Transforming Teaching Evaluation:
A Toolkit of Resources for Departmental Level Change

Sarah E. Andrews, Alanna Pawlak, Cynthia Hampton, Dena Rezaei, Noah D. Finkelstein

[Last Update: 12/17/2021]

This Toolkit is in beta testing, and we welcome your feedback. If you would like to share your thoughts, please contact Sarah Andrews at sarah.andrews-1@colorado.edu.

This resource is part of the TQF project at CU Boulder, which is supported by NSF (DUE-1725959), the Association of American Universities, and CU Boulder Colleges of Arts & Sciences and Engineering & Applied Science. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF or other funding sources. For more information about the Teaching Quality Framework, please visit our website: www.colorado.edu/teaching-quality-framework. See also http://teval.net/ for information about the broader NSF-funded multi-institutional teaching evaluation project.

Table of Contents

This Toolkit is in beta testing, and we welcome your feedback. If you would like to share your thoughts, please contact Sarah Andrews at sarah.andrews-1@colorado.edu.

Overview of the Toolkit

Overall Purpose

This teaching evaluation toolkit provides resources and guidance for departments interested in assessing teaching in more robust ways. The toolkit draws from the work of the Teaching Quality Framework (TQF) initiative at the University of Colorado Boulder, and is designed to help individuals, groups, and departments implement transparent, scholarly approaches to quality teaching assessment for formative development of teaching and for summative teaching evaluation for merit, reappointment, promotion, and tenure processes.

This Toolkit focuses on resources for departments engaged in this work, however we acknowledge that it will be important for any departmental level efforts to be connected to campus-level work. Thus, an institutional level toolkit that focuses on guidance for campus / system level transformation of teaching evaluation is forthcoming.

How to Use the Toolkit

This toolkit is written in two main sections. The first section focuses on supporting departmental teams through the ***processes*** of transforming their teaching evaluation practices. The second section provides ***resources*** departments may find helpful as they go through those processes. The toolkit has been organized this way for the sake of clarity, however we do not assume that you will read through the entirety of this toolkit in order. While some could follow it sequentially to develop a working group and overhaul their teaching evaluation systems, we encourage you to peruse the table of contents and use the sections that are most helpful to you at each stage in your work. Because this toolkit is not designed to be read from start to finish in order, some information may appear in multiple places, and sections often refer to one another.

Table 1 below provides some examples that indicate which parts of the toolkit might be useful based on your goals. If you would like to learn more about the Teaching Quality Framework Initiative and how this toolkit was developed, you can find that information in the Appendix. There you can also find suggestions for finding additional resources on your campus.

This Toolkit is in beta testing, and we welcome your feedback. If you would like to share your thoughts, please contact Sarah Andrews at sarah.andrews-1@colorado.edu.

Table 1. Relevant Toolkit Sections for Various Goals

<table>
<tr><th>Goal</th><th>Location for Processes</th><th>Location for Resources</th></tr>
<tr><td>I need to determine if my department is ready to reconsider teaching evaluation practices</td><td>Assess Departmental Readiness</td><td>TQF approaches for cultivating interest and TQF readiness for change (external links)</td></tr>
<tr><td>I want to create a working group to reconsider teaching evaluation practices in my department / unit</td><td>Form a working group</td><td>TQF Guidelines for forming departmental working groups and TQF methods for recruiting group members (external links)</td></tr>
<tr><td>I want to create a tool for peer evaluation of teaching</td><td rowspan="3">Adopt/Adapt New Tools Across Three Voices</td><td>Peer Voice: Developing a Holistic Peer Course Evaluation Plan</td></tr>
<tr><td>I want to create a tool to guide faculty self-reflection on teaching</td><td>Self Voice: Meaningful Self-reflection on Teaching</td></tr>
<tr><td>I want to create a tool for students to evaluate faculty teaching</td><td>Student Voice: Using multiple measures to collect student feedback</td></tr>
<tr><td>I want to learn how to facilitate and manage working meetings</td><td>Set Team Responsibilities
Establish Meeting Logistics
Develop Meeting Structures and Norms</td><td>Meeting Resources</td></tr>
<tr><td>I have multiple tools for evaluating teaching but want to learn how to create a rubric for assessing a dossier of evidence generated by these tools</td><td>Develop a Complete Teaching Evaluation System</td><td>Adapting the rubric into a summative teaching evaluation tool</td></tr>
</table>

Department Level Processes

In this section we describe several potential steps that departmental working groups could engage in as part of the process of improving departmental evaluation of teaching. While these steps are based on our experiences facilitating multiple working groups, and represent our best recommendations for a successful approach, they are meant to be flexible and adaptable to departmental goals. We have divided these steps into two parts: ***Preparation Processes*** and ***Working Group Processes***.

In Preparation Processes, we provide guidelines for assessing departmental readiness, forming a working group, setting team responsibilities, and establishing meeting logistics. We also provide a curated list of recommended reading / resources that working group members may want to review in advance of meeting for the first time.

In Working Group Processes, we provide recommendations for developing meeting structures and norms and outline five key processes for transforming teaching evaluation:

1) Externalize current teaching evaluation practices
2) Assess needs and set goals
3) Adapt new tools across three voices
4) Implement tools and processes in the department
5) Develop a complete teaching evaluation system

For each of these steps we include expected outcomes, suggested preparation, and guiding questions for discussion. Throughout this document we have provided estimates of time where appropriate. However, any timeline given is an example and is dependent on the department and context. While we have organized these steps in one possible sequence, we note that there are many other possible routes of action a working group might take depending on their goals and department culture (see for example, the table of "Relevant Toolkit Sections for Various Goals" above).

Preparation Processes

The preparation process is important to consider before beginning the work of transforming teaching evaluation. Note, even if your department already has formed a working group, we recommend at least reading the following sections: set team responsibilities, establish meeting logistics, and conduct recommended preparatory reading.

- ▢ Assess department readiness
- ▢ Establish a working group
- ▢ Set team responsibilities
- ▢ Establish meeting logistics
- ▢ Conduct recommended preparatory reading

Assess Departmental Readiness

To assess whether or not a department is ready to move forward with the work of transforming teaching evaluation, the department chair and/or any person in the department leading the teaching evaluation initiative should discuss and address the following questions:

- Does the department have enough interested faculty to recruit 2-3 working group members?
- Does the department (and/or college/university) have a proper plan and resources to recognize working group members' effort (e.g., monetary reward, service credit)?
- Where does the department stand in terms of teaching evaluation and where does it need to be?
- What is the proper time frame for implementing the teaching evaluation initiative in your department?
- See in particular TQF approaches for cultivating interest and TQF readiness for change.

Form a Working Group

After the department has confirmed they have enough interest and resources to engage, the next step is to put together a working group. Thinking strategically about who will participate in the working group and how those participants will be rewarded is crucial. Departments have utilized several methods for recruitment that include (from TQF methods for recruiting group members):

- Formation out of existing committees or structures (e.g., a Faculty Learning Community (FLC) or an Undergraduate Studies Committee)
- A department-wide call for volunteers, taking whoever was interested (attending to characteristics in TQF Guidelines for forming departmental working groups, where possible)
- Assignment to the team by the department chair
- Individual recruiting by the department lead (i.e., the person coordinating with TQF central)
- Some combination of the above (e.g., a department-wide call followed by assignment by a chair or recruiting additional members to an existing committee or structure)

We suggest that working group members include a mix of full professors, junior faculty, and members of key campus teaching initiatives to allow for a wide representation of views. Refer to TQF Guidelines for forming departmental working groups for additional guidance on forming teams. In particular, it is helpful for at least one member to have connections to existing committees in charge of merit raises, coordination of faculty peer observations, or other departmental teaching evaluation functions.

Working group members should be provided with incentives to participate through various mechanisms such as teaching credit, service credit, etc., and the explicit support of their department chair.

Set Team Responsibilities

Once a team has been formed and any reward structures are in place, before meeting (or in a first meeting) we recommend establishing some team responsibilities and ensuring that each team member has a role to play. As you are setting team responsibilities, work to ensure that work is equitably divided. Unless you have an external facilitator, here are some responsibilities/roles you might discuss. Who will:

- Reserve a meeting room or set up a virtual room?
- Put together agendas?
- Take notes during the meeting?
- Ensure that all participants have a voice in the meeting?
- Send reminders?
- Follow-up with the group on homework goals?
- Other departmental specific considerations?

Establish Meeting Logistics

Along with setting team responsibilities, meeting logistics should be discussed early on, with special attention paid to how you will monitor and maintain a productive working culture. Below are some items you might wish to discuss. Note that some of these items are discussed in more detail in the Working Group Process section.

- Meeting frequency (momentum, progress)
- Structures for agendas, note taking, collecting resources etc. (e.g., Google Drive) [note/link to examples in resources]
- For in person meetings, reserving a room with appropriate space (consider accessibility) and resources (e.g., projector, whiteboard)
- For virtual meetings, establishing a platform (e.g., Zoom, Microsoft Teams)
- Set standards for asynchronous work (e.g., agree upon degree of in-between meeting work)
- Pathways for regular communication (e.g., email, Microsoft Teams, Slack, co-editing Google documents, etc.)
- Community standards, developing collaborative norms, constructive discussions of contentious ideas

Recommended preparatory reading

Numerous forms of literature are available to aid in understanding and familiarizing team members with transforming teaching evaluation. In particular, the suggested readings below include information on the three voices, TQF rubric, and types of assessment. These readings can also be shared when developing departmental teams as a form of orienting potential team members to the TQF. Context-specific reading may be needed to orient working group members to departmental or unit processes related to internal processes for reappointment, promotion, tenure, and other pertinent evaluation processes that currently occur within said department or unit.

- What is the Teaching Quality Framework?
    - Teaching Quality Framework 1- pager that provides a very brief overview of our framework and overall process
    - TQF Rubric
    - Sample tools across Three-voices (student voices, peer evaluation, self-voice)
    - TQF FAQs

This Toolkit is in beta testing, and we welcome your feedback. If you would like to share your thoughts, please contact Sarah Andrews at sarah.andrews-1@colorado.edu.

    - [Transforming Teaching Evaluation in Disciplines: A Model and Case Study of Departmental Change]

These readings can help to orient working group members around common knowledge and language to begin the working group process.

Working Group Processes

In this section, we outline the process through which working groups transform their units' teaching evaluation practices. The process is described by the six steps below. These steps don't necessarily need to be done in the order presented; in particular, teams may find shifting the order of the last three steps to be productive depending on their local context.

- [ ] Develop meeting structures and norms
- [ ] Externalize current teaching evaluation practices
- [ ] Assess needs and set goals for teaching evaluation transformation
- [ ] Adopt / adapt new tools across three voices
- [ ] Implement tools and processes in the department
- [ ] Develop a complete teaching evaluation system

Develop Meeting Structures and Norms

Throughout your team's work, it is important to consider how your team will collaborate during meetings and how you will structure your meetings. Meetings are typically comprised of a review of the meeting agenda, group discussion of questions and issues related to teaching evaluation, and sometimes, synchronous work on teaching evaluation tools and processes as they are developed. Below are some structures and guidelines for implementing team meetings.

Meeting Agendas
- Prepare an agenda (that is accessible to the team, e.g. via Google Drive, Dropbox, Canvas, etc.) for each meeting that includes the major discussion topics and how much time will be spent on them (see the Meeting Resources section for an example)
- Begin each meeting with a review of the meeting's agenda and an opportunity for team members to make additions

Group Discussions
- Prepare specific question prompts related to the discussion topic to help guide the conversation (the following sections include suggested questions tailored to each topic)
- Set clear goals for the discussion - is there a decision to be made? Are you brainstorming around a topic? Are you trying to choose a path forward?
- Allow enough time for all team members to contribute meaningfully to the discussion
- If some team members haven't shared their thoughts, consider asking them specifically for their ideas
- Set team norms for productive discussions and refer to them as needed
- If possible, have a team member or members take notes on the discussion
- Summarize and synthesize the outcome of the discussion for the team to ensure that everyone agrees on the takeaways and any decisions made

Synchronous Work
- Set clear goals for what work is to be accomplished together in the meeting
- Encourage team members to open the relevant materials on their own devices
- If some team members aren't sharing their thoughts, consider asking them specifically for their ideas
- Check in with team members while working to ensure that everyone agrees with the progress being made on your materials

Closing Meetings
- Reserve five minutes at the end of meetings to summarize what was accomplished and to set goals for the next meeting
- Make a note of any unfinished business or open questions to be revisited in future meetings

These guidelines represent on-going considerations for the structure and facilitation of your team meetings. In the following sections, we present specific topics you may wish to focus on in your team meetings.

Externalize Current Teaching Evaluation Practices

In your team's first meeting(s), we recommend focusing on ensuring that everyone in the working group understands the current teaching evaluation practices in your department. Depending on the transparency of these systems and how familiar all working group members are with the processes, this is expected to take 1-2 meetings.

Expected outcomes

- Everyone in the working group will:
  - understand the current teaching evaluation processes for both annual merit and reappointment, tenure, and promotion
  - understand strengths and weaknesses of the current teaching evaluation systems
- Collectively the team will be prepared to move on to assessing departmental needs and setting specific goals for changes to your teaching evaluation systems (next section)

Suggested materials / preparatory work

- Plan on making copies of teaching evaluation tools and processes available to all team members. At minimum these should be brought to the meeting or made available electronically for review in-meeting. Ideally everyone would review these materials in advance of the meeting. Examples include department policies for evaluating teaching for merit and reappointment, promotion, and tenure; existing tools used for evaluating teaching (e.g., guidelines around use of Student Evaluations of Teaching (SETs, or FCQs at CU Boulder), peer observation protocols, guidelines for writing teaching statements, etc.)

Guiding questions for discussion

▢ **What materials are submitted for annual review and reappointment, tenure, and promotion cases?**

- Which of these are required? Which, if any, are optional?
- How do these materials satisfy “multiple measures”?
- Are all three voices (student, peer, self) represented? If not, where are they missing?
- What are the specific protocols/tools that are used to generate/inform these materials?

▢ **How are submitted materials / data for annual review and reappointment, tenure, and promotion cases evaluated?**

- Is there a rubric or some other process in use that provides guidance on how these materials are used to assess teaching quality?
- What are the implicit and explicit criteria currently used by the department?

▢ **How transparent are your teaching evaluation processes?**

- Do faculty across ranks and roles, including non-tenure-track instructors, understand expectations for teaching evaluation?
- Are these expectations written and easily accessible to faculty in all ranks?
- Are there any aspects of your teaching evaluation processes that are not written down?
- If there are gaps in written processes/policies - is someone in the working group familiar enough with the process that they can explain it / write it down? If not, who do you need to talk to in order to understand how they are done?
- Are there other aspects of your teaching evaluation process that lack transparency?
- Could processes be made more transparent so faculty are clear about the process and data can be easily assessed and analyzed by faculty?

▢ **What are the strengths and weaknesses of your current teaching evaluation systems?**

- What are positive aspects you wouldn’t want to change in your current evaluation system? What’s working?
- Are there particular parts of the annual merit or reappointment, promotion, and tenure processes that cause tension in your department?
- To what extent do your current teaching evaluation practices encourage and reward the use of evidence-based teaching practices?
- To what extent do your current teaching evaluation practices encourage and reward equitable and inclusive teaching practices?

▢ **How malleable are the processes and tools?**

- Do you expect department resistance to changing policies around teaching evaluation? If so, are there tools that could be worked on (e.g., standard protocols for gathering data) that still align with policies?
- Do you have an internal system for annual review or is it tied to a university-wide system that you can’t really change at the department level? If the latter, are there small changes that you could make at the department level? Is there a process for making recommendations for changes at higher levels?

▢ **Are there specific processes for formative assessment in use in your department?**

- I.e., are there structured opportunities for faculty to get feedback on their teaching practices that aren’t required to be submitted for formal evaluations?

Assess Needs and Set Goals for Teaching Evaluation Transformation

After ensuring that your working group understands your current teaching evaluation practices (see preceding section "Externalize Current Teaching Evaluation Practices"), you should be prepared to assess specific department needs around teaching evaluation and set targeted goals for what you hope to achieve. This step is expected to take 1-2 meetings.

Expected outcomes

- Collectively the working group will:
  - Generate a list of ideas for particular aspects of the department's teaching evaluation systems (annual merit and reappointment, promotion, and tenure) that the group could focus on, i.e., targets for action
  - Have ranked these ideas in some order based on importance and/or feasibility, with specific goal(s) for immediate next steps
  - Be prepared to begin work adopting or adapting new tools/processes across three voices (see next section) or adopting/adapting an assessment rubric for your unit (see section in Resources)
  - A "parking lot" of these ideas can be housed in the meeting notes (see the Meeting Resources section for an example)

Suggested materials / preparatory work

- Ensure that the full team understands the department's current annual merit and reappointment, tenure, and promotion systems (see "Externalize Current Teaching Evaluation Practices" above)
- Review the TQF Framework model and the full TQF Rubric, and be familiar with the seven dimensions of quality teaching. See the "TQF rubric" resource below.
- Have at hand any departmental/unit policies and procedures for teaching evaluation and copies of any tools you currently use (e.g., a peer observation protocol, a classroom interview protocol, a list of questions asked on student evaluations of teaching (SETs; known as FCQs at CU), a solicitation you use for getting student letters, etc.).
- Create an "idea parking lot" to list out all of your ideas and help you prioritize them
- Consider using the "Identifying Gaps in Current Teaching Evaluation Systems" resource below to help you assess needs and set goals in a way that aligns with the TQF framework.

Guiding questions for discussion

▢ **Based on your conversations, could you generate a list of potential targets for action within annual review and reappointment, tenure, and promotion?**

- What specific teaching evaluation tools would you like to revise, adapt, or create?
  - What procedures need to be developed around the use of these tools?
  - What gap(s) in your teaching evaluation system are you intending to fill using these new/revised tools? For example,

This Toolkit is in beta testing, and we welcome your feedback. If you would like to share your thoughts, please contact Sarah Andrews at sarah.andrews-1@colorado.edu.

        - Will this tool fill in a gap in one of the voices? Will this tool provide better / more information than an existing tool? (in terms of quality of information and/or in terms of aligning w/ the TQF rubric).
        - If SETs/FCQs are overly emphasized, how can additional forms of student voice and other voices be better incorporated to create a more balanced system?
- What policies, if any, around teaching evaluation in your department would you like to recommend revisions for?
- Do you have a sense of whether you'd like to begin with annual merit evaluations or reappointment, tenure, and promotion? What are the pros and cons of selecting one versus the other to start with?

▢ **Once you've generated a list of ideas for what you'd like to change, how will you prioritize those goals?**

- E.g., is there a particular tool/process that is most time sensitive? Or something that is perceived (by your working group and/or department wide) as most important? Or something that is otherwise low-hanging fruit that you believe would be comparatively easy to implement in your department?
- What goals do you hope to achieve by the end of the current term? Year?

▢ **How will you structure your time, meetings, and resources to achieve your short and long term goals?**

- If not already decided, how often will you meet? What amount of time can you realistically devote to in-between-meeting work?
- What structures will you use to make progress (e.g., shared google docs for asynchronous work)?

Adopt/Adapt New Tools Across Three Voices

After assessing your department's needs and goals, you can begin work on developing and adapting tools and processes that are aligned with the dimensions of quality teaching and span the three voices (self, student, and peer). As you finish tools and processes, consider whether you would like to next: a) implement them in the department, b) start working on a new tool or process, or c) begin pulling them together into a complete revised evaluation system. Developing a new tool or process is expected to take 1-4 meetings, depending on how heavily you pull from existing resources.

Expected outcomes

- Collectively the working group will:
    - Create new tools or adapt existing tools to assess teaching quality that include the self voice, student voice, and peer voice
    - Articulate processes for how these tools will be implemented
    - Be prepared to create a complete revised teaching evaluation system
    - Be prepared to implement teaching evaluation tools and processes in the department

Suggested materials / preparatory work

- Have assessed departmental needs with respect to teaching evaluation, set goals, and prioritized these goals (see "Assess Needs and Set Goals for Teaching Evaluation Transformation" above)
- Review the TQF Framework model and full TQF Rubric, and be familiar with the seven dimensions of quality teaching. See the "TQF rubric" resource below.
- For the piece of teaching evaluation you are working on currently (based on the prioritization established in the "Assess Needs and Set Goals for Teaching Evaluation Transformation" step), review:
    - your department's existing tools and processes and have them available at your team's meetings.
    - the specific guidance relevant to your focus presented in the "Departmental Level Tools for Assessing Teaching Quality" section.
    - the sample evaluation tools and processes in the "Departmental Level Tools for Assessing Teaching Quality" section and have them available at your team's meetings.

Guiding questions for discussion

▢ **Present a few example teaching evaluation tools or processes (see the "Departmental Level Tools for Assessing Teaching Quality" section for a collection of samples). What are your thoughts on these examples?**

- In what ways do these align or not align with the needs and goals you've identified?
- In what ways do these assess the dimensions of quality teaching?
- Which do you prefer? Why?
- Which elements of these do you like? Which elements do you not like?
- Do any of these fit well enough that you'd like to adapt them to your context?
- If you'd like to adapt one of these to your context, what needs to be changed? What would you like to keep the same?

▢ **Present your department's existing teaching evaluation tools or processes. What are your thoughts on these?**

- In what ways do these align or not align with the needs and goals you've identified?
- In what ways do these assess the dimensions of quality teaching?
- Which elements of these do you like? Which elements do you not like?
- Do any of these align with your goals well enough that you'd like to revise them rather than replace them with new materials?
- If you'd like to revise an existing tool or process, what needs to be revised? What would you like to keep the same?

▢ **If you are developing a new tool or process (not adapting or revising an existing one), how will you ensure it meets your needs?**

- How will your tool or process assess the dimensions of quality teaching?
- How will your tool or process align with the goals you've identified with respect to teaching evaluation?

This Toolkit is in beta testing, and we welcome your feedback. If you would like to share your thoughts, please contact Sarah Andrews at sarah.andrews-1@colorado.edu.

- How will your tool or process help fill any gaps you identified in your department's current teaching evaluation practices?

▢ **If you are working on a teaching evaluation tool, what are the associated departmental processes you need to articulate?**

- When and how will this tool be utilized?
- What guidance do faculty need in order to engage with this tool?
- How will the evidence collected by this tool be assessed?
- Who will assess the evidence collected by this tool?

▢ **For more specific discussion prompts aimed at tools and processes associated with the voice your team is working on, see the relevant subsection of the "Departmental Level Tools for Assessing Teaching Quality" section.**

Implement tools and processes in the department

Once your team has completed a draft of a particular evaluation tool or process, we recommend working on departmental implementation rather than developing a complete revised evaluation system before starting implementation. This framing is suggested because it can be more challenging to generate buy-in for your materials and to make revisions to your materials if you wait until you've developed a complete evaluation system before sharing and implementing it. In this section, we present some suggested approaches for departmental implementation of evaluation tools and processes. This step is expected to last between one semester and a year as materials are put into use. During this period, you may not need to dedicate very much meeting time to this step, beyond checking in for updates and feedback.

Expected outcomes

- Collectively the working group will:
    - Implement a new evaluation tool or process in the department
    - Collect feedback on the implementation and make revisions to the tool or process as needed
    - Be prepared to develop a complete revised evaluation system

Suggested materials / preparatory work

- Have developed a draft evaluation tool or process that is ready to be implemented in the department
- Identify key individuals or small groups (e.g., a relevant committee) that can offer feedback and who may be willing to pilot the tool or process
- Review any departmental policies regarding the implementation of new teaching evaluation practices

Guiding questions for discussion

▢ **What can you do to generate departmental buy-in to new teaching evaluation tools and processes?**

- Who are the key individuals or groups you can reach out to for feedback and/or support for your work?
- How can you keep the department aware of your work?
- What are the potential concerns and push back you may encounter?
- How can you preemptively address those concerns or push back?
- How will you frame and present your tools or processes to faculty less familiar with your work?

▢ **What level of implementation are you seeking? (official department-wide use, a small pilot run, etc.)**

- Would a pilot run help you ultimately get official department approval?
- Are there department policies that affect what level of implementation is possible?
- What level of implementation does your team/department have the resources to support?

▢ **What resources and information do faculty need in order to implement your tool or process?**

- What explanation does your tool or process require for others to understand how to use it?
- How will you give faculty access to the materials they need to implement your tool or process?

▢ **How will you collect and act on feedback regarding your tool or process?**

- What methods will you use to collect feedback from faculty using your tool or process?
- How will your team incorporate feedback into the tool or process?

▢ **If you are ready to seek department-wide adoption of your tool or process, what are the policies regarding this?**

- Are there committees that need to approve the tool or process?
- Is a department-wide vote required?
- How can you set your team up for success at each level of approval required? (See the question above regarding generating buy-in.)

Develop a Complete Teaching Evaluation System

At this point your working group will likely have developed and implemented a set of tools for collecting data on teaching quality from three voices of assessment. In most cases there will be a need to rethink/develop/revise the way materials are reviewed for summative evaluation (annual merit and reappointment, tenure, and promotion (RTP) cases). And ideally, whether you started with tools/processes for merit or RTP, these two systems will connect in some way. In this section we suggest a series of actions to help you develop a complete teaching evaluation system for your department that pulls together everything you've been working on. However, a working group may choose to start this step earlier in the overall process (e.g., in conjunction with or immediately following "Assess Needs and Set Goals") to outline the summative evaluation process(es) and then work on developing tools across the three voices that fit the evaluation system. The timing of this step depends heavily on where in the process your working group chooses to address this and how heavily you draw on existing model approaches.

Expected outcomes

- Collectively the working group will adapt the TQF rubric (or another scholarly framework) into an assessment tool that provides structure and guidance for how your department will review your new/existing data sources for summative evaluations (merit and reappointment, promotion, and tenure)
- If your department has already developed improved teaching evaluation tools (see section above): finalize department implementation (see section above) to include this complete system
- If your department has yet to develop improved teaching evaluation tools: be prepared to adapt/adopt new tools across three voices (see section above) that fit with your evaluation system

Suggested materials / preparatory work

- Have already reviewed the TQF Framework model and full TQF Rubric (or other scholarly frameworks), and be familiar with the seven dimensions of quality teaching. See the "TQF rubric" resource below.
- Review some example teaching evaluation rubrics (see "Adapt the rubric into a summative teaching evaluation tool" in the resources section below)
- Optionally: Have already developed at least one new tool per voice (see Adopt/Adapt New Tools Across Three Voices above)
- Have copies handy of all current and new teaching evaluation tools, processes, and policies.

Guiding questions for discussion

▢ **Which dimensions of teaching quality will be assessed?**
- These will likely be drawn from the TQF Rubric or a different rubric
- Will these dimensions differ for annual merit vs. reappointment, promotion, and tenure?

▢ **How will you use your available data sources to assess proficiency levels in each dimension of teaching, and how will they be weighted?**
- What data sources regarding teaching quality do you currently have access to? (These can be pre-existing sources and/or if applicable the new ones you have developed/implemented as part of this process)
- If your team has not begun developing new evaluation tools, what new data sources and tools do you need in order to assess the dimensions of teaching?
- See also the guiding questions in the Adapt the rubric into a summative teaching evaluation tool section in the resources section below

▢ **How will you regularly review and revise your teaching evaluation system?**

▢ **How could you connect your teaching evaluation system for reappointment, promotion, and tenure to the annual merit structures in your unit?**

▢ **Once you've developed an assessment rubric, how will you implement it in the department?**

- See the "Implement tools and processes in the department" section above

Department Level Resources

In the following section, we introduce resources that can be used simultaneously and in alignment with developing evaluation tools (peer, self, and student voice) to increase the efficiency and effectiveness of the process.

Meeting Resources

We provide here tools that those facilitating working group meetings may find useful. They include an example timeline for the teaching evaluation transformation process and a template for meeting agendas and minutes. For more guidance on how to facilitate meetings, see the "Working Group Process" section.

Table 2. Timeline example

This table presents one possible timeline for the process of transforming teaching evaluation with your working group. You may find it helpful to provide or create such a timeline with your working group, so that you can track your progress and plan your time.

<table>
<tr><th>Process Step</th><th>Description of Step</th><th>Toolkit Location</th></tr>
<tr><td>Prep / Team assembly
(2-4 weeks)</td><td>assemble the departmental team and review overall resources and guidelines around teaching evaluation</td><td>Preparation Process</td></tr>
<tr><td>Overall Framing
(1-2 meetings)</td><td>Learn about the TQF framework and set your meeting approach for the term</td><td>Recommended Preparatory Reading

Develop Meeting Structures and Norms</td></tr>
<tr><td>Evaluate current system
(1-2 meetings)</td><td>evaluate practices and identify gaps; plan work along each voice; create timeline/process</td><td>Externalize Current Teaching Evaluation Practices

Assess Needs and Set Goals for Teaching Evaluation Transformation

TQF Rubric</td></tr>
<tr><td>Review approaches from other departments
(1-2 meetings)</td><td>select examples or templates to model approach on; select first voice and tool/process to work on</td><td rowspan="4">Adopt/Adapt New Tools Across Three Voices

Department Level Tools for Assessing Teaching Quality

Implement Tools and Processes in the Department</td></tr>
<tr><td>Complete draft form of tool(s) for Voice 1
(1-4 meetings)</td><td>work on tools/processes for Voice 1
Start implementation process for those tools</td></tr>
<tr><td>Complete draft form of tool(s) for Voice 2
(1-4 meetings)</td><td>work on tools/processes for Voice 2
Start implementation process for those tools</td></tr>
<tr><td>Complete draft form of tool(s) for Voice 3
(1-4 meetings)</td><td>work on tools/processes for Voice 3
Start implementation process for those tools</td></tr>
<tr><td>Finalize cohesive assessment plan
(2-4 meetings)</td><td>review span of data collection over rubric/framework; begin work on pulling the voices into a cohesive assessment plan</td><td>Develop a Complete Teaching Evaluation System

TQF Rubric</td></tr>
<tr><td>Finalize implementation plan
(2-3 meetings & on-going maintenance)</td><td>discuss plan for formal implementation; collect and discuss feedback from key folks external to team; revise materials as needed</td><td>Implement Tools and Processes in the Department</td></tr>
</table>

Meeting agenda and notes documents

As described in the Working Group Process section, having an agenda and organized note taking document can help your working group use meeting time effectively and record your previous discussions. You can find a blank template for a meeting agenda / notes document here.

Image 1. Example meeting notes structure.

(TEMPLATE) Year-Month-Day Title
Attendees:
Location:

Issue/Topic	Discussion	Decision/Action
Title: X:XXpm-X:XXpm		
Title: X:XXpm-X:XXpm		
Wrap-up X:XXpm-X:XXpm	**Next meeting:** ● [DATE] ● Location: ● Let me know if you can't make it, and we'll reach out to get your feedback ahead of time. **For next time:**	

Image 2. Example "parking lot" structure for keeping track of future action tasks

Issue/Topic	Discussion Points	Decision/Action (Time to Address)
Shorter-Term Parking		
Longer-Term Parking (and ongoing discussion)		

This Toolkit is in beta testing, and we welcome your feedback. If you would like to share your thoughts, please contact Sarah Andrews at sarah.andrews-1@colorado.edu.

TQF Rubric

The Teaching Quality Framework Rubric is a tool for considering the different dimensions of quality teaching. The rubric categorizes seven dimensions of teaching, describes what different levels of proficiency in each dimension might look like (ranging from "entry to teaching" to "advanced"), and provides guiding questions for considering each dimension. The dimensions of quality teaching presented are founded on the scholarship on teaching and teaching evaluation, and in particular are adapted from the work of Glassick, Huber, and Maeroff, 1997 and that of Follmer Greenhoot, Ward, and Bernstein, 2017. The dimensions of quality teaching presented in the TQF Rubric are:

- Goals, Content, and Alignment - an instructor's goals for their class and students, and how those goals align with class activities and student needs
- Preparation for Teaching - an instructor's readiness for classroom mechanics and their knowledge of content and pedagogy
- Methods and Teaching Practices - an instructor's teaching strategies and activities and their implementation
- Presentation and Student Interaction - an instructor's engagement with their students and student feedback
- Student (and Other) Outcomes - an instructor's and their course's impact on students and their measures of student understanding
- Mentorship and Advising - an instructor's support of students in mentoring contexts outside the classroom
- Reflection, Development & Teaching Service/Scholarship - an instructor's engagement with and contribution to local or external teaching communities

An instructor's attention to issues of diversity, equity, and inclusion (DEI) and their implementation of strategies that support diverse learners and an inclusive classroom climate are included across all seven of the dimensions of quality teaching. Additionally, a diversity, equity, and inclusion (DEI) focused supplement is available, which provides an opportunity for faculty interested in evaluating these components within their teaching evaluation practices. [TQF DEI Supplement to the Rubric].

In the following sections of the Toolkit, we orient readers to the TQF Rubric and provide guidance for a few specific ways in which the rubric may be used:

- Identifying Gaps in Teaching Evaluation Systems
- Adapting the Rubric into a Summative Teaching Evaluation Tool

Identifying Gaps in Teaching Evaluation Systems

The following resource is intended to help departments/individuals connect sources of evidence that can be used in teaching evaluation to the seven dimensions of quality teaching outlined in the TQF Rubric. While there are multiple ways one could approach using this resource, we outline a few steps that could be taken by departments/individuals to conduct a gap analysis of their current teaching evaluation systems. Departments/individuals that complete this activity should gain a better understanding of how their current teaching evaluation tools/processes align with the TQF rubric and gain a broad sense of whether their evaluation system(s) are missing key voices and/or data sources across the rubric dimensions (this resource can also be found on our website here; see also "Assess needs and set goals for teaching evaluation transformation" in the Working Group Process section above).

Description of components in the TQF Rubric Map

- The seven dimensions of quality teaching are described in the leftmost column; more fine-grained examples are included (from the "Professional" proficiency level articulated in the full TQF Rubric) to aid understanding of the dimension
- The three key voices that evidence should be drawn from are the voice of the instructor being evaluated ("self"), the instructor's peers ("peer"), and the students that the instructor mentors, advises, and/or teaches ("student"). "Voice Rank" is the relative importance of a voice for the given rubric construct: 1 = high importance//high weight, 2 = mid importance/mid-weight, 3 = lowest importance/low weight, and NA = not applicable.
- The sources of evidence listed in the rightmost columns by voice are not exhaustive, but represent some of the most common/important sources that could be collected from each voice and align with the dimensions of quality teaching.

Suggested Approach to Using the TQF Rubric Map

Before beginning, it may be helpful to have already reviewed the full TQF Rubric and be familiar with the seven dimensions of quality teaching. It may also be helpful to pick either annual merit or reappointment/tenure/promotion (RTP) as the evaluation system you want to assess using this tool rather than trying to do both at once (noting that coordinating these in the long term is likely useful for faculty and the department). Finally, it may be helpful to have handy any departmental/unit policies and procedures for teaching evaluation and copies of any tools you currently use (e.g., a peer observation protocol, a classroom interview protocol, a list of questions asked on student evaluations of teaching (SETs; known as FCQs at CU), a solicitation you use for getting student letters, etc.).

Step 1: Identify available sources of evidence in your department

- Look at the sources of evidence in the three voice columns to the right for the first dimension (Goals, Content, & Alignment) and identify which, if any, tools your department already uses. If your department uses tools / evidence that are not represented, write them in.
- Repeat for the other six dimensions

Step 2: Consider caveats

- For data sources you already use, note whether they are typically required for evaluations or if they're only optional and/or if optional whether they are frequently used or only sporadically. For example: perhaps for peer voice everyone is expected to have a classroom observation on file but while you have access to syllabi/course materials, their review is optional not required and are infrequently included in evaluations.

- Also consider whether the data sources your department has available, as currently specified, would be able to provide evidence for the given dimension of quality teaching. For example: perhaps your department does peer observations but how you do them currently wouldn't allow you to identify from them whether the instructor has well-articulated course goals.

Step 3: Identify gaps

- Review your work - are there any gaps that stand out to you? For example:
    - Are you completely missing sources of evidence for a whole dimension of quality teaching?
    - Is one particular voice missing across several dimensions?
    - Do you have a lot of caveats for one or more sources, such that you have a source but it doesn't align with the rubric dimensions very well?
- List out these gaps
- Note: be sure to keep in mind the "voice rank" column. Some voices may be more or less important for a given dimension e.g., student voice is likely not applicable to "Reflection, Development, and Teaching Service/Scholarship" and so not having any student voice for that dimension would not be a gap.

Step 4: Discuss gaps and identify the one(s) you will begin working on

- From the list of identified gaps, which do you think may be easier/harder to tackle in your department? Or, which are you most/least interested in tackling? You may want to prioritize tackling low hanging fruit and/or gaps that you think there would be particular interest in your department for addressing. For example: If you do peer observation and letters are required in your RTP packets but they don't align with the rubric very well and/or aren't particularly structured - Do you think your department would be open to the idea of more structure? If so, the TQF has templates and a variety of department examples that you could get started on right away (See the "Department Level Tools for Assessing Teaching Quality" section below).
- Make a decision on a tool and/or process that you would like to work on first and outline a plan for other gaps you think you could work on.

Adapting the rubric into a summative teaching evaluation tool

We recommend starting from the TQF example rubric-based approach for evaluating teaching rather than starting from scratch. While this example could be adopted mostly as is, it can also be adapted as needed to better fit a unit's needs. Below are some possible guiding questions for how a team within your unit might approach reviewing / discussing this resource (these questions are also included in Appendix A within the example linked above; **note references to page numbers below refer to those in the linked example not to this toolkit).**

- Examine the forms of evidence list (page 9) - which of these do you already use and/or would you like to be able to include? Are there any that your unit will absolutely not use? Are there any forms of teaching evidence that your unit uses/would like to use that are missing from the list? Revise the list as needed, ensuring that at least three key voices are included (peers, students, and the voice of the instructor being evaluated) (note: for non-CU institutions, FCQ = faculty course questionnaire, the CU version of end-of-semester student evaluations (SETs). If you've already done the mapping exercise as part of "Identifying Gaps in Teaching Evaluation Systems", you can pull from your work there rather than repeating work).
- Would the overall structure of the assessment rubrics for each dimension (pages 2-8) and the evaluation summary (page 1) fit your unit's needs?
    - See Appendix B (page 12) for rubric templates with slight variations in structure
    - If your unit would like to explore alternative rubric forms, other examples include:
        - University of Massachusetts Teaching Evaluation Rubric; University of Kansas Benchmarks for Teaching Effectiveness Evaluation Form (link automatically downloads a docx); University of Oregon Evaluation of Teaching Criteria; see also the GSLL annual teaching evaluation rubric (CU Boulder) for an example used in annual merit.
        - For units at CU Boulder who are more familiar with / prefer the three dimensions of quality teaching (Inclusive, Goal-Oriented, and Scholarly) specified by the A&S Quality Teaching Initiative (QTI), a version of this document with rubrics for those three dimensions is available here.
- For each assessment rubric:
    - Review the basic and professional/advanced criteria
        - Are they clear/easy to understand? Are they appropriate for your discipline? Do they span the space of teaching and mentoring activities for your unit? Is anything missing? Is the split between what counts as basic and what counts as professional/advanced appropriate for your unit (i.e., are there any "basic" items you would want to move to "professional/advanced" or vice versa?)
            - Note - we recommend ensuring at least one equity / inclusion related criteria be specified as "basic" for each dimension.
    - Review the proficiency level descriptions
        - Does this example include the right number of proficiency levels? Are the levels as specified appropriate for your unit or would you want to modify them? In particular, where the language of "courses / educational environments" is used, does this need to be further specified / clarified for common educational settings in your unit?

    - Review the suggested core and supplemental sources of evidence
        - Adjust as needed based on any changes you made to the full list of forms of evidence. If you removed, added, or altered any of the evaluation criteria and/or sources of evidence, and based on the specific sources of evidence you have / will have access to, consider whether or not the sources of evidence will collectively allow a reviewer to assess the basic and professional/advanced criteria to determine a proficiency level
- Think about weighting, e.g., should all seven dimensions be equally weighted or would you weight some dimensions heavier than others? Should different forms of evidence or different voices (peer, student, self) be weighed the same?
- Review the example instructions (page 10) for instructors, evaluation committee members, and mentors. Are these appropriate groups that would need instructions? Does the language make sense for your unit?

Department Level Tools for Assessing Teaching Quality

When conducting teaching evaluation, it is important to gather data from a variety of sources and people. We recommend particularly that departments attend to three voices in their teaching evaluations: the peer voice (evaluation by fellow faculty, instructors, etc.), the student voice (evaluation by current and former students, including advisees and mentees), and the self voice (one's own reflection).
Below we describe each voice, explain how to develop tools and processes to collect data from each voice, and provide links to the generic templates and department-specific examples of these materials. As described in the Departmental Processes and TQF Rubric sections above, the evaluation tools developed should align with an overarching assessment rubric. In this section, we also provide guidelines for facilitating departmental team meetings around creating these teaching evaluation tools and processes.

Peer Voice: Developing a Holistic Peer Course Evaluation Plan

Unstructured peer classroom observations, i.e., those that are not based on a set of core criteria, can result in inconsistency, may include biases, and do not always address teaching practices that are valued by a department (AAAS 2012). For this reason, scholarly literature on teaching evaluation recommends that academic units articulate the best teaching practices for their field and define core criteria to use in the observation process (AAAS 2012). Feedback on teaching can be more effective in promoting growth and improvement when it focuses on specific issues, contains concrete information, and is based on specific data (rather than general impressions) (Brinko 1993). Toward these ends, departmental working groups can develop standardized peer observation protocols, peer evaluation plans, and associated materials.

Peer observation of classroom teaching is the most common method of engaging peer voice in evaluation. Formative peer review that is nonjudgmental and serves to provide collegial feedback to instructors is crucial in developing an instructor's teaching efficacy. Making faculty aware of the fact that formative peer review is an opportunity for growth rather than a summative assessment will allow both reviewer and reviewee to focus on providing feedback that can be integrated into the classroom (Keig and Waggoner, 1994). A strong, holistic course evaluation plan focuses not only on classroom observations, but on pre and post observation activities such as pre-consultation, mentoring, and post-consultation practices (described more below).

Peer Evaluation Development Guidelines

Here are guidelines on how to navigate the creation and implementation of peer evaluation tools and processes:

1. Ensure everyone on the team understands the current promotion/tenure, reappointment, and merit processes in the department. It may be helpful to bring and review the following materials:
    a. Department's current procedures, processes, and protocols around peer observations and evaluation.
    b. Examples of peer observation letters

    c. Any existing procedures for how peer observations/letters are judged by tenure committees
    d. A complete example or mock up of a tenure packet to see what the peer observations are complemented by in the tenure process.

2. Decide on which aspects of peer evaluation the team would like to focus. Below we provide a list of potential aspects of peer evaluation that teams may wish to focus their work on. Later in this section, we provide sample peer evaluation materials teams have developed in these areas.
    a. [insert something related to timing and frequency, e.g., how often are they done and timing of when they are done, including any differences in timing/frequency based on instructor rank]
    b. Selecting observers (to articulate more explicitly than "the chair will do this")
    c. Pre-observation procedures (request/review class materials, consultation between observer and instructor, etc.)
    d. Classroom observation
        i. how many observations an observer would do of a given course in a given semester
        ii. how a protocol would be used by the observer
        iii. how the protocol would inform evaluation
        iv. frequency of observation
    e. Post-observation procedures (consultation between observer and instructor, opportunity for instructor to review the observer's report, etc.).
    f. Observation data management (Where will observation reports be stored? Who will have access?)

3. Develop the chosen peer evaluation materials. We recommend reviewing the templates and examples provided later in this section and considering the following:
    a. What parts or questions do you want to include or not include?
        i. Discussion on the items to include
            1. Discuss any revisions you may want to make to the items you want to keep
            2. Discuss and resolve any items with low agreement on whether to include or not
    b. What are your overall impressions on the structure of the form?
    c. What potential additional items (e.g., classroom interviews, optional open-ended questions, including ratings scales) might you want to add?

4. Discuss other relevant issues around how the department conducts peer evaluation and teaching evaluation more broadly. (e.g. should active learning be explicitly assessed? How does the department evaluate online courses? etc.)

5. Plan for when and how the peer evaluation materials will be implemented in the department
    a. Share updates on your work with the department to keep them aware of your progress and to get their input
    b. Consider when to request a formal faculty vote on adopting new evaluation

materials

   c. Anticipate and plan for any pushback you may encounter. Strategize around how you can create buy-in.
   d. Consider intermediate options if full adoption of all materials may be challenging
      i. Have an opt-in trial period for faculty to choose to use the new materials
      ii. Use a peer observation protocol as a guideline for writing narrative letters rather than having it serve as the report alone

6. Attend to feedback on the peer evaluation materials from faculty and make revisions as needed, keeping in mind the goals that you have for peer evaluation of teaching

7. Formalize peer teaching evaluation procedures in the department
   a. Ensure there are structures in place for continuing the use of the peer observation protocol (and getting feedback on it)
   b. Tie to mentoring processes if your department has any formal/informal mentorship in place or if you are thinking about developing one for the department.
   c. Articulate how the new peer evaluation procedures will interface with the department's other evaluation practices

Peer Observation Protocols

Peer observation protocols are intended to provide structure and consistency to peer classroom observations.

- Templates
  - Model A [link] (see GSLL, MCEN, and IPHY below for examples)
  - Model B [link]
  - Model C [link] (see MCDB below for an example)
  - Master list of peer observation protocol items [link]
- Department examples
  - Germanic & Slavic Languages & Literatures (GSLL) [language courses link] [content courses link]
  - Mechanical Engineering (MCEN) [link]
  - Integrative Physiology (IPHY) [link]
  - Molecular, Cellular, & Developmental Biology (MCDB) [link]

Peer Evaluation Plans

Peer evaluation plans should explicate departmental processes around how peer observation fits into their teaching evaluation plans, how frequently observations will be conducted and by whom, and how the peer observation protocols will be used by observers. Where relevant, peer evaluation plans could also explain how other forms of peer voice may align with, supplement, and/or replace classroom observations (e.g., peer review of a course portfolio, peer evaluation of classroom interview feedback or other student feedback, etc.).

- Template [link]
- Department examples
  - GSLL [link]
  - MCEN [link]

    - MATH [link]
    - MCDB [link]

Letter Writing Guides for Peer Observation

In some departments, a standardized peer observation protocol is not directly used for evaluation, but rather is designed to provide guidance to peer reviewers around what should be discussed in a peer observation letter. Similar to peer observation protocols, peer observation letter writing guides should include structured, consistent categories for review.

- Example from MATH [link]

Peer Observation Cover Letters

Because peer observation materials compiled for reappointment, tenure, and promotion dossiers may differ, in some cases significantly, from what had been submitted in the past, a cover letter that can be attached to peer observation materials can give readers guidance on how they should be interpreted and evaluated.

- Template [link]
- Example from GSLL [link]

Active Learning Addendums

Many departments have chosen to explicitly include active learning as a category for peer observation. In departments where faculty may have varying degrees of familiarity with active learning, a more detailed addendum may be included with the peer observation protocol to provide additional guidance.

- Template [link]
- Example from MATH [link]

Self Voice: Meaningful Self-reflection on Teaching

Self-reflection is a key element of a complete teaching evaluation. Often the instructor's voice is lacking in the evaluation process either for reappointment, tenure, promotion, and merit purposes, or for any formative assessments. Many departments are working on incorporating self-reflection (e.g., developing guidelines for writing teaching statements) into their evaluation processes. Examples of teaching-related activities that can be included in the self voice are a teaching statement that can cover information such as courses taught, developmental activities, service (observation of others, workshops etc.), grants, awards, publication, mentee feedback, etc.); a teaching portfolio including materials from courses, revisions or novel courses (including materials & description); a survey / report of practices; student learning measures and other assessment outcomes, etc.

Self Reflection Development Guidelines

Here are guidelines on how to navigate the creation and implementation of self evaluation tools and processes:

1. Ensure each team member understands the current promotion/tenure, reappointment, merit in the department. e.g. What are the department's current procedures, processes, and protocols regarding writing self-reflection? It may be helpful to bring and review the following materials:
    a. Department's current procedures, processes, and protocols around self evaluation.
    b. Examples of teaching statements or teaching reflections.
    c. Any existing procedures for how self-reflection statement are judged by tenure committees
    d. A complete example or mock up of a tenure packet to see what the self-reflections are complemented by in the tenure process.
2. Decide on which aspects of the self-evaluation/reflection the team would like to focus. Later in this section, we provide sample self evaluation /reflection materials teams have developed in this area.
3. Develop the chosen self-reflection tool. We recommend reviewing the self-reflection templates and teaching statement examples provided later in this section and considering the following:
    a. What parts do you want to include or not include?
    b. Discuss any revisions you may want to make to the items you want to keep.
    c. Discuss and resolve any items with low agreement on whether to include or not.
    d. What are your overall impressions on the structure of the document?
    e. What potential additional items might you want to add?
4. Plan for when and how the self evaluation materials will be implemented in the department.
    a. Share updates on your work with the department to keep them aware of your progress and to get their input
    b. Consider when to request a formal faculty vote on adopting new evaluation materials
    c. Anticipate and plan for any pushback you may encounter. Strategize around how you can create buy-in.

    d. Consider intermediate options if full adoption of all materials may be challenging.
5. Have an opt-in trial period for faculty to choose to use the new materials and attend to feedback from faculty and make revisions as needed.
6. Formalize self-reflection procedure in the department
7. Ensure there are structures in place for continuing the use of the guidelines.
    a. Tie to mentoring processes if your department has any formal/informal mentorship in place or if you are thinking about developing one for the department.
    b. Articulate how the self-evaluation procedures will interface with the department's other evaluation practices

Example: Teaching Statements

While a teaching statement is often required for reappointment, tenure, and promotion dossiers, typically there is very little guidance on what should be included in these statements. Giving faculty guidance on what they might discuss in a teaching statement can help standardize the evaluation process, make writing statements easier for faculty, and ensure that evaluation committees get the information they need from teaching statements.

- Example from MCEN [link]
- Example from IPHY [link]
- Example from Math [link]

Example: Self-Reflection in Merit

One way of incorporating the voice of the faculty member being evaluated into the evaluation process is to allow space for them to reflect on their own teaching practices in the annual merit process. Typically departments have created a list of guiding questions that align with their values and the TQF rubric, from which faculty can select the most relevant to their own practices.

- Template [link]
- Example from GSLL [link]

Student Voice: Using multiple measures to collect student feedback

Many units on campus rely heavily on end-of-semester student evaluations of teaching (SETs, aka FCQs at CU), particularly for annual merit. While students are one of the three key voices in evaluation of teaching, over-reliance on SETs as the primary or sole measure of student voice can be problematic (see “Role of Students in Evaluation: Student Voice” in our FAQ).

Guidelines on the Use of Student Feedback in Teaching Evaluation

These guidelines are helpful in identifying the changes your department might need to make when incorporating student feedback into teaching evaluations. We also suggest reviewing the Adopt/Adapt New Tools Across Three Voices section and referring to the examples provided in the toolkit below when working through this process. We recommend departments consider the following steps when working on student voice:

1. Ensure each team member understands the current promotion/tenure, reappointment, merit in the department. e.g. What are the department’s current procedures, processes, and protocols regarding student voice? (e.g., see “Externalize Current Teaching Evaluation Practices”)
    a. Understand the current practices of incorporating students’ voice in merit/RT&P such as SETs considerations, classroom interviews, student letter solicitation.
    b. Discuss and identify areas for growth/improvement.
2. Decide on which aspects of the student evaluations the team would like to focus.
    a. Discuss how to incorporate Student Evaluation of Teaching (SETs) in a more effective and valued way.
    b. Using classroom interviews to supplement SETs
    c. Better guidelines for student letter solicitations
3. Develop the chosen student voice tool. We recommend reviewing the templates and examples provided later in this section and considering the following:
    a. What parts do you want to include or not include?
    b. Discuss any revisions you may want to make to the items you want to keep.
    c. Discuss and resolve any items with low agreement on whether to include or not.
    d. What are your overall impressions on the structure of the document?
    e. What potential additional items might you want to add?
4. Plan for when and how the student materials will be implemented in the department.
    a. Share updates on your work with the department to keep them aware of your progress and to get their input.
    b. Anticipate and plan for any pushback you may encounter. Strategize around how you can create buy-in.
    c. Consider intermediate options if full adoption of all materials may be challenging.
5. Have an opt-in trial period to use the new materials and attend to feedback from faculty (and students or other stakeholders) and make revisions as needed.
6. Formalize use of student voice procedure in the department
7. Ensure there are structures in place for continuing the use of the guidelines.

Considerations for using SETs for Student Voice

Despite their wide use, Student Evaluations of Teaching (SETs) can produce inadequate representations of student voice if practitioners do not first consider what aspects of instruction

SET questions are informing. The following can be unintended consequences of the use of SETs:

- Ask questions students are ill-equipped to evaluate (e.g. how much they learned) (Kunter & Baumert, 2006);
- Rarely correlate with learning gains (Uttl, White, & Gonzalez, 2017);
- Trigger biased ratings for women/minority instructors, particularly in STEM (Li & Benton, 2017), yield lower scores for quantitative classes (Uttl, White, & Morin, 2013), and bias outcomes based on low response rates for online SETs (Nissen, Jariwala, Close, & Van Dusen, 2018);
- Incentivize unproductive teaching practices (e.g. grade inflation) in poorly designed SETs (Stroebe, 2016).

Reframing the process of SETs can lead to effectively capturing more representative data and student voice. A factor analysis of the Teaching Behavior Checklist (TBC) (Keeley et al., 2009) found that the qualities measured by SETs can be grouped into two distinct categories: the instructor's professional competency/communication skills, and caring/support provided to students. Both categories are highly relevant with recent changes in teaching modalities and shifts to online environments. Questions ideally are focused on the student perspective, such as their interactions with faculty, their awareness of course goals, and the demands on them in and out of class. SETs can be used to aid in the following:

- Flag potential concerns of poor instruction
- Capture typically unmeasured qualities such as inclusivity (if SET questions can be chosen or are open-ended) that can encourage actionable change

The use of SETs is also dependent on response rate from students. Practices that can aid in the process of administering SETs include ("How Can Instructors Encourage Students to Complete Course Evaluations and Provide Informative Responses?", 2020) employing strategies that improve response rate by:

- providing explanations and answering questions about the process (e.g. providing visual tutorial on how to access SETs)
- reserving class time for students to complete SETs and/or categorizing the SET as an assignment on the syllabus
- using real time response rate to inform students on class completion rate of SETs
- informing students about the use of SETs (e.g. merit, class improvement, etc.)
- teaching students about unconscious biases and how to overcome them (Peterson et al., 2019)

The following question and answer format is an example of what explanations and resources could be provided to students for better understanding of SETs and the students' role in communicating their feedback (a completed example can be seen here):

- What is the purpose of the student evaluation of teaching?
- Can my responses be harmful?
- How should I approach filling out the student evaluation of teaching?
- Are my responses anonymous?
- Can I see average scores of the student evaluation of teaching for my instructors/classes?
- If I have other questions about student evaluation of teaching, where should I go?

For departments able to pick open ended questions, replace questions that students are ill-equipped to evaluate with ones they are better posed to answer, and/or eliminate questions that ask overall course/instructor satisfaction, do so. Questions that can replace omnibus questions:

- In this course, I was encouraged to reflect on what I was learning or how I was learning.
- In this course, I was encouraged to evaluate arguments, evidence, assumptions, and conclusions about key concepts (critical thinking).
- In this course, the instructor maintained an environment that was respectful of diverse students and diverse points of view.
- In this course, the instructor seemed personally invested in student success.
- In this course, I was challenged to develop my own knowledge, comprehension, and conceptual understanding.
- In this course, I was challenged to provide feedback on my work that helped me improve my performance.

SETs should not be a standalone method of evaluation to capture student voice. It is encouraged that SETs are used with multiple measures such as classroom interviews. Additionally, frequent and formative feedback collected over the course of the semester can aid in more robust and applicable information when combined with SET use.

Example: Classroom Interviews

A classroom interview is a way to obtain constructive student feedback on teaching in a consensus-based format. Students work in groups to agree on course feedback and report out to a facilitator. As departments explore better ways for incorporating student voice, many have begun working on guidelines for conducting classroom interviews. While many departments allow faculty to submit this confidential feedback as part of their merit and/or reappointment, tenure, and promotion processes, several departments have developed modified versions that can be incorporated into their peer observation processes (see Peer Observation Protocol examples above).

- Template [link]
- Example from Math [link]

Your departmental working team can start from the list below to discuss ideas or as a starting point for classroom interviews. The following questions can be guides:

- Were formal assessments in class consistent with instructional objectives (homework tests, quizzes, etc.)?
- Did the instructor establish an environment that encouraged student participation out of class (e.g., office hours, TAs)?
- Did the instructor make appropriate connections to other areas of the course subject matter or to real-world applications of the topic?
- Did the instructor help to create an inclusive learning environment?

Example: Student Letters

In most departments, there is a lack of clear instruction on how to solicit, handle, and incorporate student letters for the evaluation purposes. The following guidelines for students writing feedback letters is one way to help your team to develop your department's customized tool.

1. Start with reviewing the current department's guidelines
2. Review previous students' letters to figure out what they cover and what needs to be covered.
3. Refer to the example [link] and the template [link] for the student letter guide. Based on your discussion in the previous steps, use or adapt any parts of the letter you find relevant.
    a. You can choose from the list of questions in the template and come up with a tailored list for the department.
    b. What items do you think are most important/least important?
    c. Is there anything missing?
    d. Adding department-specific examples
4. Finalize a list of items to include in the document
5. Work on the opening paragraph guidance in the document
6. Finalize plans to take it to rest of department
    a. Conduct a review by small number of faculty and/or students before bringing to department [revise to more strongly encourage student feedback on this particular item]
    b. Request a review from relevant departmental committees
    c. Other review based on the context of the department

Appendices

Teaching Quality Framework Initiative Background

As part of a broader multi-institution consortium (TEval) and a growing national dialogue (e.g., National Academies of Sciences, Engineering, and Medicine [NASEM], 2018; President's Council of Advisors on Science and Technology, 2012; Seymour & Hewitt, 1997), the TQF uses a common scholarly framework (TQF framework model and rubric) to help departments make improvements to the ways teaching is evaluated. Departments are a key unit for change within research universities, as they can impact multiple scales (e.g., individual faculty members, the department as a whole, and upper administration) (Bradforth et al., 2015; Henderson, Beach, & Finkelstein, 2011; Marbach-Ad, Hunt, & Thompson 2020). However, institutional context varies and the location for successful change initiatives may differ depending on local context (e.g., the dean level may be even more essential at many smaller institutions). We believe that any unit interested in making changes to improve evaluation of teaching will find these departmental level processes and resources useful.

Anticipated Longer-Term Outcomes

Departments engaged in this work could also see longer term outcomes such as improved student learning outcomes, increased value of (and rewards for) high quality teaching, increased use of evidence-based and inclusive teaching practices, and improved practices for hiring decisions. A scholarly approach to teaching evaluation can have specific benefits for:

- Tenure-Track Faculty, Teaching Professors, Instructors: By looking at faculty growth over time, it is possible to promote faculty learning, rather than just one-shot evaluation.
- Contingent Faculty: Peer review and reflection processes are mechanisms to support the development of all faculty, not just those seeking tenure or with long-term careers at the institution.
- Departments: Having a suite of well-designed, community-supported assessment tools will make it easier to more accurately and consistently evaluate faculty teaching for merit, reappointment, promotion, and tenure. A framework for teaching quality can also be used to assess whether the student experience is comparable in courses that have multiple sections taught by different faculty, graduate students, and learning assistants. This can help reduce the current overreliance on student evaluation of teaching (SETs) for the evaluation of faculty.
- Administration: Having better defined, multiple measures of teaching evaluation drawn from three voices (student, peer, and self) can signal internally and externally that the university values education while simultaneously enhancing its ability to improve education on campus. It can also support the accreditation process and sends clear messaging to outside constituencies about the types of innovative scholarship of teaching and learning that are taking place on campus.
- Students: As faculty teaching development is supported through transformed evaluation, students will benefit from improved instruction. New teaching evaluation systems can also provide students with more productive ways to provide feedback, beyond traditional SETs.

How the Toolkit has been Developed

This toolkit is derived from the work of various departments across the CU Boulder campus and the Center for STEM Learning. Over a period of several years, the TQF has facilitated working groups in multiple departments across three colleges (Arts & Sciences, Engineering & Applied Science, and Business) as they developed or adapted better tools and processes for evaluating teaching across three key voices (self, student, and peer) that align with the common framework while adapting it to their discipline and departmental values and culture (see Andrews et. al 2020 for a more detailed overview of the TQF process and a case study of one unit). Said departments went through the processes described in this toolkit, guided by facilitators who are also the authors of this toolkit. Departments had both common and unique experiences as they created and/or adapted tools and processes. Four research associates and a principal investigator from the Center for STEM Learning developed this toolkit through extensive iterations based on feedback from colleagues in the evaluation of teaching community. Each individual working on this toolkit has experience in STEM Learning (i.e., Physics Education, Biology Education, and Engineering Education) and served as facilitators for department faculty instituting change in their departments. The team of authors met regularly to discuss updates and edits, as well as solicited feedback from colleagues on necessary additions and changes.

We have built this toolkit to compile tools for teaching evaluation with a suggested approach / guidelines that departments can follow to improve their own teaching evaluation systems in the absence of an external facilitator. This Toolkit has been developed based on our experiences facilitating these working groups, TQF program evaluation, and the underpinning scholarly work on higher education (e.g., Bernstein, 2002; Bernstein et al., 2010; Glassick et al., 1997), institutional and organizational change in academic contexts (e.g., Corbo et al., 2016; Elrod & Kezar, 2017; Kezar, 2013; Kezar & Miller, 2020; Reinholz et al., 2017), and scholarly approaches to teaching and learning (e.g., Beach et al., 2016; Boyer, 1990; Fairweather, 2002). Although this toolkit is based upon work conducted within departments at CU Boulder, context was considered where appropriate to ensure that the processes described herein could be applied outside of CU Boulder.

Locating Additional Resources at Your Institution

As an additional consideration, finding, accessing, and utilizing campus resources may be necessary to help to support the following processes. If you are unsure of where to start in locating resources at your institution, consider starting with centers focused on teaching and learning and/or faculty development. These centers can be helpful places to obtain information on tools such as classroom observations, teaching reflection guidelines, and other resources to aid in your evaluation of teaching processes. Your department may have additional formal or informal networks related to the quality of teaching such as a pedagogy committee housed in your academic department or affiliation with national networks related to the evaluation of teaching (e.g. TEval, etc.). Your university may have programs that recognize teaching excellence and offer incentives or support for faculty attempting to improve their teaching. The following are types of campus resources for supporting the evaluation of teaching, with specific examples that are available at CU Boulder:

- Centers for Teaching and Learning - Provides support to faculty for teaching practices and professional development
    - The Center for STEM Learning
    - The Center for Teaching and Learning
- Departmental bylaws and processes for merit, promotion, tenure, and reappointment - Provides information on internal committees and processes, as well as a baseline for how your teaching is valued by your department
    - Departmental bylaws example
- Groups focused on excellence in teaching - Provides a potential community around teaching and teaching evaluation
    - ASSETT at CU Boulder
- Groups focused on the evaluation of teaching -
    - Quality Teaching Initiative at CU Boulder

[invitation to contact us or check the website for more info / support?]