

Validating the Critical Teaching Behaviors Midterm Feedback Instrument for enhanced teaching effectiveness and student engagement

Claudia Cornejo Happel, Chad Rohrbacher, Teha Cooks,
Jenna Korentsides, Joseph R. Keebler, Carmen Van
Ommen, Kimberly N. Williams, and Lauren Barbeau

Abstract

This study, conducted at a private STEM university in the Southeastern United States, presents a series of qualitative and quantitative data validating the Critical Teaching Behaviors Midterm Feedback Instrument (CTB-MFI). The instrument's validity and reliability were tested through a mixed-methods approach involving faculty surveys, statistical analyses, and a student experience survey and focus groups. A confirmatory factor analysis (CFA) suggests that the CTB-MFI model has excellent fit across multiple indicators, as well as excellent convergent validity. However, low discriminant validity indicates possible correlations between categories of critical teaching behaviors. Data further show that faculty and students perceive the instrument as valuable and easy to use: the CTB-MFI offers insightful midterm feedback, leading to improved teaching quality and student learning experiences. The CTB-MFI supports flexible use across various disciplines and provides a shared framework and language for discussing teaching effectiveness. This research offers valuable implications for educators, administrators, and researchers interested in improving

learning outcomes and diversifying approaches to evaluating teaching effectiveness.

Keywords: midterm student feedback, student evaluation of instruction, documenting teaching, teaching effectiveness, assessment of teaching

When discussing and documenting effective teaching in higher education, faculty are in a unique situation—teaching is a significant component of their day-to-day job and their performance evaluation, yet few faculty receive pedagogical training. Additionally, while teaching “excellence” or “effectiveness” is a stated goal in the mission and vision of many universities and colleges, a definition of what this looks like in practice is often vague or missing. To gauge teaching effectiveness, many institutions rely almost exclusively on student end-of-course (EOC) evaluations and, possibly, peer observation. Often, feedback received from students and colleagues directly impacts performance reviews and career advancement rather than providing opportunity for reflective discussion and growth. Furthermore, the validity and reliability of EOC evaluations have been increasingly critiqued, specifically due to significant problems of bias against certain faculty populations (Clayson, 2022; Fan et al., 2019; Kreitzer & Sweet-Cushman, 2022). In this context, many institutions have begun to look for more holistic approaches to assessing teaching effectiveness.

Student midterm feedback is an established formative practice that encourages feedback on instruction to improve the student learning outcomes and experience during the semester a course is taught, rather than waiting for EOC feedback to apply to instructional efforts in a future semester. The Critical Teaching Behaviors Midterm Feedback Instrument (CTB-MFI) is an important addition to the field due to its focus on assisting faculty in identifying specific areas of strengths and improvement in their teaching practice. Grounded in the Critical Teaching Behavior (CTB) framework (Barbeau & Cornejo Happel, 2023), this tool also provides a language and process for instructors

to document and discuss their teaching in focused but nuanced ways. The CTB framework defines six categories of effective instructional behaviors—Align, Include, Engage, Assess, Integrate Technology, Reflect—that are broad enough to apply across disciplines and instructional modalities while providing a shared language to discuss effective teaching.

The CTB-MFI was tested for its reliability and validity through a mixed-method study design that included (1) faculty feedback surveys, (2) statistical analyses of quantitative data gathered using the CTB-MFI, and (3) a student experience survey and focus groups. The researchers hypothesized that the CTB would demonstrate high internal consistency and construct validity. The research questions include:

- (1) Relevance: Is information provided through the survey useful to faculty in identifying and documenting their teaching strengths and reflecting on areas of growth in teaching?
- (2) Reliability: Is the survey valid and reliable based on statistical measures, specifically a confirmatory factor analysis (CFA)?
- (3) Usability: Is the survey clear and easy to use for students? Specifically, when answering the quantitative questions concerning critical teaching behaviors, do participants interpret them consistently?

Based on the data gathered, the CTB-MFI offers educators a reliable and valid tool to collect midterm feedback from students. The CTB-MFI report provides instructors with a helpful overview of areas of teaching strength and potential for improvement along with insights into their overall use of effective instructional strategies. This feedback is useful for improving the quality of teaching and learning in the classroom, enhancing teaching effectiveness and student success. The results of this study have implications for educators, administrators, and researchers interested in developing and using midterm feedback surveys as a means of improving teaching and learning outcomes and in diversifying data sources considered in the evaluation of instruction.

End-of-Course (EOC) Evaluations

Feedback from students is typically completed through student EOC evaluation of teaching. Student feedback involves students expressing their values, opinions, beliefs, and perspectives, which can be used to shape instructional approaches. In recent years, multiple studies have raised concerns regarding the problematic nature of evaluating instructor performance based exclusively on student EOC evaluations. Of particular concern is the impact of bias that has been demonstrated to negatively affect EOC responses and feedback scores for female and minoritized instructors (Boatright-Horowitz & Soeung, 2009; Boring et al., 2016; Clayson, 2022; Fan et al., 2019; Kreitzer & Sweet-Cushman, 2022; Mitchell & Martin, 2018; Reid, 2010). Furthermore, feedback received through EOC evaluations is not timely; when faculty receive the feedback provided by students at the end of the term, they are often no longer able to respond effectively. Conversely, midterm feedback surveys are generally voluntary, low-stakes processes that collect anonymous data about the course with the explicit goal of improving the teaching and learning experience during the semester they are facilitated.

Midterm Feedback

Midterm feedback is a particularly effective way for faculty to gather information from students to refine their classroom pedagogy. In the 1970s, the University of Washington created a midterm feedback process called the Small Group Instructional Diagnosis (SGID), which was intended to provide instructors with early feedback on their teaching methods (Murray, 1984; Redmond, 1982). Compared to EOC student evaluations, students generally find this midterm feedback process more valuable (Mauger, 2010). They are also more engaged in the process and provide more useful feedback (Veeck et al., 2016). Participation in facilitated midterm feedback sessions was also found to

positively impact student in-class engagement and study behaviors (Hurney et al., 2021). For faculty, midterm feedback seems to have a bigger impact on instructional practice than EOC feedback. Diamond (2004) found that instructors change their in-class assignments, teaching techniques, and assessment methods after receiving midterm feedback, whereas it is unclear if the EOC evaluations have the same impact on teaching.

To emphasize the formative rather than evaluative goal of midterm feedback, survey tools most frequently used to gather midterm student feedback prioritize qualitative, open response items, often to the exclusion of quantitative responses (Hurney et al., 2021). However, instructors can benefit from receiving feedback on teaching in multiple formats. Depending on the faculty member's perspective, either quantitative or qualitative data might present an easier entry point for engaging with the feedback. Combining quantitative and qualitative data can better contextualize the feedback and provide richer insights into students' experiences in classes, allowing instructors to identify specific opportunities for growth.

There are few validated instruments specifically designed for midterm feedback. The Reiser and Dick Instructional Planning Model (1996) provides specific feedback on planning and delivery of instructional activities, though it was primarily used in the K–12 context. Donlan and Byrne (2020) noted there are few midterm evaluation surveys that are grounded in education literature and that have “held up to psychometric scrutiny” (para 3). They developed the Mid-Semester Evaluation of College Teaching (MSECT), which evaluates four constructs of teaching: classroom climate, classroom content, teaching practice, and assessment—to be used in higher education contexts.

The CTB-MFI is based on research that defines effective teaching through behavioral categories (Barbeau & Cornejo Happel, 2023). Survey items aligned with these categories are phrased with a focus on specific, observable instructor actions to increase clarity, limit student misunderstanding, and avoid asking students to comment on things outside their expertise. For example, instead of asking broadly if the

course is structured well, which students cannot adequately answer due to the ambiguity of the question, the CTB-MFI asks students to report whether faculty “explain how assignments, lessons, and course activities help you develop knowledge and skills related to course goals.” Specificity matters when it comes to items that ask students to report perceptions or evaluations of teaching behaviors on various aspects of the course (McKeachie, 1997; Murray, 1997). Specific, focused questions, as opposed to broad, generic items, prompt feedback more likely to help instructors understand and respond to student learning needs. Moreover, the CTB-MFI survey questions intentionally “focus student attention on the prevalence of instructor behaviors to help mitigate some of the biases that tend to emerge in quantitative student ratings” (Barbeau & Cornejo Happel, 2023, p. 132). The pilot for a separate study on the CTB-MFI concerning bias (in progress) resulted in promising initial findings that indicate little to no correlation between instructor ratings on the CTB-MFI and select demographic factors. This suggests that implicit biases might play less of a role when questions included in student evaluations of teaching focus specifically on observed frequency of concrete, observable instructional behaviors.

Methodology

Study Overview

This research project is based on insights from three distinct data sources. To validate the tool and test its usefulness in practice, the researchers conducted (1) faculty feedback surveys, (2) statistical analyses of quantitative data gathered using the CTB-MFI, and (3) student experience survey and focus groups. The data informing this study were collected at a private STEM university in the Southeastern United States specializing in aerospace and aviation with approximately 350 full-time faculty. In the 2021–2022 academic year, it served approximately 8,000 students.

The survey instrument analyzed in this study includes a total of 15 quantitative ranking items; there are three Likert-style items for each of the five behavior categories as shown in Appendix A. For each item on the CTB-MFI, students are asked to report the frequency with which they see instructional behaviors aligned with these categories enacted in their classroom; response options range from *never* (1) to *always* (5). The survey also includes three open response questions asking students to provide more detail on what is working well, what is not working well, and any course concepts they might be confused about. Typically, it takes students no more than 10 minutes to complete the survey.

Student feedback gathered using the CTB-MFI is compiled in the instructor report. The instructor report mirrors the survey instrument provided to students but also includes an overview of five of the CTB categories, including a definition for each category, along with the three associated survey items. (The final category, "Reflect," is not assessed in the CTB-MFI because the behaviors, such as instructor self-assessment and engagement in professional development, are not visible to students.) To emphasize that midterm feedback is a formative process that benefits from faculty follow-up with students, the report includes a prompt encouraging instructors to document their insights and identify action steps they will take based on feedback received. The process followed (and recommended) for facilitating midterm feedback using CTB-MFI is published in Appendix B.

The CTB-MFI has been used at the location of this study since Spring 2021 with over 120 faculty representing all colleges completing the survey in at least one section of their course(s). The three data sources contributing to this research project have been implemented in phases since Spring 2021 as outlined below. The following data were collected with the goal of providing a holistic assessment of the validity and reliability of the CTB-MFI.

- **Faculty Survey (Spring and Fall 2021):** The CTB-MFI (v. 1.0; 10 quantitative items) was used in a pilot limited to College of Arts and Sciences (COAS) in Spring 2021; in Fall 2021 the CTB-MFI was

used for all midterm feedback sessions in the COAS, College of Business (COB), and College of Engineering (COE); researchers conducted faculty survey to assess faculty response and gather feedback on potentially problematic survey items.

- **Confirmatory Factor Analysis (Fall 2022):** The CTB-MFI (v. 2.0; 15 quantitative items)¹ was used consistently for all midterm feedback sessions in COAS, COB, COE, and College of Aviation (COA); the data were utilized to conduct a CFA of student responses to establish reliability.
- **Student Survey and Focus Groups (Spring 2023):** Researchers conducted student surveys and focus groups with undergraduate and graduate students to evaluate whether student interpretation of the quantitative CTB-MFI questions were consistent across users and aligned with the intention of the question as defined by the research team.

For the sake of clarity, the following sections will report on each of the three areas of data collection and analysis separately before overarching insights of this research project are presented in the conclusion.

Faculty Feedback

Methodology—Faculty Feedback

The CTB-MFI was implemented on a limited scale in Spring 2021 with 11 faculty from the COAS only. At the end of the semester, researchers reached out to faculty who had conducted the CTB-MFI in their

¹ Researchers expanded the list of items after the analysis of an initial data set collected using the original CTB-MFI confirmed the statistical validity of the survey as a holistic measure of effective teaching but could not validate the questions as a measure of five distinct categories of critical teaching behaviors. Completion of an omnibus test (using all 10 survey items) resulted in a Cronbach's alpha value of .891, which indicates good inter-item reliability for the survey overall as a single dimension.

courses to request feedback on the CTB-MFI format, clarity, and usefulness. In Fall 2022, after implementing the CTB-MFI process with a wider audience, faculty participants were again invited to complete a version of this survey focused on faculty perceived usefulness of the quantitative and qualitative feedback responses included in the CTB-MFI by responding to the following questions:

- When reviewing your midterm feedback report, how useful was the overview (data table and overview chart) of feedback students provided by answering the ranking questions?
- Please explain your rating. Consider, for example, what did you find useful/ not useful? How did you use the data from ranking questions? What would make this type of data more useful?
- When reviewing your midterm feedback report, how useful was the narrative overview of students' open response feedback on your teaching behaviors, including representative student comments?
- Please explain your rating. Consider, for example, what did you find useful/ not useful? How did you use the data from narrative feedback questions? What would make this type of data more useful?

In Spring 2021, faculty were additionally asked to respond to the following open response question:

- CTB Questions: Anything missing? After reviewing the ranking questions on the survey, were you hoping to receive feedback on any additional ranking items not included in the survey?

Responses to questions asking faculty to rate the usefulness of the quantitative data and narrative data overview were reported on a 5-point Likert scale with a response anchor at each rating point (e.g., 1 = *not at all useful* and 5 = *very useful*). Both surveys were voluntary to complete and facilitated online through MS Forms; no demographic information beyond the instructor's college affiliation was collected.

Results & Discussion—Faculty Feedback

A total of 37 faculty from COAS and COE responded across the two semesters (see Table 1). In Spring 2021, most faculty indicated that they did not feel anything was missing from the questions asked and offered no direct suggestions for changes, though few provided important feedback in response to prompts asking them to explain their ratings that led to minor revisions of survey items to increase clarity and relevance. For example, based on feedback received, researchers added a list of potential technology options, including the university's learning management system (LMS), to the question that asked students to report the frequency with which their instructor "[u]ses technologies and/or apps that enhance your learning experience in the course." Additionally, a faculty comment stating "I was looking for more focused questions on the clarity of assignments and grading rubrics, rather than just the opportunities to have learning assessed" prompted the revision of questions to more specifically gather student input on how frequently assignments expectations are clearly communicated. Faculty responses to the survey questions rating the usefulness of the CTB-MFI tool are summarized in Table 1.

These data show faculty consider both types of feedback to be useful, but overall they perceive the insights gained from qualitative student responses to be more useful than those provided by quantitative data, a finding that aligns with insights from an earlier study showing that faculty prefer student comments to Likert scores (van Wyk &

Table 1. Faculty Average Rating of Usefulness of Quantitative and Qualitative Feedback

Semester	College	Participants (total $n = 37$)	Usefulness of quantitative feedback (average score)	Usefulness of qualitative feedback (average score)
Spring 2021	COAS	11	4.55	4.90
Fall 2021	COAS	13	4.62	4.85
	COE	13	4.62	4.62
	Total (Fall 2021)	26	4.62	4.73

Mclean, 2007). Interestingly, this does not hold true for engineering faculty, who on average considered both qualitative and quantitative feedback to be equally useful, indicating a marked difference between faculty preferences in arts and sciences (COAS) and engineering (COE).

When asked to explain their ratings, faculty elaborated that the quantitative scores were useful because they provided an at-a-glance overview of effective teaching behaviors and how “well” faculty were integrating them into their courses; one COE instructor stated, “[t]he quantitative data of different categories were helpful to see how I was doing in the different categories of effective teaching behavior, and quick [sic] understand the areas I need to improve on.” Faculty also mentioned that the data provided a baseline that they would be able to track over time: “I suspect that tracking these changes over time would be more helpful, especially if they could be cross-referenced with end-of-term evaluations to show trends of growth over time.” Faculty were aware of how the data might assist them in summative evaluations. One faculty member stated, “I’m glad we had [quantitative data] to consider, especially since numbers are often what higher admin are most likely to pay attention to!” That said, several survey respondents mentioned that they found quantitative data useful only in combination with the qualitative feedback provided.

The qualitative student comments provided faculty with more specific insights and concrete guidance for understanding student concerns and making pedagogical changes. Specifically, faculty stated that comments helped them identify pedagogical opportunities: “The narrative data helped me understand what the students are confused about, where I need to change my teaching methods and what I need to clarify moving forward.” Faculty also appreciated how the qualitative feedback provided much needed context: “Having some explanation of the scores in the quant section and specific feedback about what’s happening in the course (and how widespread agreement with that concern is) was the most helpful part.”

Finally, one aspect of the CTB-MFI that goes beyond student feedback responses and separates it from other instruments is its explicit

purpose to be used as a framework intended to guide both instructor reflection and discussion around pedagogy. Faculty feedback indicated that they found the instrument extremely useful in both areas. For instance, one participant said, “[the CTB-MFI] gave us [the educational developer and faculty member] a structure for the discussion,” recognizing how the tool was able to productively guide both the discussion facilitator and the faculty receiving feedback through the debrief conversation. Another shared how the analysis of the feedback data was a positive experience: “I enjoyed seeing how [the CTB-MFI] was explained and [it] gives me concrete feedback for areas I need to focus on to improve my work in the future.”

Summary—Faculty Feedback

Overall, faculty found both quantitative and qualitative feedback helpful in understanding student learning experiences and perceptions of teaching in their courses. Some faculty emphasized that quantitative data serves as a useful addition that provides at-a-glance insights and a foundation for documenting teaching strengths for an external audience. Other faculty emphasized that qualitative feedback must remain an essential component of midterm feedback as it provided context and depth to the survey data. The instrument and associated report contribute to the establishment of a common definition and framework of effective teaching and shared language for pedagogical review and discussion.

Confirmatory Factor Analysis

Methodology—CFA

Teaching center staff administered the midterm feedback survey via Microsoft Forms in Fall 2022 in 151 sections taught by 76 faculty. Table 2 provides more details on the disciplines represented and course level of the sections in which the CTB-MFI was administered.

Table 2. Student Responses per Discipline and Course Level

		Number of total responses	Number of responses in CFA data set
Total participants		2,978	1,487
Discipline	Humanities and Social Sciences	461	223
	Sciences	596	304
	Engineering	1,456	729
	Aviation	353	179
	Business	103	49
	Other	9	3
Course level	100	1,064	531
	200	735	359
	300	568	285
	400	461	233
	500+ (graduate level)	150	79

Based on faculty preference, teaching center staff either visited the classroom to introduce the survey and facilitate its completion in the absence of the faculty or provided a link to the faculty member that was shared with students via the institutional LMS (Canvas). When shared online, students would complete the survey asynchronously. In the classroom, the survey was administered as a standalone feedback tool or in combination with a whole-group class discussion (modeled on the last step in the SGID protocol); when combined with a discussion, students first individually completed the survey before discussing their feedback and observations with classmates. In all cases, students completed the survey anonymously; no identifying data was collected. For this study, only the quantitative portion of the CTB-MFI data collected was considered for the CFA.

Quantitative responses were exported from MS Forms into a spreadsheet; data were randomized and divided into two comparable groups for analysis. A CFA was conducted using one of these data groups to assess the degree to which the specified survey or scale measurement fits the observed data. The CFA helps researchers test whether the items within the survey measure the intended underlying constructs or factors. By doing so, a CFA enables researchers to evaluate the construct validity of surveys and other measurement

instruments and validate their theoretical assumptions. Within the output of a CFA, the fit indices, standardized regression weights, assessments of discriminant and convergent validity, and correlations between factors are of particular interest. The following sections present the results of the CFA.

Results & Discussion—CFA

All analyses were conducted utilizing AMOS graphics 26.0. We assessed the fit of the survey using a chi-squared test, the Tucker-Lewis index (TLI), comparative fit index (CFI), and the root mean square error of approximation (RMSEA) (Byrne, 2016; Hu & Bentler, 1999).

Fit indices are used to assess how well the measurement model fits the data. The fit of the survey will be evaluated using several factors including chi-squared, which tests the difference between the observed and expected covariance matrices (Byrne, 2016). In this case, the chi-squared test is significant, indicating a good fit. The TLI and CFI measure how well the model reproduces the observed covariance matrix (Byrne, 2016). Values above 0.95 indicate good fit; therefore, in this case, both values indicate that collected data support the proposed five-category model as a good fit, as shown in Table 3. Finally, the RMSEA value estimates the discrepancy between the hypothesized model and the population covariance matrix (Byrne, 2016). For RMSEA, values below 0.08 indicate good fit, which, as seen in Table 3, is the case for the present model.

The standardized regression weights represent the strength and direction of the relationships between the latent constructs, or factors (e.g., CTB categories such as Align, Include), and their observed variables (e.g., aligned items or questions within the survey) (Byrne, 2016). Each factor has multiple items within the survey, and the weights indicate how strongly each item is related to its corresponding factor. Standardized regression weights typically range from -1 to +1 (Byrne, 2016), with larger absolute values indicating stronger relationships between the constructs (latent variables) and their related survey

Table 3. Fit Indices for the Survey

Model	χ^2 p value	Degrees of freedom	TLI (> .95)	CFI (> .95)	RMSEA (< .08 /> .05)
CTB-MFI survey	537.823 ² $p < .001$	80	.950	.962	.062 ³

Table 4. Standardized Regression Weights

Model	Estimate
LEARNING_OUTCOMES < Align	.737
TIME_MANAGEMENT < Align	.759
ALIGNED_ASSESSMENT < Align	.667
STUDENT_PERSPECTIVE < Include	.713
RANGE_PERSPECTIVES < Include	.779
COMMUNITY < Include	.749
PARTICIPATION < Engage	.701
REAL_APPLICATION < Engage	.728
COMMUNICATION < Engage	.775
FEEDBACK < Assess	.722
SCAFFOLDING < Assess	.611
TRANSPARENT_EXPECTATIONS < Assess	.808
ONLINE_ORGANIZATON < Tech Integration	.775
TECH_ENHANCE < Tech Integration	.812
TECH_TRAINING < Tech Integration	.796

items. Table 4 shows the different items for each construct (e.g., in the Align category: LEARNING_OUTCOMES, TIME_MANAGEMENT, ALIGNED_ASSESSMENT; see Appendix A for item details) and their standardized regression weights (e.g., 0.737, 0.759, 0.667).

A CFA also allows us to assess convergent and discriminant validity. Convergent validity assesses the degree to which the indicators within a construct are related to each other, whereas discriminant validity evaluates the degree to which different constructs are distinct from one another (Byrne, 2016; Hu & Bentler, 1999). To interpret

2 A significant chi-squared test is usually an indicator of good fit. The chi-squared test reported in this column is significant. Although it is standard to report chi-squared tests for CFA/SEM analyses, they are almost always significant due to large sample size. Therefore, other fit indices are provided that account for the large sample constraints of these techniques.

3 Indicates the value for the model is in the acceptable range of fit indices.

Table 5. Assessment of Discriminant and Convergent Validity

Constructs	CR	AVE	MSV	MaxR(H)
Align	0.765	0.521	0.826	0.895
Include	0.791	0.559	0.974	0.925
Engage	0.779	0.541	0.974	0.941
Assess	0.759	0.516	0.826	0.951
Tech Integration	0.837	0.631	0.731	0.838

convergent and discriminant validity within a model, several tests are considered including the average variance extracted (AVE) and maximum shared variance (MSV). The AVE measures the amount of variance captured by the indicators of a construct, with values above 0.5 indicating good convergent validity, as shown in Table 5 for the present model (all above 0.5). This indicates that all items within each construct converge to measure that corresponding single construct. The MSV represents the maximum amount of variance that a construct shares with any other construct (Byrne, 2016; Hu & Bentler, 1999). To have good discriminant validity, the AVE should be greater than the MSV. Table 5 provides the AVE, MSV, and other related measures for each construct (e.g., Tech Integration, Align, Include).

Furthermore, the composite reliability (CR) value represents the reliability or internal consistency of the constructs in the model, with a value of 0.7 or higher being considered acceptable. While the model appears to have excellent convergent validity, the data indicates moderately poor discriminant validity. First, the square root of the AVE for Align, Include, Engage, Assess, and Tech Integration is less than one, the absolute value of the correlations with another factor. Second, the AVE values for Align, Include, Engage, Assess, and Tech Integration are all less than the MSV. This could be an indication of overlapping items within several constructs or concept overlaps. Data showing correlations between different constructs, presented in Table 6, support this assumption; the numbers on the diagonal represent the squared correlation between each construct and its manifest variables (survey items), all of which portray high correlations with one another (with

Table 6. Correlations Between Factors

Constructs	Tech Integration	Align	Include	Engage	Assess
Tech Integration	0.794				
Align	0.814	0.722			
Include	0.732	0.894	0.747		
Engage	0.787	0.906	0.987	0.735	
Assess	0.855	0.909	0.787	0.854	0.718

Note. Numbers on the diagonal represent the squared correlation of that factor with its manifest variables.

values > 0.5 indicating a strong, positive correlation). In other words, if we theoretically merged these constructs into one overarching latent variable, such as “effective teaching,” the combined constructs would measure it very well.

Summary—CFA

Overall, the results suggest that the model had excellent fit across multiple indicators (e.g., RMSEA, TLI, CFI). This tells us that our data covariance matrix aligns very well with the hypothesized model’s covariance matrix and that our respective measures (e.g., the survey items) are reasonably good indicators of their latent constructs (e.g., the CTB categories). As for construct validity, the model had excellent convergent validity, indicating that the items for each construct are good indicators of that construct. However, the model had moderately poor discriminant validity, which assesses whether the individual latent variables (e.g., the CTB categories) are distinct from one another. In some instances, our test for discriminant validity demonstrates that constructs are moderately related to each other, as seen in Table 5. The overlap between constructs shown in the data comes as no surprise as any instructional behavior may often accomplish multiple purposes. For example, while active learning tasks allow an instructor to Engage students in the learning process, they also have been linked to promoting equitable learning environments (Include) and, if they Align well with learning outcomes, allow us to Assess student progress and understanding. Active learning

tasks might also be facilitated by strategic Tech Integration, such as student response systems. In short, connections between categories of critical teaching behaviors often serve as an asset rather than a shortcoming. While good convergent validity of the model allows us to assume that items within each construct provide us with reliable insights into achievements related to the instructional behavior assessed in each category, the low discriminant validity reminds us that ultimately all behaviors are correlated, and good teaching must be considered holistically.

Student Feedback

Methodology—Student Feedback

Researchers solicited student feedback to evaluate consistency of student interpretations of each question and to understand students' thought processes when completing the CTB-MFI; student comments were expected to provide further insights into the construct validity of survey items by allowing researchers to assess whether student interpretations were consistent across participants and aligned with the intended purpose of each item. Undergraduate and graduate students ($n = 8$) participated in a 45-minute feedback session after being recruited through outreach to student organizations and by posting flyers on approved student communication boards on campus. An overview of participant demographics can be seen in Table 7.

Three feedback sessions with two to three participants each were conducted. First, participants completed a survey that asked them to subsequently read each of the 15 ranking items included on the CTB midterm feedback tool and then (1) restate the question in their own words and (2) describe what kind of evidence they would consider when rating their instructor for this item. Next, students participated in a short focus group conversation to discuss their insights and reaction to the midterm feedback items and to share general comments.

Table 7. Demographics of Student Participants in Survey and Focus Group

	First-year	Sophomore	Junior	Senior	Graduate
Male	1	2		1	2
Female			1		1

Focus group conversations were audio recorded and transcribed. After completing the focus group sessions, the researchers analyzed the results from the written responses and the focus group conversations to identify themes.

Results & Discussion—Student Feedback

Student survey responses, summarized in Appendix C, enrich our understanding of student interpretations of CTB-MFI items and the evidence they would consider in rating their instructors' behavior. Written survey responses suggest that students' interpretations of the ranking questions asked on the CTB-MFI were largely consistent across participants with small variations. For example, when restating the item asking whether the instructor "selects examples and activities that represent a range of perspectives and experiences," student responses showed general agreement with the idea that the instructor should incorporate "a multitude of examples and applications that can benefit the entire class"; specifically, though, some students expected that the examples included would show a diversity of applications across different careers, whereas other students expected instructors to demonstrate that "history has suppressed minorities and that they are actively working to be more inclusive and diverse."

Students' responses restating the CTB-MFI items indicate each of the CTB-MFI items is clearly perceived as asking a distinct question. However, responses to the second survey question—"What evidence would you consider when evaluating your instructors' efforts in this area?"—indicated some overlap between evidence considered in the assessment of different teaching behaviors. For example, students considered instructor availability outside of class time—in-office

hours and by email—as an important piece of evidence informing their feedback on how often their instructor “Invites students’ questions, examples, and experiences and listens carefully when students speak” (Include) as well as how frequently their instructor “Establishes regular and open communication” (Engage). This insight mirrors CFA findings indicating some expected overlap between the categories in this model.

Completing the survey required students to engage with and reflect on the ranking items included in the CTB-MFI and provided them with a thorough overview of the tool and the process. In the focus group conversations that followed, students were asked to provide specific feedback on the CTB-MFI tool, particularly if they thought any questions were redundant, irrelevant, or missing, and general impressions of the midterm feedback process. Students agreed that none of the questions were confusing or unclear. They had few suggestions for specific changes to the tool. One group considered that the questions associated with the Tech Integration category were similar and suggested eliminating the following item “Uses technologies and/or apps that enhance your learning experience in the course (e.g., Canvas, multimedia content, polling)”; none of the other groups suggested eliminating any of the items. One focus group responded especially positively to the CTB-MFI items:

S1: “Well, I’d say these questions are a lot more helpful than those [on the EOC evaluation]. . . . They are more helpful to evaluate the teaching and learning experience of the student. Whereas some of the questions [on the EOC evaluation]—I can’t remember.”

S2: “Well, some of them, I would agree, they felt a little bit more course specific.”

This conversation suggests that students appreciate the request to provide feedback on specific, observable teaching behaviors rather than more abstract or ambiguously defined ideas or qualities. Students in all three focus group sessions remarked that midterm feedback

sessions should be more strongly encouraged and incentivized for all instructors. One student stated, "I feel like a midterm feedback is more important than end-of-semester survey. End-of-semester survey is just as important, but it's kind of like when in your classes you don't know how you're doing until just like the very end and that's, that's not helpful at all." Several students agreed that midterm feedback is at least as important as EOC feedback, maybe even more so, because it provides valuable information for the instructor to make changes and improvements to benefit students during the current course.

Summary—Student Feedback

Overall, insights from the student feedback sessions support that the CTB-MFI items are clear, relevant, and interpreted consistently across respondents. Each item on the survey is perceived as clearly distinct from the others, although there is some overlap in the evidence students consider when evaluating each item. Students further agreed that they consider midterm feedback in general an important avenue for providing feedback that should be incentivized so more instructors seek formative feedback from students rather than wait until the end of the semester.

Discussion

The purpose of this study was to validate the CTB-MFI through multiple measures and to evaluate its reliability, relevance, and usability. A CFA ($n = 1,487$) conducted to confirm statistical validity and reliability suggested that the proposed model with its five behavioral categories had excellent fit across multiple indicators (e.g., RMSEA, TLI, CFI). As a result, data gathered with this tool can be considered to provide valid and reliable insights into categories of strengths in an instructor's teaching as well as indicate categories in which there is potential for improvement to further increase instructional behaviors that lead to

student success. Data also reveal that while the CTB-MFI model has excellent convergent fit, confirming that the three survey items for each category are reliably related to one another, moderately poor values for discriminant validity indicate some expected correlations between the five CTB categories. These correlations remind us that, even though we can benefit from feedback on how well our teaching practices align with the five critical teaching behaviors assessed in this survey, good teaching ultimately must be considered holistically.

In addition to validating the CTB-MFI through statistical means, we gathered faculty and student feedback to evaluate its usability and relevance. Student responses to survey and focus group questions indicated that the questions asked on the CTB-MFI were perceived as clear and relevant. Interpretation of CTB-MFI items was largely interpreted consistently across student participants; these findings further confirm the construct validity of survey items, suggesting that the questions included indeed serve as reliable measures of the instructional behaviors they are intended to assess.

However, these findings would be irrelevant if instructors did not see value in the insights provided through the CTB-MFI tool and process; as a result, this study started by asking for instructor input on its perceived usefulness. Their feedback suggested that the addition of quantitative items to the midterm feedback process, which generally prioritizes student responses to open-ended questions, is perceived as helpful by instructors across disciplines. Instructors reported that quantitative data gathered using the CTB-MFI offered useful information that directly impacts their teaching practice and ability to discuss teaching with colleagues and administrators by providing an at-a-glance overview of strengths and potential areas for growth, the possibility for tracking growth over time, and guidance on framing their teaching effectiveness for others.

The researchers believe that this study points to the CTB-MFI as a valid and reliable midterm feedback instrument that has the potential to add value for instructors and students. While this study offers promising insights into the reliability and validity of the CTB-MFI based on a

thorough analysis of multiple quantitative and qualitative data points, including feedback from faculty and students as well as statistical analysis of data, there are some limitations that should be recognized. All data was collected at a single, private STEM-focused institution in the Southeast. All participating courses were taught in a face-to-face modality. We encourage potential users to verify the reliability of this tool for other instructional modalities and with local student populations.

Conclusion

Midterm feedback is a process that provides faculty with information concerning their teaching effectiveness and the student learning experience. It has been proven to positively impact faculty behavior and pedagogy in the classroom (Diamond, 2004; Knol et al., 2013) and to have a lasting impact on student impression of the course as evidenced by increased student motivation (Lewis, 2001; Redmond, 1982; Svinicki, 2001), engagement (Hurney et al., 2021), and higher EOC evaluations (Bubb et al., 2013; Cohen, 1980; Knol et al., 2013). Few validated instruments exist (Donlan & Byrne, 2020; Hampton, 2000). The CTB-MFI fills an important gap by specifically soliciting focused feedback on teaching behaviors under the instructor's control. This focus intends to minimize the impact of potential student biases related to faculty identity on the evaluation of instruction. Furthermore, the focus on evidence-based categories of instructional behaviors provides a concrete language that students, faculty, and administrators can use to discuss effective teaching across disciplines and offers a framework faculty can use to document their teaching practice and achievements.

This study indicated many opportunities for future research. Of particular interest is the expansion of a pilot study to identify whether CTB-MFI feedback items, which focus on specific, observable teaching behaviors, can help reduce the negative impact of implicit biases on

student ratings of instruction for minoritized faculty groups. Additionally, future research might explore how the CTB-MFI impacts faculty decision to change or refine their teaching practice and conduct a validation study of the tool for courses delivered in hybrid and/or fully online modalities.

Midterm feedback empowers students with a sense of agency over shaping their learning experience and allows instructors to gain insights into the student experience to address miscommunications and make timely adjustments as appropriate. By focusing on specific, observable instructional behaviors associated with teaching that fosters student success, the CTB-MFI supports the creation of a shared frame of reference and common language to discuss teaching effectiveness. Furthermore, the CTB-MFI provides faculty with actionable data to refine their courses, diversify and strengthen their teaching portfolio, and clearly document their teaching for promotion and tenure portfolios.

Ethical Considerations

This study was approved by the Institutional Review Board (IRB) at Embry-Riddle Aeronautical University. Participation in the present study was voluntary, and participants were free to withdraw at any time. Prior to participation, participants were informed of the purpose of the study and that their responses were anonymous (survey) or confidential (focus group). All focus group participants gave written informed consent.

Biographies

Claudia Cornejo Happel, PhD, EdS, is the Director of the Center for Teaching and Learning Excellence at Embry-Riddle Aeronautical University, Daytona Beach. Her research interests include documenting and assessing teaching, scholarship of teaching and learning (SoTL),

and inclusive instructional practices. Her co-authored book with Lauren Barbeau, *Critical Teaching Behaviors: Defining, Documenting, and Discussing Good Teaching* (Routledge, 2023), builds on her expertise and interests to provide evidence-based guidance and practical tools for faculty to enhance teaching effectiveness and document instructional achievements.

Chad Rohrbacher, EdD, is the Senior Associate Director at the Center for Teaching and Learning Excellence at Embry-Riddle Aeronautical University, Daytona Beach. He works primarily with the College of Engineering and coordinates faculty mentoring initiatives across campus. His research areas include faculty classroom assessment, faculty peer observations, and the scholarship of teaching and learning (SoTL).

Teha Cooks, PhD, is an Associate Director of the Center for Teaching and Learning Excellence at Embry-Riddle Aeronautical University, Daytona Beach. She earned a PhD in Curriculum and Instruction (STEM education) from Texas Tech University and has taught science courses in K–12. Dr. Cooks's interests include experiential and project-based learning, pedagogical partnerships, mentoring GTAs, scientific argumentation, and global collaboration.

Jenna Korentsides, MS, is a PhD candidate at Embry-Riddle Aeronautical University, Daytona Beach. She works as a graduate research assistant in the Small Teams Analog Research (STAR) lab under the advisement of Dr. Joseph R. Keebler. Her research expertise expands from experimental and applied psychological research and statistical analysis to usability and user experience studies. Jenna also focuses on research topics related to long-duration space exploration, human-machine/human-AI interaction, teamwork, training, technological innovation, healthcare, and transportation systems.

Joseph R. Keebler, PhD, has over 15 years of experience conducting experimental and applied research in human factors, with a specific

focus on training and teamwork in medical, military, and consumer domains. He has partnered with multiple agencies and has led projects aimed at the implementation of HF/E in complex, high-risk systems to increase safety and human performance. This work includes command and control of tele-operated unmanned systems, communication and teamwork in medical systems, and simulation-/game-based training for advanced skills including playing guitar and identifying combat vehicles. His work includes over 170 scientific works, including over 60 peer-reviewed journal publications.

Carmen Van Ommen is a PhD candidate at Embry-Riddle Aeronautical University, Daytona Beach. She works in Dr. Barbara Chaparro's Research in User eXperience (RUX) lab, where her research is mostly focused on user experience research with a concentration in accessibility and gaming. Her recent projects include evaluating video game satisfaction with people with disabilities, the user experience of an adaptive game controller, and development of a scale to assess perceptions of technology product inclusivity.

Kimberly N. Williams, MS, is a PhD candidate at Embry-Riddle Aeronautical University, Daytona Beach. She conducts research in the Research Engineering and Applied Collaborations in Healthcare (REACH) lab under the advisement of Dr. Elizabeth Lazzara, where she predominantly focuses her research on training and assessment of learners in healthcare professions. She has worked on many projects that span psychometric development and user experience, including usability studies of educational tools and medical devices.

Lauren Barbeau, PhD, is the Assistant Director for Learning and Technology Initiatives in the Center for Teaching and Learning at the Georgia Institute of Technology. Her co-authored book with Claudia Cornejo Happel, *Critical Teaching Behaviors: Defining, Documenting, and Discussing Good Teaching* (Routledge, 2023), offers instructors a framework for identifying, implementing, and documenting effective teaching

behaviors as well as aligned peer observation and student feedback instruments to help them gather external perspectives on their teaching. Lauren earned her PhD in English from Washington University in St. Louis.

Conflict of Interest Statement

No potential conflict of interest was reported by the authors.

Data Availability

The data reported in this manuscript are available on request by contacting the corresponding author.

References

- Barbeau, L., & Cornejo Happel, C. (2023). *Critical teaching behaviors: Defining, documenting, and discussing good teaching*. Routledge.
- Boatright-Horowitz, S. L., & Soeung, S. (2009). Teaching White privilege to White students can mean saying good-bye to positive student evaluations. *American Psychologist*, 64(6), 574–575. <https://doi.org/10.1037/a0016593>
- Boring, A., Ottoboni, K., & Stark, P. (2016, January 7). Student evaluations of teaching (mostly) do not measure teaching effectiveness. *ScienceOpen Research*. <https://doi.org/10.14293/S2199-1006.1.SOR-EDU>
- Bubb, D. K., Schraw, G., James, D. E., Brents, B. G., Kaalberg, K. F., Marchand, G. C., Amy, P., & Cammett, A. (2013). Making the case for formative assessment: How it improves student engagement and faculty summative course evaluations. *Assessment Update*, 25(3), 8–12.
- Byrne, B. M. (2016). *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. Routledge.
- Clayson, D. (2022). The student evaluation of teaching and likability: What the evaluations actually measure. *Assessment & Evaluation in Higher Education*, 47(2), 313–326. <https://doi.org/10.1080/02602938.2021.1909702>
- Cohen, P. A. (1980). Effectiveness of student-rating feedback for improving college instruction: A meta-analysis of findings. *Research in Higher Education*, 13(4), 321–341. <https://doi.org/10.1007/BF00976252>

- Diamond, M. R. (2004). The usefulness of structured mid-term feedback as a catalyst for change in higher education classes. *Active Learning in Higher Education*, 5(3), 217–231.
- Donlan, A. E., & Byrne, V. L. (2020). Confirming the factor structure of a research-based mid-semester evaluation of college teaching. *Journal of Psychoeducational Assessment*, 38(7), 866–881. <https://doi.org/10.1177/0734282920903165>
- Fan, Y., Shepherd, L. J., Slavich, E., Waters, D., Stone, M., Abel, R., & Johnston, E. L. (2019). Gender and cultural bias in student evaluations: Why representation matters. *PloS ONE*, 14(2), e0209749. <https://doi.org/10.1371/journal.pone.0209749>
- Hampton, S. E. (2000, February 16–20). *A review of literature on formative evaluation of teachers through mid-term student feedback and how the Reiser and Dick Instructional Planning Model can enhance this feedback* [Paper presentation]. Association for Educational Communications and Technology International Convention, Long Beach, CA, United States.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Hurney, C. A., Renner, C. M., & Troisi, J. D. (2021). *Midcourse correction for the college classroom: Putting small group instructional diagnosis to work*. Routledge.
- Knol, M. H., in't Veld, R., Vorst, H. C. M., van Driel, J. H., & Mellenbergh, G. J. (2013). Experimental effects of student evaluations coupled with collaborative consultation on college professors' instructional skills. *Research in Higher Education*, 54(8), 825–850. <https://doi.org/10.1007/s11162-013-9298-3>
- Kreitzer, R. J., & Sweet-Cushman, J. (2022). Evaluating student evaluations of teaching: A review of measurement and equity bias in SETs and recommendations for ethical reform. *Journal of Academic Ethics*, 20, 73–84.
- Lewis, K. G. (2001). Using midsemester student feedback and responding to it. *New Directions for Teaching and Learning*, 2001(87), 33–44. <https://doi.org/10.1002/tl.26>
- Mauger, D. (2010). *Small group instructional feedback: A student perspective of its impact on the teaching and learning environment* [Doctoral dissertation, George Fox University]. UMI Dissertation Publishing, ProQuest. <https://www.proquest.com/docview/305248543/abstract/8C25D04C818C4F09PQ/1>
- McKeachie, W. J. (1997). Student ratings: The validity of use. *American Psychologist*, 52(11), 1218–1225. <https://doi.org/10.1037/0003-066X.52.11.1218>

- Mitchell, K. M. W., & Martin, J. (2018). Gender bias in student evaluations. *PS: Political Science & Politics*, 51(3), 648–652. <https://doi.org/10.1017/S104909651800001X>
- Murray, H. G. (1984). The impact of formative and summative evaluation of teaching in North American universities. *Assessment & Evaluation in Higher Education*, 9(2), 117–132. <https://doi.org/10.1080/0260293840090204>
- Murray, H. G. (1997, March). *Classroom teaching behaviors and student instructional ratings: How do good teachers teach?* [McKeachie Award address]. 78th Annual Meeting of the American Educational Research Association, Chicago, IL, United States.
- Redmond, M. V. (1982). *A process of midterm evaluation incorporating small group discussion of a course and its effect on student motivation*. <https://files.eric.ed.gov/fulltext/ED217953.pdf>
- Reid, L. D. (2010). The role of perceived race and gender in the evaluation of college teaching on RateMyProfessors.Com. *Journal of Diversity in Higher Education*, 3(3), 137–152. <https://doi.org/10.1037/a0019865>
- Reiser, R. A., & Dick, W. (1996). *Instructional planning: A guide for teachers*. Allyn and Bacon.
- Svinicki, M. D. (2001). Encouraging your students to give feedback. *New Directions for Teaching and Learning*, 2001(87), 17–24. <https://doi.org/10.1002/tl.24>
- van Wyk, J., & Mclean, M. (2007). Maximizing the value of feedback for individual facilitator and faculty development in a problem-based learning curriculum. *Medical Teacher*, 29(1), e26–e31. <https://doi.org/10.1080/01421590601032435>
- Veeck, A., O'Reilly, K., MacMillan, A., & Yu, H. (2016). The use of collaborative midterm student evaluations to provide actionable results. *Journal of Marketing Education*, 38(3), 157–169. <https://doi.org/10.1177/0273475315619652>

Appendix A: CTB-MFI Categories and Questions

Category	Definition	CTB-MFI items	Code
Align	Instructors who align components of learning experiences start with clear learning goals. Measurable outcomes, teaching and learning activities, assessment tasks, and feedback build on each other to support student progress toward these goals.	States the learning outcomes (development of specific skills and knowledge) to be accomplished in the course assignments and activities	LEARNING_OUTCOMES
		Uses time effectively and efficiently toward achievement of course learning outcomes	TIME_MANAGEMENT
		Gives exams and assignments that reflect course readings, lectures, and class activities	ALIGNED_ASSESSMENT
Include	Instructors who create an inclusive learning environment promote equity by using accessibility standards and learner-centered strategies when designing and delivering content. They cultivate an atmosphere in which students see themselves positively represented and experience a sense of belonging conducive to emotional well-being for learning.	Invites students' questions, examples, and experiences and listens carefully when students speak	STUDENT_PERSPECTIVE
		Selects examples and activities that represent a range of perspectives and experiences	RANGE_PERSPECTIVES
		Builds community and trust between students	COMMUNITY
Engage	Instructors who engage students purposefully select research-based techniques to ensure that students actively participate in the learning process and take responsibility for their intellectual development.	Encourages participation from all students through meaningful individual and/or small group activities in the classroom and/or online	PARTICIPATION
		Connects content to real-life applications and examples and/or current research in the field	REAL_APPLICATION
		Establishes regular and open communication	COMMUNICATION

(Continued)

(Continued)

Category	Definition	CTB-MFI items	Code
Assess	Instructors who assess learning develop and facilitate transparent, meaningful tasks to provide students with timely feedback on their learning and to measure achievement of learning outcomes. They frequently review data to improve instruction.	Gives timely and specific feedback that helps you improve on future assignments	FEEDBACK
		Schedules regular tasks (quizzes, homework, discussions, project drafts, etc.) that help you prepare for bigger assignments	SCAFFOLDING
		Clearly communicates how to succeed on assessments by providing grading criteria or examples	TRANSPARENT_EXPECTATIONS
Integrate Technology	Instructors who integrate technology responsibly use tools to design accessible, high-quality instructional materials and engaging learning opportunities beyond traditional barriers of place and time.	Shares course materials on the online learning platform in a way that makes it easy to find and access them	ONLINE_ORGANIZATON
		Uses technologies and/or apps that enhance your learning experience in the course (e.g., Canvas, multimedia content, polling)	TECH_ENHANCE
		Trains students to use course technology/apps and provides support	TECH_TRAINING < TechIntegration

Appendix B: CTB-MFI Suggested Process

- **INITIATE MIDTERM FEEDBACK:** The midterm feedback process is initiated by the instructor who invites a peer or educational developer into class.
- **FACILITATE CTB-MFI SYNCHRONOUSLY:** To increase response rates, facilitate the survey during class meeting time.
- **CONTEXTUALIZE FOR STUDENTS:** The consultant describes the entire process prior to providing the instrument to students; information shared includes a reminder that feedback is anonymous for students and confidential for faculty; participation in the CTB-MFI is voluntary for faculty—faculty opt in to this process because they are interested in understanding and improving the student learning experience in their course.
- **ENSURE ANONYMITY OF STUDENT RESPONSES:** To ensure anonymity of student responses, the faculty member should not be present while students complete the CTB-MFI. Consultants analyze quantitative data and summarize main themes emerging from qualitative student responses to further ensure students cannot be identified from their individual responses.
- **CREATE REPORT:** Consultants summarize qualitative data and capture recurring themes along with relevant examples of student comments. Using an Excel spreadsheet, consultants tabulate quantitative data and create a summary table and associated bar graph showing the distribution of responses as well as average scores for each survey item.
- **DISCUSS FEEDBACK:** Consultant provides feedback directly to the faculty member in written form and schedules time to discuss and debrief insights. During the debrief session, the consultant provides an overview of the CTB categories and discusses quantitative feedback in conjunction with relevant student comments and themes emerging from qualitative responses. Consultant and instructor work together to determine how to respond to the feedback that was received. The debrief should be a collaborative, meaning-making opportunity that ultimately identifies at least one practice to refine.

- **CLOSE THE LOOP:** Faculty should close the assessment loop by having a conversation with their class to discuss what they learned from feedback, address misunderstandings, clarify expectations, and explain changes they plan to make in response to feedback.

Appendix C: Summary of Student Survey Responses

CTB category	CTB-MFI item	<i>Synthesis of responses:</i> What does this mean to you? Restate what you think this item is asking you to evaluate.	<i>Synthesis of responses:</i> What evidence (e.g., in-class behaviors, course materials) would you consider when evaluating your instructors' efforts in this area?
Align	States the learning outcomes (development of specific skills and knowledge) to be accomplished in the course assignments and activities	Instructors present an overview of learning objectives/outcomes and "goals of what learning in this class should result in." They give students a sense of "what we will learn throughout the semester" and what "I'm meant to get out of an activity or assignment."	Students look for evidence primarily in the course syllabus but also in assignment prompts and "everything requiring submission of work."
	Uses time effectively and efficiently toward achievement of course learning outcomes	Instructors "use class time to do things that bring value to students" and discuss curriculum that "could be seen on the final exam" and other course assessments. Instructors are respectful of students' time and do not "waste time" through poor time management (e.g., "first couple of minutes spent on turning everything on") or going off topic.	Students look for evidence primarily in the relation between how class time is spent and what is assessed. They also consider the overall pace of the course and whether they are released on time. One student stated, "I usually evaluate this based on how long it takes for me to start writing notes in class . . . i.e., how long until I actually need to pay attention, and how long my attention is maintained."
	Gives exams and assignments that reflect course readings, lectures, and class activities	Instructors assign work that is "relevant to the learning outcomes and topics in the classroom" and provide the "means and materials to perform well" through in-class lecture, discussion, and assigned materials.	Students look for evidence primarily in scaffolding materials provided by the instructor, whether that is clear and focused instructions, relevant homework, guided reviews, or the availability of models or "examples of assignments that convey expected work."

(Continued)

(Continued)

CTB category	CTB-MFI item	Synthesis of responses: What does this mean to you? Restate what you think this item is asking you to evaluate.	Synthesis of responses: What evidence (e.g., in-class behaviors, course materials) would you consider when evaluating your instructors' efforts in this area?
Include	Invites students' questions, examples, and experiences and listens carefully when students speak	Instructors are open to conversation, receptive to student feedback, and show their care by "tak[ing] the time to listen to their students."	Students look for evidence primarily in instructor availability outside of class time, for example, in-office hours and responsiveness to emails. They also consider the extent to which instructors encourage student questions and "participation beyond being able to raise your hand in the classroom."
	Selects examples and activities that represent a range of perspectives and experiences	Instructors incorporate "a multitude of examples and applications that can benefit the entire class" regardless of chosen major or career path. Instructors go beyond one-sided narratives and provide multiple, diverse sources of information.	Students look for evidence in the assigned materials, examples, and applications incorporated into class—Do they represent multiple perspectives and consider diversity of students' goals?
	Builds community and trust between students	Instructors are "trustworthy" and "personable with students" and promote teamwork. They create an environment where students have an opportunity to build relationships with peers and the instructor and feel comfortable participating and reaching out.	Students look for evidence in instructor interactions with students: Do instructors remember student names? Do they make an effort to connect with students by sharing a little about themselves and asking for student feedback? Do they find a balance between serious and light-hearted interactions? Additionally, students consider whether group activities and class-led discussion are a part of the course.

(Continued)

(Continued)

CTB category	CTB-MFI item	Synthesis of responses: What does this mean to you? Restate what you think this item is asking you to evaluate.	Synthesis of responses: What evidence (e.g., in-class behaviors, course materials) would you consider when evaluating your instructors' efforts in this area?
Engage	Encourages participation from all students by incorporating meaningful individual and/or small group activities in the classroom and/or online	Instructors "have in-class activities that extend beyond the lecture" and create opportunities for students to "share their experiences, knowledge, and wisdom in an appropriate forum" either in class or online.	Students look for evidence primarily in the presence of activities such as small group discussions, group projects, student-led presentations, and online discussions.
	Connects content to real-life applications and examples and/or current research in the field	Instructors explicitly "apply learning outcomes to real-life examples" and "mak[e] connections to further help understand topics in class and how what we learn can apply to our careers." They incorporate recent events and address contemporary and emerging research in the field.	Students look for evidence primarily in the integration of application examples/ stories from industry experience, discussion of case studies, and the presence of hands-on activities. Student responses point out that connections need to be made explicit by the instructor, as "if students think their work is meaningless and a waste of time, they are not going to put as high-quality work into it."
	Establishes regular and open communication	Instructors are available and approachable to answer student questions in and outside the classroom, e.g., in-office hours or email.	Students look for evidence primarily in instructor availability during office hours and timely response to emails.

(Continued)

(Continued)

CTB category	CTB-MFI item	Synthesis of responses: What does this mean to you? Restate what you think this item is asking you to evaluate.	Synthesis of responses: What evidence (e.g., in-class behaviors, course materials) would you consider when evaluating your instructors' efforts in this area?
Assess	Gives timely and specific feedback that helps you improve on future assignments	Instructors return graded work within a few days of the submission deadline—"at least the amount of time students were given notice of the assignment"—and provide clear, specific feedback that "can lead to you actually being able to improve future coursework."	Students look for evidence primarily in the timely return of graded work—at a minimum, graded work should be returned "well in advance of any other major assignment/exam deadlines so that we have the opportunity to learn from our mistakes." Specifically, students are also looking for feedback providing "an explanation of what went wrong" if their work did not meet expectations.
	Schedules regular tasks (quizzes, homework, discussions, project drafts, etc.) that help you prepare for bigger assignments	Instructors scaffold work to "build up your ability and knowledge to be able to accomplish a more extensive assignment" and promote "incremental progress to the final grade." Regular assignments and homework "help [students] understand classroom topics."	Students look for evidence in the number, distribution, and point value of assignments across the semester—Are there "plenty of materials to prepare you for those large class benchmarks?"
	Clearly communicates how to succeed on assessments by providing grading criteria or examples	Instructors clearly define and explain grading criteria prior to due dates. They provide rubrics and discuss examples.	Students look for evidence in the availability of rubrics and examples of student work, which ideally are "clear and discussed ahead of time."

(Continued)

(Continued)

CTB category	CTB-MFI item	Synthesis of responses: What does this mean to you? Restate what you think this item is asking you to evaluate.	Synthesis of responses: What evidence (e.g., in-class behaviors, course materials) would you consider when evaluating your instructors' efforts in this area?
Integrate Technology	Shares course materials on the online learning platform in a way that makes it easy to find and access them	Instructors provide supplementary and required materials on the LMS; they are organized in a way that "makes sense" and "it is easy to differentiate between assignments, course materials, in-class notes, etc."	Students look for evidence in the LMS presence of a course, and they consider both the availability of materials (e.g., recorded lectures, lecture slides/ notes, supplementary materials) and the ease of use (e.g., modules/materials available early, transparent organization).
	Uses technologies and/or apps that enhance your learning experience in the course (e.g., LMS, multimedia content, polling)	Instructors use technology to supplement lectures and provide alternative or "novel ways to share and interact [with] course materials" to enhance learning and communication.	Students look for evidence in instructors' use of technology tools outside the LMS. Some examples mentioned include polling apps, Kahoot, YouTube, TikTok, etc.
	Trains students to use course technology/apps and provides support	Instructors provide "proper support and instructions [so] technology isn't a barrier to completing the work." This includes "go[ing] through the [LMS] page . . . to show students the 'ropes' of the site."	Students look for evidence in materials provided by the instructor, such as "written directions," "video instructions," and/or "demonstrations."