations check" survey should be optional. We will communicate this to all instructors, along with the benefits of distributing this survey if they choose to do so. The same "foundations check" questions will remain on the second "end-of-quarter" survey which is the main survey.

4. *Desire for optional questions for instructors to select from.*

   The TEIC agrees that it should be possible to add optional questions to the SET. We had a lengthy discussion about how to do this. Some members emphasized maintaining the same criteria for optional questions as were used to develop the required questions in the campus-wide SET instrument (these criteria are listed in appendix 3). Other members emphasized offering greater flexibility to instructors. A key disagreement arose over whether an "overall experience" question (e.g., "Was the instructor helpful overall for your learning?") should be allowed as an optional question. Such a question would not meet the criteria, but some members felt that it nonetheless has value. By a vote of 12-1-2 (for-against-abstain, with 3 members not voting) the committee approved a compromise whereby the SET will have three question categories:

i. **Campus-Wide required questions:** These are the questions listed in appendix 1. By design, they meet the question criteria. Responses will be captured in the auto-generated report that is uploaded to eFile.

ii. **Campus-Wide optional questions:** These also must meet the question criteria. Instructors may select a small number of these (perhaps from a pre-populated drop-down menu) to include in their SET, if desired. Responses are available to the instructor and may be included in the auto-generated eFile report.

iii. **Instructor-Discretionary questions:** These are optional and do <u>not</u> have to meet the question criteria. Instructors may add a small number of these directly to their SET (as text entries), at their discretion. Responses will be available to the instructor but <u>cannot</u> be included in the auto-generated eFile report. However, instructors may choose to present the results in their self-statement.

Currently there are no category (ii) questions. Going forward, instructors may propose questions to add to this category. We believe that any changes/additions to categories (i) or (ii) should be vetted by a committee with oversight responsibility, and should adhere to the same criteria used to develop the required questions (more details are in appendix 3). This approach will help to maintain the integrity and consistency of the SET, and aligns well with the shared governance recommendations in our January 9, 2023 memo to the Senate. Depending on if/when the Senate chooses to act on those recommendations, the TEIC could vet the initial set of proposals for optional questions before the Senate takes over this responsibility.

Related to the topic of optional questions, committee members also disagreed on whether departments should be able to endorse questions that would be included in the eFile report. The co-chairs felt that this issue of "departmental sovereignty" in the teaching evaluation process is both weighty and beyond the scope of our charge, so we did not address it. We expect this issue will be brought to the Senate in the future.

5. *Concerns that terms like "useful", "appropriate", and "timely" are subjective.*

The question wording subcommittee discussed the use of terms like this at length. Some changes were made, and more details are provided in appendix 1. It is worth noting that it is not possible to remove all subjectivity from a survey that elicits responses about personal experiences. Moreover, it is worth understanding those personal experiences even if an instructor might disagree with them because it reveals opportunities for improvement – even if the improvement is just clearer communication with students. The TEIC discussed how aversion to subjectivity may derive from instructors' concerns about low scores and how they might affect the merit/promotion process. Because the new SET is designed to provide instructors with actionable feedback for improving their teaching, we urge everyone involved in the review process to place more emphasis on how an instructor responds to low scores than on the low scores themselves. Repeatedly low scores are certainly worrisome, but severe consequences for any low

scores will have a chilling effect on instructional innovation and student learning.

6. *Suggestion to pursue a larger pilot test before full campus roll-out.*

   The TEIC agrees that an expanded pilot will be beneficial. We originally planned to conduct this pilot in fall 2023, but ITS has experienced delays in procuring the replacement technology for iEval. At this time, we anticipate conducting the pilot in winter 2024 concurrent with the planned ITS pilot of the new technology. If successful, the full campus roll-out would happen no earlier than spring 2024. Among other benefits, this will provide additional time for educating and socializing students and instructors about these new systems.

Table of Changes in Response to Feedback		
Comments on Overall Structure and Content		
Suggestion or Feedback	**Action Taken**	**Committee notes**

NOTE: in this table, the two parts of each question are referred to as **Tier A** and **Tier B.** In the instrument, the students will see the two tiers as #a and #b. For example, 1a and 1b.
 	
1. Each question has two parts. This two-tier question format generated the question, "How should the instructor interpret TIER A and TIER B to improve their teaching?"
	• See Committee notes	• **What is the purpose of the instrument?**
		  ○ To give instructors access to student-reported experiences of teaching and learning.
		  ○ To meet the recommendations of both UCOP and the UCR ad-hoc committee that any SET instrument should **avoid averaging qualitative data and take care to avoid bias.**
		• **Why does the instrument have this structure?** This instrument meets its purposes by accessing student reports of their experiences of evidence-based, well-established best practices in teaching.
		  ○ <u>Tier A</u> = **Frequency of best teaching practice**—did students experience it, and if so, how often? Tier A is meant to help students identify and focus on an observable best practice, inviting their memory and reflection on the class before they evaluate efficacy in Tier B.
		  ○ <u>Tier B</u> = **Effect of instructor's use of Best teaching practice (BTP)**– did the students perceive the instructor's use of the BTP as contributing to their learning?
		  ○ <u>The comment box</u> then asks students to explain why the perceived a best practice as helping or not helping. Thus, open comments are invited in a meaningful context after students have remembered and reflected.
		• **How will results be reported?** This will depend on the platform ultimately chosen, but the intention is:
		  ○ All data will be grouped by domains: A) Course Foundations, B) Course Experiences, C) Graded Work.
		  ○ Within each domain, each question will include a histogram of answers A and B.
		  ○ Comments will also be grouped by domain, allowing instructors to **read the TIER A/TIER B results in the context of student explanations**.
		• **How should results be interpreted?** All three components of the answer to a question should be read together.
		  ○ <u>Results of Tier A should not be read as a ranked result</u>. "Always" does not mean "Best." A student answering "Always" to a Tier A question is not equivalent to an instructor receiving a high score.
		  ○ Thus, Tier A is only meaningful when the instructor reads the pairing of frequency (Tier A) and quality (Tier B) alongside the students' comments at the end of the domain.
		  ○ **For example**, an instructor might always implement a best practice, but do so unsuccessfully, resulting in a student recognition that the practice was present (Tier A), but a perception that did not contribute to their learning (Tier B).
		  ○ An answer to TIER B of "Probably didn't" or "Definitely not" might motivate instructors to improve their **use** of the best practice. To better understand this result, the instructor should also read the comments for that domain. (A second example is provided in Question 5).
		• **Does this instrument improve teaching?** Not on its own.
		  ○ For this instrument to improve teaching, students would need to be experts in pedagogy, which they are not.
		  ○ Rather, it is an instrument for accessing students' *experiences* of teaching, which might be a first step for instructors who wish to seek out resources and training.
		  ○ The instrument cannot prescribe a solution to any problems detected because there are many ways to solve the problem. There are many resources on campus to help instructors improve their implementation of best practices, including the Academy of Distinguished Teaching, which offers classroom observation, and XCITE.
		  ○ We agree that this instrument should not be the only measure of quality of teaching. This is why the APM requires 2 or more measures.

2. Concern that the instrument's mentions of pedagogical terms infringe on academic freedom.	• See Committee notes	• Nothing in this SET should be perceived as a mandate; we are giving students an opportunity to comment on whether they observed evidence-based, well-established best practices in teaching. • We relied on research to support each question.
3. Some felt the instrument was too long.	• We deleted an entire domain, "Learning Support," and 1 question in "Graded Work."	• We will continue to test time to completion, which now appears to be 2-3 minutes, though we have only piloted this draft of the proposed instrument on faculty. • The prior draft of the proposed instrument was tested on 700 students; time to completion averaged 7 minutes.
4. **For TIER B:** The original wording: "Was this useful for your learning?" generated some concern that "useful" may vary by cultural context; students may have varying understanding of the question.	• Changed question to "Did the instructor's use of this practice contribute to your learning?"	• We also discussed "support your learning," but decided "support" implies "under and separate," while "contribute," emphasizes an integration of teacher effort and student effort. • We considered "did this contribute to your achievement of course goals," but we did not want to link the question too tightly to SLOs. • Our goal was to frame the question so the student's shared responsibility for their learning was embedded in the question (i.e., not all burden of learning is placed on the instructor).

5. Many advocated for use of a 5-point Likert scale in lieu of the proposed 3.	**Scale was changed.**
	Tier A questions "Did the instructor…"
	• Always,
	• Most of the time
	• About half the time
	• Rarely
	• Never
	• Not sure
	
	Tier B questions "Did the instructor's use of this practice contribute to your learning?"
	
	• Definitely yes
	• Probably did
	• I'm neutral
	• Probably didn't
	• Definitely not
	• N/A (not applicable)
		<u>Tier A</u>: We chose a scale that emphasize *frequency* of the instructor's use of a teaching practice.
		  • We considered the following other options:
		    ○ Definitely true, probably true, neither true nor false, probably false, definitely false
		    ○ Definitely yes, probably yes, might or might not, probably not, definitely not
		    ○ A great deal, a lot, a moderate amount, a little, none at all
		  • We included **"Not Sure,"** because students may be uncertain about **how often** a practice occurred in a course.
		  • "Not sure" acknowledges the students' uncertainty about whether the practice occurred.
		  • We moved "Not sure" to the end because it doesn't fall between the other options.
		  • We did not use N/A, because N/A indicates the student is sure the item is not applicable. This can be captured with "Never."
		  • Although this scale is 5 points, <u>it should not be interpreted as a ranking.</u> It accesses frequency of a best practice, but as noted in item 1 and item 14 of this table, not all best teaching practices need to be always used. Tier A should be read with Tier B and comments.
		  • Note: Tier A allows for instructors to receive credit for their use of a best practice, even if Tier B indicates it was used imperfectly. This report of use of practice can be used by the instructor in their Teaching Statements to discuss their efforts to improve their teaching.
		
		<u>Tier B</u>: We chose a scale that emphasized *degree* of contribution.
		  • Because this is referring to quality, "I'm neutral" is offered for the student who feels the practice neither contributed nor did not contribute.
		  • We included N/A to accommodate students who answered Tier A with "Never" or "Not sure."
		  • Some students who answered Tier A with "Never" or "Not sure" might respond that this didn't contribute to their learning. An instructor could interpret the combined feedback from Tier A and Tier B as "I didn't use this practice, and students perceived this absence as not contributing to their learning, so perhaps I might try using this practice in the future."

6. Suggestion to insert an open comment box after every question	• We kept 1 comment per domain. (There are now 3 domains).	• Our original intent was a comment box for each question. • In our pilot of 700 students, we found that in large classes, a comment box for every question led to 50 pages of comments, rendering the results too cumbersome.
7. Suggestion that departments be allowed to draft and add their own questions.	• We have produced guidelines so that departments can craft questions to submit for approval. (See Appendix).	• This instrument was developed by pedagogy experts after extensive research on the literature of bias in SETs, pedagogy, and instrument design. All questions here are research-based and deliberately crafted to avoid bias. • SETs varying by department can lead to inconsistency across campus, undermining attempts to control for bias. • We advocate for a SETs oversight committee, which could vet questions composed according to the following guidelines: **Guidelines for Questions- They should be:** • **Observable**: can the student see the teaching practice? • **Specific**: does the question encompass only one teaching practice? • **Evidence-based**: does the question describe a best teaching practice? • **Neutral**: Does the question avoid words, especially adjectives, that invite value judgments? • **Clear**: Does the question resist ambiguity to preclude multiple interpretations? • **Accessible**: Does the question avoid jargon that students won't know? One optional question already crafted: Did the instructor create an inclusive classroom environment? For example, did they encourage diverse contributions and/or provide inclusive class experiences?
8. Some asked why there wasn't a broadly overarching question about, for example: • Overall learning • Overall experience • Overall instructional quality	• We did not add this type of general, overarching question.	• A wide body of research indicates that SETs in their current form, which often include questions like these, are vulnerable to bias, especially against women and minoritized people. This was the conclusion of the Academic Council Teaching Evaluation Task Force, and was supported by the Ad Hoc Committee on Evaluation of Teaching that preceded our committee. • General questions in particular tend to be extremely prone to bias, as they are not referencing any instructor action that is observable and specific. Focusing on specific, observable practices mitigates potential bias. • We are measuring *teaching* not learning. **Learning is best assessed by instructors.** We trust instructor's measures of student learning through their carefully designed graded work and instructor evaluations used to assign grades. • The default practice for some evaluators of instructors has been to consider only this question while ignoring other results of the SET. This strategy is prone to bias. It also makes it possible to ignore strengths in one domain and weaknesses in another. Reviewing bodies should work to develop a more nuanced assessment of an instructor's effectiveness instead of relying exclusively on one summative question. • The practices in the instrument are research-based best pedagogical practices. If the practices surveyed are in place, the overall learning and experience will be excellent.
9. Some felt that students were being positioned to be content experts in teaching.	• Discussed at length.	• Students are experts in *their own experience*. This report asks them to report on their perceptions of their experiences. • The instrument is prompting students to evaluate the use of evidence-based practices. The definitions of such practices are provided, and the students don't need any knowledge beyond the provided definitions to answer the questions.

10. Some felt that the instrument was akin to a customer satisfaction survey.	• Discussed at length.	• We are asking for the students' experiences, but we are prompting them with specific best practices that we are defining, so we are moving away from a feelings- or satisfaction-based survey to a report of an observable skill and its effect on them. • This is a reason that we aren't asking, "Overall was this effective," because that question is very open to interpretation and bias. • We agree that this instrument should not be the only tool used to evaluate instruction. It should be used in combination with other measures.
11. Some felt that the students should be asked if the workload was reasonable and if the instructor designed the class appropriately for students' skill level.	• Discussed at length.	• Having high expectations for students being educated at a college-level class is reasonable (also, reasonable is a sliding scale, dependent on student performance). • As with other value-laden words, "reasonable" invites bias as it may vary with a student's performance in the class. A student who did poorly might see the workload as unreasonable, while a student who did well might see it as reasonable. • It is not possible to identify all students' incoming skill, especially in large classes, so classes can't be leveled to each student's ability.
12. Some felt that there should be more fine-grained feedback on specific teaching behaviors, such as use of classroom technology.	• We did not add this.	• We were concerned that such fine-grained scrutiny would impinge on academic freedom.
13. Some felt that students should be asked if grading was fair.	• The question most related to fair grading was changed to "Did the instructor provide information about grading criteria and/or standards for how the work was graded?"	• If the students are unhappy about their grade, they are likely to label the grade as unfair. • The question we propose describes the circumstances that are most likely to foster fair grading.
14. In original draft, Tier B questions used frequency. Some were concerned that students could not determine how frequently a best practice contributed to their learning.	• Changed to • Definitely yes, Probably did, I'm neutral, Probably didn't, Definitely not, N/A	• We agree that **degree of effect** of the practice is more important than frequency of the practice for Tier B. • However, the three components (Tier A, Tier B, and Open Comments) are meant to be read together to achieve meaningful results. (See in this table, p.1, Item 1 for an example).

15. There were some concerns that the wording of the second question, "Did this contribute to your learning" would create an unclear result for instructors because "this" was ambiguous (i.e., What does "this" refer to?)	• Changed to "Did the instructor's use of **this practice** contribute to your learning?"	• In response to one suggestion, we considered, "If the instructor had always used this practice, how much would it have contributed to your learning?" But this asks students to engage in a hypothetical, which breeds inaccuracy and bias. Also, many best practices are very effective even if used only occasionally. • We agree that the clearer definition of "this" offers clarity. • The students have an opportunity to comment on ambiguous results. Thus, all results should be viewed in context with their open-ended comments. This is why there is a comment box at the end of each section. • We considered: "Did the instructor use this practice **well enough** to contribute to your learning?" But the problem is the same as with "fairness" and "reasonable;" the qualifier invites responses that vary with the student's grade and their emotions about their grade.

(Continued)

Appendix 1

Comments on Domain A: Course Foundations Check

16. Comment that early administration of "Foundations Checklists" during Week 2 may not be possible. Some students haven't added until Wk 3.	• The administration of the early-quarter preliminary Foundations Checklist is now **optional.**	• We agree that there may be logistical barriers to implementing a 2-part evaluation. • Some faculty may wish to take advantage of this option to ensure that students are aware of all information early in the course. • The early poll may offer faculty valuable opportunities to correct any oversights or student misunderstandings before the final, required evaluation.
17. **Learning Outcomes** Concern that the inclusion of the term "Learning Outcomes" might confuse students, impinge on academic freedom, and be too abstract for students. Some suggested substituting "Course Outcomes."	• **We highlighted the definition** (which precedes the term) in bold lettering.	• "Learning Outcome" is the accepted term in teaching and learning. • We are confident that students will read the embedded definition: **"What you should be able to do or know by the end of the course."** • The concept of "Course Outcome" is equally high diction and refers to a different concept. • One concern was about a "cookie cutter" approach to learning. We are confident that LOs are capacious and can allow room for creative course goals. This is supported by best evidence in scholarship of Teaching and Learning.
18. Comment that mention of SDRC etc. information should be removed.	• Mention was limited to the Foundations Check and removed from another question.	• We agree that information about SDRC, CAPS, and other UCR support services are under the purview of the university and are readily available on the UCR website, but such services provide important support that may help students with their learning. • Additionally, such services support the UCR mission of diversity, equity, and inclusion. • Students may not be aware of the resources, and providing the information at the course level may be the best chance to provide it to students and reinforce the UCR-wide information. • The information is prepopulated in the CANVAS syllabus template in a collated and easily copied format. • As elsewhere, nothing on this checklist is a requirement
19. Concern that a mention of Canvas means that the SET will require instructors to use Canvas.	• We left the mention in place.	• Canvas is one of many modalities mentioned, including paper handouts. Instructors have the freedom to choose any or none of these modalities. • Our intention is to remind students that Course Foundations may be delivered in many ways, and to jog their memories about where such information might have been located.

A. Course Foundations Check

Appendix 1

Comments on Domain B: Course Experiences (formerly "Class Experiences")

<div style="writing-mode: vertical">Comments on Class Experiences</div>

20. "Class Experiences" title no longer fits given edits to questions in this section.	• Changed to "Course Experiences"	• We chose "Course Experiences" to better capture a wide range of instructional modalities, including synchronous, asynchronous, and individual meetings. "Class" implies "in the classroom."
21. Comment that a question about **"Respectful language . . .and . . . appropriate language"** should be 2 questions.	• We changed the question to "Did the instructor create a respectful environment?"	• We agree that the concepts differ and should be in separate questions. • However, after much discussion, we deleted "appropriate language" entirely. We concluded that "appropriate" was culturally specific, and would vary student by student, and thus could not be a useful, specific, or observable term.
22. Concern that "Respectful Language" would be too varied by culture, and that some of the provided examples may not apply to an instructor.	• We deleted the examples. • We changed the question to "Did the Instructor create a respectful environment?"	• We trust that students understand "respectful." • The open comment box immediately follows this question, allowing students to explain if they feel class wasn't respectful. • The example about "responsive to questions" was already addressed in "opportunities for student engagement" question), so we deleted that example. • The examples could introduce bias, so we deleted them. • Changing to "Environment" is more capacious and allows for respect beyond word choice (such as answering questions).
23. Concern that "Appropriate language" differs by culture and student.	• We deleted reference to "Appropriate."	• We agree the term can have multiple interpretations, inviting bias. It also different by course content.
24. Question about asking for outside help moved to this domain (from Learning Support section, now deleted), as it is part of Course Experience.	• Question now reads, **"If you asked for outside help, was the instructor responsive?"** • Added **"N/A – I did not ask for help"**	• We adjusted question to ask about instructor responsiveness and deleted and mention of "timeliness," which we had initially included. As mentioned elsewhere, this term varies widely by student. • Tier A captures objectively whether the instructor responded. • The key component of the initial question is the responsiveness of the instructor. • Whether the response was timely or helpful is captured in Tier B (Did it contribute to learning), which they can then explain in the comment box.

25. A suggestion that the instrument ask about whether the instructor accommodated "Varied learning styles."	• See Committee notes	• We discussed at length. We acknowledge that a variety of classroom experiences is good pedagogy. It is important enough to give students a chance to give feedback.
		• To avoid infringing on academic freedom, we avoided more specific language (e.g., "variety of activities," or "active learning"). "Accessible" also has the potential to be prescriptive.
		• We considered "that were accessible to you." We worried an instructor might provide numerous opportunities that students did not access, leading to a distorted result.
		• Also, "Accessible" might distract students from "varied" to only focus on "accessible."
		• We brought this issue to the full committee because it was such a complex issue.
		• Students whose learning styles were not accommodated can say that graded work or course experiences did not contribute to their learning, and then elaborate in the comment box.
		• We concluded that the issue of learning styles is captured by instrument questions such as:
		  • **B.2** (opportunities for student engagement)
		  • **Each Tier B question:** (Did the instructor's use of this practice contribute to your learning?)

Appendix 1

Comments on Domain C: Graded Work

Suggestion or Feedback	Action Taken	Committee notes
26. Concerns that "Clear directions" may be ambiguous and not observable.	• We deleted "clear."	• We agreed that "clear" varies by student. • Asking the follow-up "Did the instructor's use of this practice contribute to your learning?" captures whether it was clear.
27. Concern that the link between learning outcomes and assignments may be obvious to student and need not be stated.	• See Committee notes	• It is useful to know if the linkage was obvious or not obvious to students, which they can say in the comment box.
28. Concern that providing grading criteria won't be possible for some assignments and that it will create extra work for instructors.	• The question now reads, "Did the instructor provide **information about grading criteria and/or standards for how work was graded**?"	• "Criteria" means the standards by which an assignment is judged. • The standards may be as simple as "Your assignment will be judged by its correctness" or as complex as a variety of skills that must be exhibited in the assignment, especially for process-oriented assignments. • Clarity and consistency in standards are evidence-based best teaching practice. • Some instructors might not provide criteria until after grades are issued or only if asked. • Thus, we changed the question to encompass all points (before, during, after) when criteria might be revealed.
29. Suggestion to delete mention of rubrics, as some instructors don't use them.	• We deleted "rubrics."	• While rubrics are the accepted best practice in scholarship of Teaching and Learning, to avoid impinging on academic freedom, we deleted mention of rubrics. • "Did the instructor provide **information about grading criteria and/or standards for how work was graded**?" is expansive enough to include rubrics when instructors use them.
30. RE: C.1 "Directions for completing graded work" - concerns that some students might not perceive some assignments (e.g., quizzes) as learning tools, but rather only as assessment tools.	• See Committee notes	• We agree that students do not always know if something contributed to their learning. • Thus, each Tier B question, "Did the instructor's use of this practice contribute to your learning" has the option, "Not sure"– it appears at the end. "Not sure" is akin to "N/A."
31. RE: C.4 Concerns that "timely manner" is subjective.	• We changed to "Did the instructor **return graded work in time for you to monitor your performance throughout the course**?"	• Learning benefits from a regular formative assessment provided in time for students to change their course performance. • Also, students need to know their standing in class, even if a student is an A student. • We included a broad student purpose (knowing their progress) to encompass many motives for wanting to know their standing.
32. Concerns about the term "Assessment" being difficult for students to understand.	• Changed to "Graded work"	• We concluded that "Graded work" offers more clarity to students.

33. Concern that linking graded work to learning outcomes would create significant effort for instructors.	• See Committee notes	• Including learning outcomes (LO's) is an evidence-based core best practice in teaching. • Research indicates that linking assignments to those outcomes leads to a more coherent and organized curriculum that is easier for students to follow and thus to learn. • Research also shows that students who understand how an assignment will help them overall course goals (LO) tend to learn more, especially under-represented groups with minimal cultural capital. • There is no mandate for how the instructor creates this linkage (e.g. it can be a verbal explanation in class, written on the assignment itself, implied by assignment name, etc.).
34. Concerns regarding returning grades efficiently "What if TA is a bad/slow grader?"	• No action taken	• If the TA is grading, faculty may need to discuss a negative answer in their teaching statement, but scholarship in Teaching and Learning indicates this is important information for assessments of teaching. • If an instructor has a poor TA in one class, criticism/low scores won't be consistent over many classes; if it is an instructor problem, it will be consistent over classes. • Results may also be used to argue for more Teaching Assistants assigned to the course (if grading consistently cannot be completed in a timely manner for students to know their standing in the course).
35. Concerns about the term "Feedback" • varies by class • impossible in very large lectures	• Deleted this question	• We agree that not all classes provide feedback - some only provide a grade. • Also, "feedback" is implicitly included in the previous question about returning graded work.
36. Suggestion that we rework comment box to ask if the graded work assessed understanding of course material.	• No further changes made to the comment box.	• The series of questions captures this. If the instructor provides clear guidelines and criteria, the graded work should accurately assess the student's understanding. • Especially questions 2 and 3, which allow students an opportunity to say that the assessment did not contribute to their learning. The current phrasing is capacious enough for students to elaborate when needed.

Appendix 1

Comments on Domain D: "Learning Support

	Note: "Learning Support" domain was deleted (see notes below)	
37. Comment that Learning Support should be confined only to in-class requests.	• This question was moved to Domain B: "Course Experiences" and changed to "If you asked for help outside of class, **was the instructor responsive** (for example in office hours, scheduled appointments, emails, etc.)?"	• We discussed at length and concluded that questions B.2 & B.3 encompass **in-class** requests. • Thus, we framed the question to be about outside help. • Contact information is also included in Course Foundations Checklist, Q3: "Information on how to ask for help that was accessible to you (for example: office hours, email, Zoom)."
38. Comment that learning support should have an N/A option, as some students don't ask for help.	• N/A option added.	
39. Suggestion that outside assistance belongs in "Class Experiences."	• Moved to "Course Experience" (domain name changed – see notes in "Comments on Course Experience"	• Domain name changed to "**Course** Experiences." • "Course " is broad enough to include any experience, whether online, in person or in other contexts.
40. Concern that not all faculty provide optional supplemental materials	• Question deleted	• We agree that an effective instructor's curriculum as described in the syllabus should be sufficient. • While it might be kind or helpful to provide supplemental materials, it is not necessary that an effective instructor do so.

Appendix 2

Part I: Administered early in the quarter (i.e., Weeks 2-3). This part is **<u>OPTIONAL</u>**.

General Instructions: Please answer the following questions about the course. It should take you a few minutes to complete.

Course Foundations Check

All of the following provide a framework for learning in this course. These items may be found in the syllabus, on Canvas, and/or in handouts. Are they present in at least one of these places?

1. A clear description of **what you should be able to do or know by the end of the course** (learning outcomes)
   Yes/No

2. A clear description of the grading system for this class
   Yes/No

3. Information on how to ask for help that was accessible to you (for example: office hours, email, Zoom appointments, etc.)
   Yes/No

4. A list of course topics
   Yes/No

5. Assignment due dates
   Yes/No

6. Guidelines for academic integrity
   Yes/No

7. Course materials (may include readings, software, textbooks, recordings, and other resources)
   Yes/No

8. Information on how to ask for accommodations/support from the university (for example: Contact information for SDRC, CAPS, ARC, Title IX office, etc.)
   Yes/No

9. Were the above items compiled in an accessible location (e.g., a document, a Canvas page, or other format)?
   Yes/No

10. Are you enrolled in a lab for this course? IF NO, skip 11.
    Yes/No

11. IF YES: have you received instructions for taking appropriate safety measures in lab settings?
    Yes/No/ N/A—my lab does not require special safety measures

Appendix 2

General Instructions: The following evaluation has four sections. It should take you approximately 8-10 minutes to complete.

A) Course Foundations Check

All of the following provide a framework for learning in this course. These items may be found in the syllabus, on Canvas, and/or in handouts. Are they present in at least one of these places?

1. A clear description of **what you should be able to do or know by the end of the course** (learning outcomes)
   Yes/No
2. A clear description of the grading system for this class
   Yes/No
3. Information on how to ask for help that was accessible to you (for example: office hours, email, Zoom appointments, etc.)
   Yes/No
4. A list of course topics
   Yes/No
5. Assignment due dates
   Yes/No
6. Guidelines for academic integrity
   Yes/No
7. Course materials (may include readings, software, textbooks, recordings, and other resources)
   Yes/No
8. Information on how to ask for accommodations/support from the university (for example: Contact information for SDRC, CAPS, ARC, Title IX office, etc.)
   Yes/No
9. Were the above items compiled in an accessible location (e.g., a document, a Canvas page, or other format)?
   Yes/No
10. Are you enrolled in a lab for this course? IF NO, skip 11.
    Yes/No
11. IF YES: have you received instructions for taking appropriate safety measures in lab settings?
    Yes/No/ N/A—my lab does not require special safety measures

12. Did the instructor announce changes to the structure of the course (e.g., due dates, dropping assignments, etc.) in a timely manner?
    Yes/No/ N/A–No changes were made to the course structure.

Comment Box: What modifications, if any, to the list above would benefit FUTURE STUDENTS?

Appendix 2

B) Course Experiences

In the following questions, the phrase "course experiences" refers to things in the course context like lectures, group work, discussions, activities, fieldwork, lab work, guest speakers, videos, emails, office hours, discussion boards, etc., whether synchronous or asynchronous.

1.
   a. Did the instructor establish a **clear relationship (connection)** between **course experiences** and **what you should be able to do or know by the end of the course** (learning outcomes)**?**
      Always/Most of the time/About half the time/Rarely/Never/Not Sure
   b. Did the instructor's use of this practice contribute to your learning?
      Definitely Yes/Probably Did/I'm Neutral/Probably Didn't/Definitely Not/N/A

2.
   a. Did the instructor provide **opportunities for student engagement** during course experiences?
      Always/Most of the time/About half the time/Rarely/Never/Not Sure
   b. Did the instructor's use of this practice contribute to your learning?
      Definitely Yes/Probably Did/I'm Neutral/Probably Didn't/Definitely Not/N/A

3.
   a. Did the instructor create a **respectful learning environment**?
      Always/Most of the time/About half the time/Rarely/Never/Not Sure
   b. Did the instructor's use of this practice contribute to your learning?
      Definitely Yes/Probably Did/I'm Neutral/Probably Didn't/Definitely Not/N/A

4.
   a. If you asked for help outside of class, **was the instructor responsive** (for example in office hours, scheduled appointments, emails, etc.)?
      Always/Most of the time/About half the time/Rarely/Never/Not applicable–I did not ask for help
   b. Did the instructor's use of this practice contribute to your learning?
      Definitely Yes/Probably Did/I'm Neutral/Probably Didn't/Definitely Not/Not applicable–I did not ask for help

Comment Box: If any of these aspects of the instructor's approach to **class experiences** contributed to your learning, please share why. If any aspects did not contribute to your learning, what could the instructor improve or change about **class experiences** to better support future students' learning?

C) Graded Work
In the following questions, the phrase "graded work" refers to any graded class component. For example: exams, quizzes, projects, homework assignments, student presentations, performances, papers, etc.

1.
   a. Did the instructor provide **directions** for how to complete each piece of **graded work**?
      Always/Most of the time/About half the time/Rarely/Never/Not Sure
   b. Did the instructor's use of this practice contribute to your learning?
      Definitely Yes/Probably Did/I'm Neutral/Probably Didn't/Definitely Not/N/A

2.
   a. Was the **graded work connected** to **what you should be able to do or know by the end of the course** (learning outcomes).
      Always/Most of the time/About half the time/Rarely/Never/Not Sure
   b. Did the instructor's use of this practice contribute to your learning?
      Definitely Yes/Probably Did/I'm Neutral/Probably Didn't/Definitely Not/N/A

3.
   a. Did the instructor provide **information about grading criteria and/or standards for how work was graded**?
      Always/Most of the time/About half the time/Rarely/Never/Not Sure
   b. Did the instructor's use of this practice contribute to your learning?
      Definitely Yes/Probably Did/I'm Neutral/Probably Didn't/Definitely Not/N/A

4.
   a. Did the instructor **return graded work in time for you to monitor your performance throughout the course**?
      Always/Most of the time/About half the time/Rarely/Never/Not Sure
   b. Did the instructor's use of this practice contribute to your learning?
      Definitely Yes/Probably Did/I'm Neutral/Probably Didn't/Definitely Not/N/A

Comment Box: If any of these aspects of the instructor's approach to **graded work** contributed to your learning, please share why. If any aspects did not contribute to your learning, what could the instructor improve or change about **graded work** to better support future students' learning?

Appendix 3

SET Question Criteria

To help ensure that Campus-Wide SET questions avoid bias, adhere to research-based best practices, and exhibit consistency, we advocate for the creation of an oversight committee that is charged with vetting any proposed Campus-Wide question (whether required or optional) for inclusion in the SET. Below, we provide six criteria for vetting proposed questions. These criteria are based on the latest research in the Scholarship of Teaching and Learning and are designed to avoid potential bias as much as possible. By adhering to these criteria, the oversight committee will allow flexibility in question inclusion while also maintaining a consistent evaluation that does not compromise its quality. These criteria do not apply to Instructor-Discretionary questions which may be directly added by an instructor to the SET for each of their courses.

1. **Observable**: Each question should allow students to provide feedback on a teaching practice that **they can directly observe during the course.**
   a. **Exemplary:** Did the instructor provide instructions for how to complete graded work?
   b. **Unexemplary**: Did the instructor prepare the graded work well?
      This is not observable because students cannot see the instructor while preparing.
2. **Specific:** To maintain clarity and precision, each question should focus on a single variable related to teaching practice and avoid ambiguous or broad inquiries.
   a. **Exemplary:** Did the instructor provide information about grading criteria and/or standards for how work was graded?
   b. **Unexemplary**: How was the overall class experience?
      This question is not specific since it asks about the overall class experience that could include many elements. It is not about a specific practice.
3. **Evidence-based:** All questions should seek to gather feedback on teaching practices that are supported by evidence and recognized as best practices from teaching and learning research.
   a. **Exemplary:** Did the instructor provide opportunities for student engagement during course experiences?
   b. **Unexemplary**: Did the instructor primarily use multiple-choice questions in the exams?
      While multiple-choice questions can be helpful, there is no evidence that this is a best practice for any class.
4. **Neutral**: To ensure a fair and unbiased evaluation process, questions should avoid biased language, especially adjectives, that invite value judgments.
   a. **Exemplary:** Did the instructor provide a clear description of what you should be able to do or know by the end of the course (learning outcomes)?
   b. **Unexemplary:** Did the instructor demonstrate a comprehensive and deep understanding of course materials?
      This question is prone to bias, as "comprehensive and deep understanding" is a value judgment. The terms are not neutral and invite bias.

Appendix 3

5. **Clear**: To avoid multiple interpretations, questions should be crafted to resist ambiguity, leading to unambiguous and straightforward responses from students.
    a. **Exemplary:** If you asked for help outside of class, was the instructor responsive (for example in office hours, scheduled appointments, emails, etc.)?
    b. **Unexemplary:** Was the instructor approachable during this course?
        The word "approachable" is not very clear. It could be interpreted as the instructor being friendly, respectful, quickly responsive to questions, or many other things that are known to the student, but unknown to anyone reading the response.
6. **Accessible**: Jargon and technical terms should be avoided in the questions to ensure that students can easily understand and respond to them without confusion.
    a. **Exemplary:** Was the graded work connected to what you should be able to do or know by the end of the course (learning outcomes)?
    b. **Unexemplary**: Did the instructor use a combination of formative and summative assessments?
        An average student might not know what "formative" and "summative" mean or how to determine if an assessment is either of these.