Instruct

PIDP 3230

EVALUATION OF LEARNING

Author: Yuri Tricys, Date: June 15th, 2025

REFLECTIVE WRITING ONE

vcc Build Hello

Objective

In "The Art of Evaluation. A Resource for Educators And Trainers", Tara Fenwick and Jim Parsons elucidate the 'Principles of Authentic Evaluation', positing that authentic evaluation must be:

- Ongoing
- Ensured in validity and reliability
- Comprehensive
- Effectively communicated
- Methodologically diverse

In a modern educational environment, in which students are more likely to utilize artificial intelligence to produce assignments, it seems prudent to emphasize the relevance of 'validity' and 'comprehension', as Fenwick and Parsons define them.

As such, the emphasis of this reflective writing will be on reaffirming the critical importance of 'validity' and 'comprehension' in authentic evaluation, to ensure not just educational integrity, but also effective education.

Reflective

In terms of 'effective education', Fenwick and Parsons say "evaluation is valid if it measures what it's supposed to measure."

> *Evaluation is valid if it measures what it's supposed to measure. As an example, too often adult learners are subjected to paper-pencil tests asking them to report isolated facts related to a particular topic of instruction, when the actual instructional goals focus on developing broad conceptual understanding in the learners.*
>
> *(Fenwick & Parsons, 2021)*

The intuition is that 'paper-pencil' tests ask learners to report isolated facts, when actual instructional goals focus on developing broad conceptual understanding.

The rise of AI-generated assignments is likely to broaden any disconnect between what is measured and what is intended to be learned, because, at least in terms of papers, like this one, AI can reproduce content without demanding the learner accumulate and apply knowledge that aligns with the learning objectives of the institution, the stakeholders, and the instructor.

My own experience with AI, however, in the context of assignments that are intended to achieve learning objectives, provides a surprisingly unexpected outcome. I've found personally that AI doesn't do a specific enough job. In other words, it doesn't recreate what I'm visualizing for a given page or paragraph. After collecting samples of results created by AI, as an author, I have to go in and read each sample, then dissect it into parts, and then reassemble the various parts with my own writing, in order to achieve my 'vision' for any particular piece … which is interesting food for thought that I will expand on in the next section.

Interpretive

The idea mentioned above, which is that AI is not intuitive enough to create a paper that matches the authors' vision without the author putting in the kind of work that connects learning goals with the process of learning during assignment

completion, does depend on the presence of a 'vision' for the paper in the first place.

My experience as a tutor helping undergraduate students write papers, suggests that a 'vision' for a paper is a higher-level output. Many of the students whose papers I reviewed, had not yet learned even basic academic paper structure. Their papers did not have introduction paragraphs that spelled out any specific thesis or point, followed by body paragraphs that supported and refuted their thesis, with a conclusion that summarized results.

AI produces such papers consistently. It is a champion of structure, which may have a positive effect on encouraging students to learn to think critically in structured patterns. But arguments in support of AI say little about alignment between intended learning outcomes and authentic evaluation.

However the student produces the paper, or writes the test, the results should be evaluated in a way that encourages the application and integration of sought after knowledge and skills. Fenwick and Parsons encourage readers of their text to remember that "what is measured often becomes what is learned" (Fenwick & Parsons, 2021). If instructors measure only recall, we inadvertently signal that memorization is the ultimate goal. Conversely, if we assess higher-order thinking—such as analysis, evaluation, problem solving, and creation—we encourage students to develop these capacities (Cornell University Center for Teaching Innovation, n.d.-b).

Decisional

The performance based nature of backend web development, which is a part of what I'm preparing to teach through this course, is a little easier conceptually in terms of evaluation. Backend web development concepts, at least at the beginning and intermediate levels, are fairly straight forward in that students either know them or they don't, and projects are either functional or they are not.

Front-end web development, on the other hand, which includes design and user experience – the principles of which can be more abstract and subjective – could require more careful design of evaluation criteria.

In general, it's just as important in the era of AI to craft learning goals that are specific, measurable, applicable, realistic, and time-bound (SMART), and to ensure that assessment methods are chosen based on the thinking skills they're intended to measure (University of Toronto Centre for Teaching, n.d.).

As a potential future instructor, I could take extra measures to employ a variety of evaluation methods to ensure more robust testing for comprehension and general learning.

These methods could include:

- Performance-based assessments, such as projects, presentations, or portfolios, which require students to apply their knowledge in authentic contexts (Cornell University Center for Teaching Innovation, n.d.-b)
- Formative assessments, such as reflective journals, peer reviews, and in-class discussions, which provide ongoing feedback and opportunities for students to demonstrate understanding as they learn (Cornell University Center for Teaching Innovation, n.d.-a)
- Summative assessments that focus on synthesis and evaluation, such as case studies or problem-based tasks, rather th mere recall of facts (Behring & Laitusis, 2022)

Moreover, incorporating self-assessment and peer-assessment strategies to foster metacognition and collaborative learning could ensure measurement of not only what students know but also how they think, solve problems, and relate new information to their prior experiences.

Finally, teaching students to visualize completed assignments against a comprehensible and achievable standard, while teaching them also how to use AI to achieve to those standards, could go a long way towards encouraging the achievement of learning outcomes.

Personally, I intend to remain attentive to all of these strategies as well as the other broader factors that influence learning, like accessibility of materials, the creation of a supportive learning community, and comprehensive and authentic approach to both material creation and assignment evaluation.

In this way, if I should teach, I hope to contribute to an educational environment where validity and comprehension are not just ideals, but everyday realities.

References

Fenwick, T. J., & Parsons, J. (2021). The Art of Evaluation. A Resource for Educators And Trainers. *Thompson Educational Publishing Inc.*

University of Toronto Centre for Teaching. (n.d.). *Developing learning outcomes.* https://teaching.utoronto.ca/resources/dlo/

Cornell University Center for Teaching Innovation. (n.d.-a). *Measuring student learning.* https://teaching.cornell.edu/teaching-resources/assessment-evaluation/measuring-student-learning

Cornell University Center for Teaching Innovation. (n.d.-b). *Learning outcome types and recommended assessment methods.* https://teaching.cornell.edu/assessing-student-learning/learning-outcome-types-and-recommended-assessment-methods

Behring, R., & Laitusis, V. (2022, November 5). *Reading comprehension assessment strategies.* HMH Education Company. https://www.hmhco.com/blog/reading-comprehension-assessment-strategies

AI Models Used in Report

The ideas, structure, writing, and editing in this paper were performed by the author. Various AI models were used in collecting data, verifying data, and formatting various arguments.

- Meta Llama 4 Maverick 17b 128e Instruct
- Deepseek R1 Distill Llama 70b
- Meta Llama 4 Maverick
- Qwen 2.5 vl 72b Instruct
- Brave LEO [Qwen 14B, Llama 3.1 8B by Meta, Mixtral 8x7B by Mistral AI]