

## High-Quality Evidence of Impact Advisory

Updated January 2026

Program approval is the process through which a sponsoring organization or licensure program receives state approval to operate. All Massachusetts educator preparation program providers are required to meet the same rigorous expectations articulated in the [2023 Guidelines for Educator Preparation Program Approval](#). Following initial approval, educator preparation programs undergo a periodic Formal Review to ensure their continued growth and effectiveness.

The [Formal Review process](#) relies on the following evidence base:

- Initial Inquiry evidence including the Program Overview, required documents, and candidate artifacts as well as surveys, focus groups, and interviews with internal and external stakeholders;
- Follow-Up Inquiry evidence including Targeted Submission worksheets and further documents and artifacts if requested; and
- State-collected data (see [Edwin for Ed Prep Overview](#)).

Taken together, this collection of evidence helps DESE and the review team determine whether each program approval criterion is rated as a commendation, met, or finding in the final report.

In the Follow-Up Inquiry, a sponsoring organization (SO) submits evidence by responding to prompts in submission worksheets tailored based on their Initial Inquiry evidence. The SO is responsible for providing high-quality evidence to address gaps and inconsistencies or elevate promising practices identified through the Initial Inquiry.

SOs may share data collected by their own organization or data collected by other organizations (including DESE-collected data). These data do not need to be (and in many cases should not be) evidence collected specifically for the review process; SOs should describe the sources of evidence they use in their continuous improvement processes, including how the evidence demonstrates the impact of their practices related to each criterion and how they use the evidence to make adjustments as necessary to support effective and equitable preparation experiences.

## Evidence of Impact

In all cases, evidence of impact is weighted more heavily than evidence that describes inputs or plans (see Hierarchy of Evidence graphic). This advisory focuses on collecting and reporting high-quality evidence of impact.

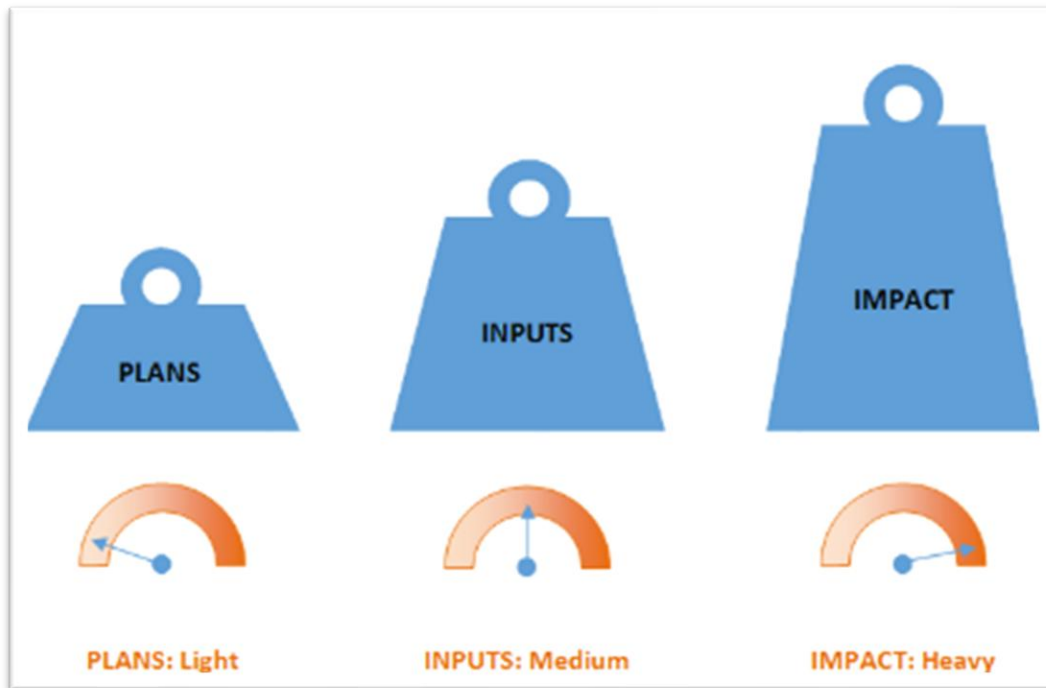


Figure 1. Hierarchy of Evidence in the Formal Review

The goal of this advisory is to support SOs' understanding of the characteristics of high-quality evidence of impact and enable SOs to evaluate the quality of evidence they are collecting for their continuous improvement processes. More specifically, this advisory will provide general guidance on incorporating data from a variety of sources, including quantitative and qualitative data, to demonstrate the impact of the SO's practices in the Follow-Up Inquiry of the Formal Review.

## What is High-Quality Evidence of Impact?

The table below describes the characteristics of high-quality evidence of impact. The "Questions to Consider" guide SOs to evaluate the quality of evidence collected and shared. **To inform their continuous improvement efforts and prepare for their Formal Review, SOs should consider what evidence they are already collecting or could begin collecting that meets these characteristics.** SOs may wish to consider this table alongside the 2023 Program Approval Criteria list to determine whether they are collecting high-quality evidence of impact aligned with each criterion.

## Characteristics of High-Quality Evidence of Impact

High-quality evidence of impact submitted through the Formal Review should be:	Questions to Consider
<p><b>Aligned with criterion:</b> Evidence of impact and the programmatic aspects it is measuring should be directly linked to the expectation described in the criterion being evaluated. Evidence sources that measure SOs' impact more broadly will be less informative at the criterion level in the Formal Review.</p> <p>Some criteria have multiple components and thus multiple sources of evidence may be needed to address all components of the criterion.</p>	<ul style="list-style-type: none"> <li>• Will this evidence help demonstrate whether and how we are meeting the specific expectations described in the criterion?</li> <li>• Which component(s) of the criterion expectation does this address?</li> <li>• What additional evidence do we need to address all aspects of the criterion?</li> </ul>
<p><b>Connected with plans and inputs:</b> Evidence should describe the impact of existing systems/structures (inputs) in place to address the criterion and/or how evidence has informed plans for future systems/structures.</p>	<ul style="list-style-type: none"> <li>• How does this evidence connect with structures or systems that support the expectation described in the criterion?</li> <li>• Based on this evidence, what plans does your SO have to maintain or modify the systems and structures?</li> <li>• Does this evidence help illustrate how your SO is making data-informed decisions to improve its systems and structures?</li> </ul>
<p><b>Outcomes-based:</b> Evidence focuses on the <i>results or impact</i> of the aspect of the program it is measuring.</p>	<ul style="list-style-type: none"> <li>• Does this evidence show whether the system/structure/decision/policy is working to improve quality and equity in programs?</li> <li>• What is your SO's interpretation of the results the evidence shows?</li> <li>• What impact on internal or external stakeholders does this evidence demonstrate?</li> </ul>
<p><b>Specific and grounded in data:</b> Evidence of impact that demonstrates how a system/structure supports effective and equitable preparation should specifically measure the intended result or impact of that system/structure.</p>	<ul style="list-style-type: none"> <li>• Is this evidence directly aligned to the intended outcome of the system/structure?</li> <li>• Does analysis of this evidence remain explicitly grounded in the data?</li> </ul>

	<ul style="list-style-type: none"> <li>• Does analysis include reasonable inferences based on the evidence, avoiding broad statements that extend beyond the evidence itself?</li> </ul>
<p><b>Measured by multiple sources:</b> Whenever possible, it is helpful to triangulate evidence from multiple sources, including both quantitative and qualitative evidence, that measure the same result or impact. Evidence should also include the perspectives of multiple stakeholders, including those most directly impacted by the system/structure/practice.</p>	<ul style="list-style-type: none"> <li>• Does your SO have multiple sources of evidence that can be used to more clearly understand the impact?</li> <li>• Does the evidence reflect multiple perspectives on the impact?</li> <li>• Do multiple sources of evidence support a similar conclusion or are they contradictory?</li> <li>• If contradictory, what additional measure(s) will your SO use to better understand the impact?</li> </ul>
<p><b>Capture multiple points in time:</b> Whenever possible, evidence should reflect change or continuity in impact over time. This could include pre- and post-implementation data; formative and summative data; and/or longitudinal data from multiple points throughout a year or across multiple years.</p>	<ul style="list-style-type: none"> <li>• Can the data be used to show trends in impact over time?</li> </ul>
<p><b>Interpreted:</b> All evidence shared should include a clear description of your SO's key takeaway(s), including what the evidence shows your SO about how it is meeting the expectation described in the relevant criterion. It is especially helpful to share your analysis and next steps you plan to take (or have taken) if the evidence indicates there may be gaps or inconsistencies in the impact of your SO's practices.</p>	<ul style="list-style-type: none"> <li>• What did your SO learn from this evidence?</li> <li>• What are the takeaways your team identified?</li> <li>• What are your decisions or next steps based on this evidence?</li> </ul>

## Tips for Completing Follow-Up Inquiry Targeted Submission Worksheets

Sponsoring organizations are expected to collect evidence and monitor the impact of their preparation programs at least annually. While the recommendations above apply to both organizations' own continuous improvement efforts and the use of high-quality evidence of impact in the Formal Review, below are recommendations specific to providing a clear, strong evidence base in the Follow-Up Inquiry.

### **Respond Directly to the Prompts**

Responding directly to the prompts ensures that the information provided is relevant, focused, and aligned with what is being asked. When answers address the prompt head-on, it becomes easier to evaluate the response, make clear connections between the response and the criterion, and contextualize the response with evidence collected in the Initial Inquiry. Direct responses also prevent confusion, reduce unnecessary detail, and keep within the word limits.

### **Focus on Submitting New and Different Evidence**

Submitting new evidence is most likely to inform the review team's understanding and ratings. While the information collected during the Launch and Initial Inquiry provides us with a baseline understanding of how your organization is meeting expectations, we are asking you to address gaps in our evidence base and/or inconsistent experiences within or across stakeholder groups. New evidence helps us further understand what we already know, reveal trends we did not yet hear about, and elevate recent changes. This allows the review team to see the most current and complete picture possible. For example, if your targeted submission worksheet notes that a specific training is used for program supervisors, it is not helpful to reiterate that this training is provided. However, it may be helpful to note if an additional training is provided that was not mentioned in the targeted submission worksheet or if your sponsoring organization has evidence that you have not previously shared with DESE that the training is effective.

### **Provide Specific Examples**

Providing specific examples allows the review team to concretely understand what is happening at your organization. While general statements can be interpreted in many ways, specific examples can illustrate exactly what is meant and reduce the chance of misunderstandings. Specific examples create clarity and support stronger analysis.

### **Consider Whether Your Organization Has Further Evidence**

Sponsoring organizations may decide not to provide further evidence for any prompt within the Follow-Up Inquiry. We recognize that it takes significant time and capacity to provide further evidence; if your organization does not have additional evidence to address gaps or inconsistencies (or further highlight promising practices) elevated in the Follow-Up Inquiry, you may decide not to provide further information. In these cases, DESE and the review team will rate the criterion based on the evidence collected in the Launch and Initial Inquiry stages.

## Using CAP Ratings as an Evidence Source in the Follow-Up Inquiry

Sponsoring organizations may use CAP feedback, evidence, and ratings for a variety of purposes within their own continuous improvement. Comparing CAP data within and across programs, among different program supervisors, or among different practicum settings can be valuable information that informs field placement practices, training and support for supervisors, course-embedded field-based experiences, etc.

In the Follow-Up Inquiry, a summary **of CAP ratings on their own is not typically considered high-quality evidence of impact**. While CAP is designed to support candidates' readiness to teach all students and DESE expects organizations to calibrate its ratings within and across programs, the purpose of CAP is to assess individual candidates' performance, not to demonstrate a program's efficacy. CAP summative ratings in isolation could be interpreted in a variety of ways that may or may not demonstrate effective and equitable preparation. The ratings on their own will not demonstrate that CAP is being used consistently "to improve practice and ensure only candidates who are ready for full responsibility in the licensure role are endorsed."<sup>1</sup>

With additional context or alongside additional sources of data, CAP ratings may tell a clearer story about your organization, making them a stronger source of evidence. For example:

- Using data from the Candidate Assessment of Performance Data Dashboard (EP901) to show strong correlations between CAP ratings at your organization and completers' Educator Evaluation ratings or Student Growth Percentiles
- Sharing longitudinal data that shows that, following additional support/training for field supervisors, CAP ratings have become more consistent within and across programs, program supervisors, etc.
- Sharing examples of high-quality feedback provided to a candidate along with evidence that their practice improved such that they earned specific CAP ratings
- Sharing examples of how multiple sources of evidence were used with a clear rationale to support CAP ratings

---

<sup>1</sup> FBE Criterion 5 from the [2023 Guidelines for Educator Preparation Program Approval](#)

## Additional Considerations for Quantitative Evidence

Below are recommendations to consider while incorporating quantitative data into Follow-Up Inquiry during the Formal Review. These suggestions are meant to ensure data provides the review team with the information they need to contextualize it in their decision-making process and triangulate it with other evidence sources used throughout the review.

Recommendation	Rationale
<p>When surveying stakeholders, SOs may wish to reference the <a href="#">DESE Stakeholder Surveys</a> as a resource for crafting items aligned to the 2023 Program Approval Criteria.</p>	<p>These survey items were specifically crafted to measure DESE’s expectations. Using these items supports alignment with criteria and allows for a clearer comparison between the SO’s survey results and results collected by DESE through the annual stakeholder survey and/or the Formal Review process.</p>
<p><b>Share data from the three most recent years</b>  <i>“X percent of 2022 completers (N=a), Y percent of 2023 completers (N=b), and Z percent of 2024 completers (N=c)...”</i></p>	<p>Although SOs may choose to include data from any year(s), focusing primarily on the most recent three years provides reviewers with an indication of experiences in the program that are most reflective of any recent changes and allows for longitudinal comparisons when appropriate.</p>
<p><b>Include the N-size</b>  <i>“In 2024, we received survey responses from 58 completers and found that...”</i></p>	<p>This will help reviewers understand how representative this data is of the population of stakeholders.</p>
<p>Include the entire survey scale for the item in the prompt response.  <i>“When asked whether, overall, their advising was effective, <b>80 percent of candidates (N=60) agreed, 10 percent somewhat agreed, and 10 percent somewhat disagreed.</b>”</i></p>	<p>Seeing the full range of responses enables reviewers to better interpret the evidence.</p>