

Beyond Outcomes: Assessing the Process of Problem Solving Using Process Education

Joshua Morrison, M.S.Ed.

Author Note

Joshua Morrison, <https://orcid.org/0009-0008-5586-9690>

I have no conflicts of interest to disclose.

Intersection: A Journal at the Intersection of Assessment and Learning
AALHE Conference Proceedings 2026, Volume 7, Issue 4

Abstract: The link between a finished course assignment and the learning required to produce it is increasingly uncertain. In the era of generative artificial intelligence, polished coursework can be created in moments without demonstrating the reasoning, judgment, or productive struggle traditionally associated with learning. Product-focused grading may therefore be evaluating the AI tool used to produce it, not the learner's actual performance. This article presents Process Education's assessment framework for assessing four links in the problem solving performance chain: definition, strategy, implementation, and validation. Rather than focusing solely on the final product, the framework makes the learner's process more visible and assessable. Practical tools include a seven-dimension process rubric, the Strengths-Improvements-Insights (SII) feedback method, coaching prompts designed to elicit process evidence, and illustrative student examples. During the related AALHE conference session, participants analyzed individual samples, identified broken process links, and drafted formative feedback intended to strengthen future performance. The article argues that process-oriented assessment offers a constructive response to generative AI by shifting attention from detecting tool use to evaluating the quality of student reasoning. It also supports growth-oriented feedback that helps learners improve their learning skills and future performances, strengthen self-assessment, and develop insights transferable across contexts.

Keywords: *Process Assessment, Process Education, generative AI, rubric design, formative feedback, Academic Integrity*

Introduction

Professor Smith is in her office on a Wednesday afternoon, working through the latest class problem sets in her solid mechanics course. The assignment asks students to calculate the load-bearing capacity of a bridge design, a problem type she has assigned for years. Alex's submission shows the correct final answer, presented neatly with a formula and the appropriate substituted values included. When Smith asks Alex during office hours the next day to walk her through how he approached the problem, his explanation unravels. Alex used a generative AI (GenAI) tool to retrieve the formula, and he can't explain why his chosen formula is a better option over a more conservative alternative. In his written work, Alex missed a unit conversion that the AI tool fixed automatically; he didn't notice the error.

While Professor Smith is fictional, examples like these are now ubiquitous in higher education. With the advent of GenAI systems, students can produce technically accomplished assignments, essays, and other artifacts in moments. Thus, the link between a submitted assignment and the learning that produced it is no longer so strong.

Further, the implications of GenAI for student learning are significant. If grading simply rewards the answer or final product rather than the reasoning and process that produced it, both individual student grades as well as program-level learning outcome data may reflect the quality of the tool(s) used, rather than the actual competency students develop. At a time when confidence in American higher education is at or near historic lows (Jones, 2026), ensuring the value of a college education in terms of its strong learning outcomes becomes even more essential.

This paper presents a process-based assessment framework grounded in Process Education (PE) that faculty and assessment professionals can adapt across disciplines. It distinguishes between product evaluation and process assessment, introduces four observable links in problem solving, explains how a seven-dimension rubric and formative feedback (using the Strengths-Improvements-Insights (SII), method) can be used to assess those links, and provides prompts and examples for capturing process evidence.

This paper and the presentation it reflects are based on my experience with PE firsthand, as a student team coach and course facilitator of Learning to Learn courses, as well as my role as a member and now President-Elect of the International Academy of Process Educators.

Assessing Process, Not Just Product

Many faculty and assessment professionals are already fluent in judging quality. Faculty provide letter grades on student assignments. Directors of Assessment may provide academic program review summaries verifying whether an academic program meets accreditation requirements. Both are examples of product-based evaluation, where a judgment is made of an artifact by assigning it a score, grade, or summative mark. Process-based assessment asks a different, more developmental question: How did the learner (or program director) get to this product? The answer is then used to improve future performance, rather than close the book on the past.

This distinction matters most in terms of the feedback it produces. A score of “95, A” or “Meets Expectations” is closed; it confirms that a certain level of quality was achieved. Process-oriented feedback provides something for the learner to act on; it is open and names what is working well, what can be improved, and what to adjust next time. This kind of feedback becomes a starting point for future development and student growth. That starting point is beneficial in two ways. For students, process-focused feedback supports a growth mindset and stronger self-direction, since learners now know specifically what can be adjusted, not just what was wrong. For faculty and assessment offices, process-oriented feedback aids in diagnosing problems, since these surface earlier, and instruction or guidance can target the reasoning students are using rather than just the artifacts they submit. This feedback can also be timelier, and thus, more relevant, to the current performance.

These benefits are not unique to Process Education; they reflect decades of research on formative assessment and self-regulated learning. What PE offers is a coherent philosophy, along with a set of robust tools, that ties those threads together into a single, practical framework for observing and improving performance (Duncan et al., 2016; Van Slyke, et al., 2021).

Conceptual Foundation for Process Assessment

Process Education synthesizes multiple educational theories and practices within a broadly constructivist perspective, treating learning as an active, contextualized process of making meaning from experience (Duncan et al., 2016). Two traditions are especially relevant to process-oriented assessment: formative assessment and self-regulated learning.

Formative assessment considers feedback as an essential catalyst for improved learning performance. Black and Wiliam (1998) found that specific, actionable feedback produced some of the largest effect sizes in the educational research literature. As shown across over 40 studies, students who received strengthened formative assessment improved at a significantly greater rate than their peers. This tradition supports PE's SII model, which identifies a learning performance's strengths, areas that can be improved, and resulting insights (realizations) that are transferable to other contexts.

Self-regulated learning provides a complementary foundation. Zimmerman's (2002) model includes forethought, performance, and self-reflection. These stages parallel the presented problem solving performance chain, which synthesizes the strategy proposed by Morgan and Williams (2010): definition and strategy involve preparing and selecting an approach, implementation involves carrying out and monitoring that approach, and validation involves evaluating the result and reflecting on future use. Both models ask learners to become intentional in their preparation, monitoring, and learning from a performance.

Process assessment also addresses the question of validity of judgments made about the quality of a performance. Messick (1989) argued that validity is about whether an assessment supports the interpretations and uses made from its results. When a finished assignment does not reveal the reasoning used to produce it, a gap is created in the evidence of the competency being assessed. When faculty utilize process evidence of the problem solving process, the learner's thinking, decisions, and analysis become more visible at the course and program levels. This visibility, then, enhances the validity of course and program level assessments of performance.

This issue is critical in the age of GenAI, when learners can produce polished prose, computer code, and many other artifacts without the requisite productive learning struggle. This is a significant assessment challenge, as product-only grading risks passing judgment on the quality of the GenAI tool that produced it, rather than the learner's actual performance. Perkins et al. (2024) addressed this with their Artificial Intelligence Assessment Scale, which establishes differing levels of permitted AI use. The problem solving performance chain serves a related purpose; it diagnoses the quality of the learner's reasoning whether GenAI tools are involved or not. Instead of focusing on a range of AI use permission structures, assessing the process of problem solving allows insight into the reasoning and critical analysis of the learner.

Faculty and assessment professionals must therefore move upstream in the product development process. Useful process evidence may include:

- Decision points: Choices among formulas, approaches, sources, or assumptions, including the rationale for those choices.
- Iterations: Revisions, abandoned strategies, corrected assumptions, or other adjustments made while the work developed.

- Error checks: Steps used to determine whether the result is reasonable and meets established quality criteria.
- Reflections: Explanations of what the student did, why it worked or did not, and what could be transferred or changed in a future performance or context.

Within PE, assessment involves observing the quality of a performance using established criteria and providing feedback that documents strengths and guides improvement (Duncan et al., 2016; Van Slyke, et al., 2021). Two elements are central to process assessment. First, performance criteria, which define what “good problem solving” looks like in observable terms. Second, growth-oriented feedback moves beyond general evaluations such as “good job” by identifying evidence, explaining its effect on the performance, and describing what the learner can do differently to enhance performance next time. The Classification of Learning Skills can provide precise language for identifying the particular skills demonstrated or needing development (Process Education, n.d.). The problem solving performance chain operationalizes these principles as a series of observable links.

The Problem Solving Chain

Using a PE framework, then, problem solving can be considered a sequence of cognitive links. Weak links can break the chain, even when a final answer looks correct, while strong links support both better outcomes (better products) and stronger learning and student growth. Table 1 summarizes common steps in the problem solving chain.

Table 1

Problem Solving Performance Chain

Link	Definition	Strategy	Implementation	Validation
What it means	Identify the real problem	Choose the right mental tools/approach	Execute the plan	Check results against performance criteria

Assessment of student learning, then, should target these links in the process, not just the final product. Each link must be made observable with criteria and evidence, so that formative feedback is possible and serves to improve the next attempt by the learner.

This feedback can be produced through a process-oriented problem solving rubric, shared by the faculty member with students to ensure a common understanding of the criteria for quality. Importantly, faculty should consider missing evidence as a flag that high levels of performance cannot be awarded; absence of such evidence is to be considered absence of demonstrated performance, not assumed competence. This rule of thumb keeps the assessment focused on the process of problem solving and its steps, and avoids assuming development and growth based solely on the final product itself. As an example, Table 2 outlines seven rubric dimensions that can be used to assess the problem solving process. This rubric is intended to be useful to faculty in most classroom environments, and was developed from the original presentation materials with assistance from Claude Sonnet 5.0 Medium (Anthropic, 2026).

Faculty can use the following scale to document performance:

- Emerging: Little or no process evidence
- Developing: Partial or inconsistent evidence
- Proficient: Clear and appropriate evidence
- Advanced: Clear, adaptive, and transferable evidence.

Table 2

Seven-Dimension Rubric for Assessing the Problem Solving Process

#	Dimension	What It Captures
1	Problem Framing	Goal, constraints, and assumptions identified
2	Strategy & Planning	Approach selection plus rationale for choosing it
3	Execution & Organization	Structured steps that align to the stated plan
4	Monitoring & Adapting	Checkpoints, error checks, and revisions made mid-process
5	Evidence/Justification	Use of principles, data, or reasons to support decisions
6	Reflection & Transfer	"What/why/next time" thinking and generalization to new contexts
7	Communication of Reasoning	Clarity, coherence, and explainability of the process to an outside reader

The seven dimensions operationalize the four links: Problem Framing corresponds to Definition; Strategy and Planning to Strategy; Execution and Organization and Monitoring and Adapting to Implementation; and Reflection and Transfer to Validation, while Evidence and Justification and Communication and Reasoning operate across the chain.

Along with a formal letter grade or score, students benefit from additional formative feedback, as the score represents a judgment of quality on a past performance. By itself, it is not useful for producing growth among learners. For this, the SII feedback method is needed.

The SII Feedback Method

In Process Education, the SII feedback method is used commonly as a catalyst for growth, and self-growth, the ability for learners to mentor their own performance enhancement over time (Duncan et al., 2016). SII incorporates a performance's strengths, including how and why these were important to the performance quality (typically at least two). The strength feedback is grounded in specific evidence and includes the impact of each strength on the work product. Potential areas for improvement identify where the performance can be strengthened, and how to implement those improvements (one or two). Insights are transferable realizations that generalize beyond the current single task to other contexts. Both the Strengths and Improvements feedback are enhanced when assessors view the performance in terms of the learning skills required to produce the work product.

The Classification of Learning Skills (CLS) can help assessors name the particular cognitive, social, affective, or assessment skills demonstrated in a performance or needing further development (Process Education, n.d.). Because faculty are interested in producing lifelong learners that grow their capabilities over time, the SII feedback method is a powerful way to grow learning skills by explicitly identifying the ways in which students use them to greatest effect, and where they can be developed more strongly, along with insights gained by completing the formative feedback process. In addition, faculty can benefit from enhanced process evidence that uncovers the learning skills and processes

students use to prepare their work products (Apple, et al., 2016). The next section articulates some suggested coaching prompts that can elicit stronger evidence of these processes.

Coaching Prompts for Process Evidence

Coaching prompts embedded in course assignments and assessments such as quizzes or exams may make otherwise hidden reasoning visible. Student reasoning can be made more explicit through oral responses, written annotations, decision logs, or reflections. The resulting evidence can support both formative feedback and evaluative grading. Such questions target the links in the problem solving process, including:

- Framing: "What are the constraints and success criteria?"
- Planning: "What other approach could work, and why did you reject it?"
- Monitoring: "Where is your first checkpoint? What would signal an error?"
- Evidence: "What would count as convincing evidence here?"
- Transfer: "Where else can you apply this strategy next week?"

One might ask whether a GenAI tool can provide viable answers to the coaching prompts, and this is an ongoing concern of many faculty. To mitigate such risk, a few options can be considered. First, instructors might require that the coaching prompts be completed by students freehand, using “blue books” or other analog means. Second, faculty might require completion of the coaching prompts as a follow up to the submitted work, as an in-class activity. Last, faculty may wish to include a mandatory short oral presentation by students or student groups, during which they can utilize one or more prompts to help verify student learning.

When faculty regularly implement such prompts, two benefits emerge. First, faculty obtain the process evidence they need to better understand the quality of student learning. Second, students get practice in considering their own processes and how they manage the problem solving process in completing their academic work. That experience is then transferable to other contexts. How this can happen in practice is outlined next, with a brief overview of Alex’s “correct error” example from the introduction, and a summary of the student cases presented at the AALHE conference.

Applying the Framework for Assessing Problem Solving

Returning to Alex, his correct answer concealed weak Strategy and Validation process links. He could not justify the formula used, identify the corrected unit error, or independently verify the result. Alex would be rated as “Emerging” in the Strategy & Planning dimension in the rubric above.

Professor Smith might provide SII feedback to Alex as follows:

- Strength: Organization of calculations for load-bearing requirements; this made the procedural steps easy to follow.
- Improvement: Analysis of the tradeoffs between alternative formulas to calculate load, including why the selected formula fits the problem’s constraints; independently verify units before accepting an AI-assisted result.
- Insight: When using GenAI in learning, the learner remains responsible for the decisions that shape the solution and the checks used to validate it.

This case demonstrates why a correct final answer is not, by itself, evidence of competency. The rubric of the problem solving process, such as the seven dimensions included above, plus SII feedback, surfaces the gap in student learning in an observable, scorable way.

Other student work case studies were presented at the AALHE conference. Each of these was developed in collaboration with Pacific Crest's Quillium Studio custom GPT, implemented in ChatGPT (OpenAI, 2026). During the conference session, participants used the cases to identify broken process links, draft a Strength and Improvement, and develop a coaching prompt. Examples of evidence of broken links as exemplified in the case studies are included in Table 3. Participant groups generally identified plausible broken links but found it more difficult to formulate questions that would elicit stronger evidence of student reasoning. The debrief therefore focused on whether participants were rewarding the answer or the logic, whether the feedback went beyond evaluative comments to growth-oriented suggestions for change, and whether the student could demonstrate the skill using another tool.

Table 3

Student Case Analyses Using Problem Solving Links and SII Feedback

Broken link	Evidence
Definition	Solution selected before the problem was scoped
Implementation	Units were not carried through a calculation
Validation	Claim lacks specific, testable evidence
Strategy	Concept does not match the intervention mechanism

The cases illustrated that each link can fail independently, and that a polished final product may conceal gaps in the reasoning process that otherwise would go unnoticed. Finally, participants were presented with an invitation to use the design template in Table 4 to apply directly to their own courses or institutional contexts. The template can aid faculty and assessment professionals in considering parts of the problem solving process and how they can obtain valid and reliable evidence for its presence, as well as coaching questions to enhance the evidence available.

Table 4

Design Template for Process Evidence

Process Link	How to "See" It (Evidence)	Coaching Question
Definition	Problem statement, constraints, success criteria	What problem are you actually trying to solve?
Strategy	Brief rationale comparing approaches	Why is this approach preferable to the alternative?
Implementation	Annotated steps, decision log, or think-aloud report	What adjustment did you make, and why?
Validation	Error check, rubric self-assessment, or reflection	What evidence shows that the result meets the criteria?

The template can be implemented in online or asynchronous courses, lab or studio settings, and also in advising or student services contexts.

Discussion and Implications

This section provides a brief discussion of process-based assessment and its implications at the class and program level.

First, process rubrics can inform program-level learning outcomes reporting without duplicating effort in grading. For example, when faculty use process artifacts in their coursework assignment evaluations, assessment professionals and others can sample these alongside the products for embedded assessment of student learning outcomes. In this way, program and department chairs can obtain stronger evidence of student learning and identify aspects of the problem solving process that are particularly well done, and those that may need improvement.

Second, when multiple members of a department's faculty adopt both process and product-oriented grading practices, multiple raters can now calibrate their assessments around the same four-link problem solving chain. This helps keep scoring consistent and understandable for students, and across individual course sections. Such consistency then adds reliability to any judgments made about program quality and student learning outcome achievement.

Third, process assessment can be a constructive, and growth-producing, alternative to detection-based attempts at academic integrity enforcement in the age of GenAI. Rather than policing AI out of student assignments, student work is assessed on the process of its development, requiring evidence that cannot be satisfied by submitting an unexplained output alone. When the use of AI is approved, students are incentivized to critically evaluate its output and must be prepared to explain its results, rather than just accept them.

One limitation is worth noting. This paper reports work that is exploratory and illustrative in nature. The cases were simulated, the conference audience was self-selected, and the problem solving process rubric has not yet been tested formally for inter-rater reliability.

Connections to Prior AALHE Scholarship

Prior AALHE scholarship in conference Proceedings and *Intersection* has emphasized the value of formative assessment and for assessment professionals to consider their practices in response to GenAI. This article extends the discussion by offering a four-link diagnostic process that makes student thinking observable, and supports the production of specific formative feedback using the SII method. Carney et al. (2022) highlight the importance of formative assessment as a catalyst for student learning improvement in a shorter time frame, and the need for faculty development on the practice. Formative assessment using rubrics, reflection, active monitoring of the learning process, and substantive feedback can support metacognition and learning outcome improvement. By utilizing process-oriented techniques like the problem solving chain and SII feedback, I extend their argument and provide a specific tool for providing such feedback.

In addition, a recent *Intersection* article on adapting to AI in assessment work points out that academic integrity is at risk with the advent of GenAI and its use by students to create work products without the requisite student learning (Miller, 2024). Utilizing the problem solving process links as a part of the student work product submission process not only increases the validity of the course assignments, but provides additional insight into student learning not otherwise captured by the product alone.

Conclusion

Generative AI has weakened the assumption that the quality of a finished assignment accurately reflects the quality of the student learning that produced it. Assessment of learning must therefore move “upstream” to the reasoning, decisions, adjustments, and validation procedures used during the product’s development. Process Education provides a practical framework for making those processes visible and assessable through the four links of definition, strategy, implementation, and validation.

The seven-dimension rubric, SII feedback method, coaching prompts, and process-evidence examples presented here offer faculty and assessment professionals feasible ways to examine student reasoning while supporting improvement in future performances. This approach does not require removing GenAI tools from learning environments; it requires students to demonstrate intentional ownership of their learning process, including their decisions, reasoning, and validation of the quality of their work.

A first step for faculty is to revise one existing assignment by adding a criterion, checkpoint, or reflection prompt that makes one problem-solving link observable. Even a single cycle of process-oriented learning assessment can provide stronger evidence of learning and more actionable feedback than the finished product alone.

References

- Anthropic. (2026, June 8). *Claude Sonnet 4* [Large language model]. Anthropic. Claude
- Apple, D., Ellis, W., & Hintze, D. (2016). 25 years of Process Education (Self-Assessment). *International Journal of Process Education*, 8(1), 59–66.
- Black, P., & William, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139–148.
- Carney, E. A., Zhang, X., Charsha, A., Taylor, J. N., & Hoshaw, J. P. (2022). Formative assessment helps students learn over time: Why aren't we paying more attention to it? *Intersection: A Journal at the Intersection of Assessment and Learning*, 4(1). <https://doi.org/10.61669/001c.38391>
- Duncan, W., Ellis, W., & Apple, D. (2016). *A primer for Process Education: Sharing 25 years of practice and scholarship*. Pacific Crest.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Jones, J. M. (2026, July 13). Confidence in U.S. higher education slips back slightly. *Gallup*. <https://news.gallup.com/poll/712322/confidence-higher-education-slips-back-slightly.aspx>
- Leise, C., Litynski, D. M., Woodbridge, C. M., Ulbrich, I., Jain, C., Leasure, D., Horton, J., Hintze, D., El-Sayed, M., Ellis, W., Beyerlein, S., & Apple, D. (2019). Classifying learning skills for educational enrichment. *International Journal of Process Education*, 10(1), 57–83.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–103). American Council on Education/Macmillan.
- Miller, W. (2024). Adapting to AI: Reimagining the role of assessment professionals. *Intersection: A Journal at the Intersection of Assessment and Learning*, 5(4), 99–113.
- Morgan, J., & Williams, B. (2010). Process Education best practices for teaching open-ended problem solving in a project context. *International Journal of Process Education*, 2(1), 35–49.

- OpenAI. (2026, April 13). *I am seeking sample student submissions for slide 12, for use in breakout sessions. I need five examples that are less than one page (approx. 400 words maximum) that illustrate potential breakdowns in student reasoning* [Generative artificial intelligence chat]. Quillium Studio.
- Perkins, M., Furze, L., Roe, J., & MacVaugh, J. (2024). The Artificial Intelligence Assessment Scale (AIAS): A framework for ethical integration of generative AI in educational assessment. *Journal of University Teaching and Learning Practice*, 21(6). <https://doi.org/10.53761/q3azde36>
- Process Education. (n.d.). *Classification of learning skills*. Retrieved July 20, 2026, from <http://www.processeducation.org/cls/web>
- Van Slyke, A., Utschig, T., & Apple, D. (2021). Improving performance using the Methodology for Developing Performance. *International Journal of Process Education*, 12(1), 3–20.
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice*, 41(2), 64–70.

About the Authors

Joshua Morrison, Director of Academic Retention Programs, University of Indianapolis,
jmorrison@uindy.edu