

Moving Beyond Standardized Metrics: Rethinking Criteria for Evaluating Educational Innovations

Engelbrecht, E.; Strobel, J. M. ; Rooij, R.M.

DOI

[10.5281/zenodo.17631658](https://doi.org/10.5281/zenodo.17631658)

Publication date

2025

Document Version

Final published version

Published in

Proceedings of the 53rd SEFI Annual Conference 2025

Citation (APA)

Engelbrecht, E., Strobel, J. M., & Rooij, R. M. (2025). Moving Beyond Standardized Metrics: Rethinking Criteria for Evaluating Educational Innovations. In R. Kangaslampi, G. Langie, H.-M. Järvinen, & B. Nagy (Eds.), *Proceedings of the 53rd SEFI Annual Conference 2025* <https://doi.org/10.5281/zenodo.17631658>

Important note

To cite this publication, please use the final published version (if applicable). Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.



Workshop

Recommended citation: Engelbrecht, E., Strobel, J., & Rooij, R. (2025). Moving Beyond Standardized Metrics: Rethinking Criteria for Evaluating Educational Innovations. In Kangaslampi, R., Langie, G., Järvinen, H.-M., & Nagy, B. (Eds.), SEFI 53rd Annual Conference. European Society for Engineering Education (SEFI), Tampere, Finland. DOI: 10.5281/zenodo.17631659.

This Conference Paper is brought to you for open access by the 53rd Annual Conference of the European Society for Engineering Education (SEFI) at Tampere University in Tampere, Finland. This work is licensed under a Creative Commons Attribution-NonCommercial-Share Alike 4.0 International License.

MOVING BEYOND STANDARDIZED METRICS: RETHINKING CRITERIA FOR EVALUATING EDUCATIONAL INNOVATIONS

E. Engelbrecht ^{a,1}, J. M. Strobel ^b, R. M. Rooij ^c

^a Delft University of Technology, Delft, NL, 0000-0001-8731-347X

^b The University of Texas at El Paso, El Paso, USA, 0000-0002-2124-1116

^c Delft University of Technology, Delft, NL, 0000-0002-2784-1892

Conference Key Areas: *Building the capacity and strengthening the educational competences of engineering educators; Improving higher engineering education through researching engineering education*

Keywords: *education innovation; educational interventions; evaluation criteria; value indicators*

ABSTRACT

The field of innovation evaluation in engineering education is growing, yet the perspectives of educators and educational support staff are often overlooked. As a result, critical factors such as relevance, context, and long-term value can go unrecognized, limiting both the effectiveness of evaluations and stakeholder engagement. This workshop builds on findings from a Group Concept Mapping (GCM) study conducted at a European university of technology, where engineering educators and support staff collaboratively identified and ranked 104 criteria for evaluating educational innovations. Early findings revealed a strong emphasis on qualitative aspects, such as *workload*, *student well-being*, and *alignment with stakeholder needs*, ranked above concerns such as *scalability* and *regulatory compliance*. The workshop has two aims: to broaden the study by involving an international group of participants from diverse professional roles, and to introduce the engineering education community to a more inclusive and stakeholder-informed view of 'quality'. Participants will collaboratively reflect on and further develop the list of evaluation criteria, and group them into conceptual themes. By the end of the session, participant groups will have explored and developed categorizations of evaluation criteria.

¹ *Corresponding Author*
E. Engelbrecht
e.engelbrecht[at]tudelft.nl

1 BACKGROUND AND RATIONALE

1.1 Current perspectives on 'quality'

There is limited research on how educators and educational support staff, the key stakeholders in implementation, define and prioritize evaluation criteria. Their perspectives are essential for developing meaningful, context-sensitive frameworks, yet both groups are often excluded from formal evaluation processes while faculties fall back on standardized metrics for quality evaluation. This limits relevance and buy-in (Avey et al., 2009; Bajada et al., 2019; Michie et al., 2002). Without their input, critical factors influencing the long-term success of innovations remain overlooked.

Co-creation with stakeholders is one of the principles of curriculum agility (Brink et al., 2024). This framework also highlights the importance of having structures in place to support innovation. The ultimate goal of our research is to develop one such structure by developing a framework for evaluating innovation with academic publication in mind.

1.2 Discovering new perspectives on 'quality'

This workshop draws on a previous study (in review) where data collected using the method of Group Concept Mapping (GCM) (Trochim, 1989). Engineering educators and educational support staff brainstormed, sorted and ranked 104 criteria for evaluating educational innovations at a university of technology in Europe. The criteria were sorted based on similarity to form conceptual themes. The ranking was done based on importance. One of our preliminary findings showed that qualitative criteria related to *workload*, *student well-being* and the *needs and goals of stakeholders* were prioritized over *scalability*, and *regulation and compliance*.

1.3 Towards an international view

During the initial study, the stakeholder group was limited to those in our more immediate, local professional network. This workshop extends our previous work and serves two purposes. Firstly, to expand on this study by including a more international group of stakeholders, occupying a wider range of professional roles. Secondly, to expose the larger engineering community to a different view on 'quality' in education.

The addressed the research question: *How do education stakeholders conceptualize evaluation criteria for educational innovations?* Participants jointly reflected on those criteria and categorize them.

2 WORKSHOP OBJECTIVES

2.1 Target audience

The session is open to anyone interested in the topic of education innovation evaluation and educational quality. The workshop will be most beneficial to those wanting to evaluate educational interventions and innovations for improving instructional design. The evaluation criteria can be used during regular quality management processes, or for educational research initiatives.

2.2 Expected learning outcomes

By the end of the workshop, participants should be able to conceptualize relevant evaluation criteria that align with specific types of innovations.

3 WORKSHOP DESIGN

3.1 Time plan

The schedule for the workshop was as follows:

Table 1. Time plan

Run time	Activity	Notes
10 min	Introduction	Introducing topic and method
10 min	Brainstorming	Discussing statements
25 min	Sorting and group discussion	Sorting criteria to create categorization
15 min	Plenary discussion	Sharing results and wrap-up

3.2 Interactivity

At the start of the workshop, we briefly shared information about our project for developing a framework for evaluating educational innovations and shared the process we went through for completing the GCM study.

Then, participants formed groups. Each group received a workshop pack with the following items:

- Discussion prompt cards for brainstorming. The discussions give participants a brief introduction to some of the evaluation criteria and help them contextualize the criteria.
- Instructions card to guide the activity.
- Persona cards which participants use to select their roles during the workshop. The roles assigned were Coordinator (for managing the group and workshop materials), Presenter (for sharing the results from the discussions), and Academics (their main function being brainstorming) assigned to the remaining group members.
- Criteria cards containing the criteria collected during the GCM study.
- Clips, sticky notes and pens for grouping and categorizing the criteria.

First, participants brainstormed criteria using the discussion prompt cards. This also served as a warm-up activity. Next, they worked through the criteria cards and grouped them according to similarity (conceptually). Each group was assigned a label to categorize them. During the plenary discussion, the categories were compared, and we had further discussions about criteria and evaluation of innovations.

4 WORKSHOP RESULTS

From the workshop, three sets of categorizations emerged. The categories are listed in Table 2, alongside the categories from the original GCM study.

Table 2. Time plan

Group 1	Group 2	Group 3	CGM study
<ul style="list-style-type: none">• Implementation (long-term operation)	<ul style="list-style-type: none">• Motivational effects	<ul style="list-style-type: none">• Sustainable Development Goals	<ul style="list-style-type: none">• Student well-being• Student learning

<ul style="list-style-type: none"> • Legal, practical and stakeholders • Scalability • Faculty (individual educators) • Student learning • Student affective • Design principles of innovation (educational design) • Vague (could be in multiple categories) 	<ul style="list-style-type: none"> • Wellbeing and inclusiveness • Educator loss and benefits • Learning outcome effects • Intended implementation • Sustainability (long-term) • Rules, regulations, policies • Support • Relevance to society/ authenticity • Unintended/ miscellaneous 	<ul style="list-style-type: none"> • Execution • Complications • Motivation and engagement • Wellbeing • Course purpose • Stakeholders • Curriculum • Restrictions 	<ul style="list-style-type: none"> • Efficiency • Student success • Impact on users • Context of innovation • Risks and Benefits • Scalability and resilience • Trade-offs • Stakeholder needs • Suitability • Regulation
--	--	--	---

Some notable similarities and differences that can be observed. All groups identified the student learning experience and included motivation, wellbeing and engagement, indicating a shared belief that successful innovations should have a positive effect on students. This was also reflected in the GCM study, where *student wellbeing*, *student learning* and *student success* were identified as distinct categories.

Likewise, each group recognized the importance of innovations to be practical and manageable over time, with criteria such as *scalability* and *sustainability*. In this regard, participants in the GCM study identified *efficiency*, *risks and benefits*, as well as *trade-offs*.

Furthermore, contextual factors such as various stakeholders and legislative aspects appeared across all groups. In the original study, the *context of innovation*, *scalability and resilience*, *suitability*, as well as *regulation* were identified.

However, the groups differed in the primary focus. Group 1 seemed to emphasize operational and design aspects, creating categories that related to implementation and educational design. Group 2 included more human-centered categories, such as motivation, inclusion, wellbeing and support. Group 3, on the other hand, adopted a broader lens, connecting evaluation of innovation to Sustainable Development Goals, factors that could hinder the innovation (complications and restrictions), and curriculum-related aspects.

All categories created during the GCM study were seen again during this workshop in some form or another, except for two categories: *context of innovation* and *trade-offs*. *Context of innovation* included criteria such as ‘transferability’ (to other courses/faculties/disciplines) and ‘impact of the innovation on other parts of the curriculum’. The workshop participants associated these criteria with curriculum design and implementation-related categories. *Trade-offs* included ‘realistic

timeframe for implementation', and 'time spent on course administration', while the workshop participants grouped these criteria under *restrictions*, or educator-related categories.

Each participant brought with them their own perspectives that are shaped by their professional roles and personal values. Additionally, each innovation is implemented to meet specific purposes in its own unique context. Consequently, it might be neither possible nor desirable to prescribe a fixed list of evaluation criteria. Instead, evaluation guidelines should remain flexible and responsive to their contexts of implementation to facilitate evaluations that are not only informative, but also meaningful.

However, evaluators can use these categorizations to pinpoint the main purpose for having implemented an innovation, and so focus their evaluation efforts to obtain more useful and relevant data that can inform decision making when implementing or improving their educational innovations.

REFERENCES

- Avey, J. B., Avolio, B. J., Crossley, C. D., & Luthans, F. (2009). Psychological ownership: Theoretical extensions, measurement and relation to work outcomes. *Journal of Organizational Behavior*, 30(2), 173–191. <https://doi.org/10.1002/job.583>
- Bajada, C., Kandlbinder, P., & Trayler, R. (2019). A general framework for cultivating innovations in higher education curriculum. *Higher Education Research & Development*, 38(3), 465–478. <https://doi.org/10.1080/07294360.2019.1572715>
- Brink, S. C., De Hei, M., Sjoer, E., Carlsson, C. J., Georgsson, F., Keller, E., McCartan, C., Enelund, M., Lyng, R., & Admiraal, W. (2024). Curriculum Agility principles for transformative innovation in engineering education. *European Journal of Engineering Education*, 1–17. <https://doi.org/10.1080/03043797.2024.2398165>
- Michie, J., Oughton, C., & Bennion, Y. (2002). Employee Ownership, Motivation and Productivity. http://theworkfoundation.co.uk/research/publications/emp_own.jsp
- Trochim, W. M. K. (1989). An introduction to concept mapping for planning and evaluation. *Evaluation and Program Planning*, 12(1), 1–16. [https://doi.org/10.1016/0149-7189\(89\)90016-5](https://doi.org/10.1016/0149-7189(89)90016-5)