Student Co-Created Formative Assessment in a Cross-Disciplinary Data Literacy Module: A Microcurricula Approach at Scale

Yalemisew Abgaz[1]
School of Computing, Dublin City University, Ireland
Corresponding author: YalemisewM.Abgaz@dcu.ie

Designing inclusive and engaging formative assessments at scale across heterogeneous programmes remains a persistent challenge in higher education. This challenge is amplified as data literacy and analytics emerge as critical transversal competency across all disciplines in an AI-shaped graduate landscape. Yet, instructor-designed assessments dominate at scale, lacking student perspective and inclusivity. To my knowledge, no prior study has examined student co-creation of formative assessment within a cross-disciplinary microcurricula framework at institutional scale. This paper presents a co-creation intervention within a microcurricula module in which students from the previous cohort co-designed formative assessments, enabling prospective students to learn directly from peer strategies and experience. These were offered as optional preparation resources for summative assessment across seven Data Literacy and Analytics Topics (DLATs). This paper focuses on two practice-heavy, technology-oriented topics. Among 1,432 and 1,387 enrolled students per topic, 392 and 216, respectively, engaged with the co-created assessments. Independent samples t-tests revealed a statistically significant improvement across both topics. Thematic analysis of student feedback further revealed enhanced metacognitive awareness and increased confidence in assessment preparedness, reinforcing the value of peer-designed resources in cross-disciplinary learning. These findings position student co-created formative assessment as a scalable, equity-aligned strategy for cross-disciplinary STEM education. In doing so, they reposition students as active agents and co-owners of the learning process, with implications for equitable curriculum design across higher education.

Keywords: Assessment as Learning, student co-creation, Active Learning, Universal Design for Learning, Metacognition, STEM education.

Introduction

Designing formative assessment that is both inclusive and engaging across heterogeneous, large-enrolment programmes remains a persistent challenge in higher education (Andersson & Palm, 2017; King, 2022). This challenge has intensified as data literacy becomes increasingly recognised as a transversal skill competency required of graduates in an AI-shaped landscape (Abgaz, 2026b; Reynolds et al., 2020). Despite this, formative assessment at scale remains overwhelmingly instructor-designed, limiting how well it reflects the perspectives, languages, and study strategies of the students who will use it.

Student co-creation offers an alternative: rather than positioning students purely as recipients of assessment, it invites them to act as designers and, in doing so, as co-owners of the learning process. To my knowledge, no prior study has examined student co-creation of

formative assessment within a cross-disciplinary microcurricula framework operating at institutional scale. This paper addresses that gap by presenting Formative Assessments Co-Created with Students (FACCtS), an optional scaffolding embedded within a large, cross-disciplinary Data Literacy and Analytics module (Abgaz, 2026a; Abgaz & Dunne, 2023). Guided by the research question—Do FACCtS improve the engagement and performance of students' summative results?—the paper tries to investigate the impact of FACCtS on the engagement and performance of students, examined through both the summative results students achieve and the way students themselves describe their experience of engaging with peer-designed resources. The paper reports on a mixed-methods evaluation result of FACCtS across two practice-heavy and technology-oriented topics. To this end, I conducted a comparative study of students who participated in FACCtS and those who did not participate. Students who completed FACCtS achieved higher summative scores, and their feedback indicated a similar pattern of positive impact on engagement, describing greater metacognitive awareness, more active engagement with the module content, and a stronger sense of preparedness ahead of summative assessment.

This paper is organised into five sections. The Literature Review Section presents related literature on formative assessments and student participation. The Background Section provides a brief introduction about the module and FACCtS design, while the Methods Section presents the methodology. The Results Section presents the findings, followed by a discussion of their implications in the Conclusion Section.

Literature Review

Formative assessment and feedback is widely regarded as a powerful tool for learning (Boud & Soler, 2016; Broadbent et al., 2018), but its evidence base in higher education remains context-dependent. Andersson and Palm (2017) found that changes to classroom formative assessment practice produced measurable gains in student achievement, but only after sustained professional development that allowed teachers to calibrate the assessment and feedback to their own students. In a systematic review of higher education specifically, Morris et al. (2021) reported that the causal evidence for formative assessment and feedback is comparatively limited, with low-stakes quizzes and peer or tutor feedback showing the most consistent benefits, and with effectiveness heavily dependent on how these approaches are implemented. King (2022) similarly found that online formative assessments were positively associated with student performance, while cautioning that design and alignment with learning outcomes were decisive factors. Taken together, this literature suggests that formative assessment is effective when it is closely calibrated to the students who use it, an assumption that becomes difficult to sustain once a module serves many hundreds or thousands of students with different backgrounds, rather than the single, relatively homogeneous classroom.

This calibration challenge is compounded by two trends in contemporary higher education. First, data literacy and analytics has been identified as a transversal skill that graduates now require, regardless of discipline, as digital and AI-shaped workplaces demand a baseline of data-handling capability across STEM and non-STEM fields alike (Reynolds et al., 2020). Second, institutions are increasingly delivering such transversal content through cross-faculty, modular structures rather than discipline-specific courses. Abgaz and Dunne (2023) describe this as a microcurricula-as-a-service model, in which a single module is designed to serve large, cross-programme cohorts asynchronously and at scale, and Abgaz (2026a) extends this into an integrated framework for interactive, inclusive, and asynchronous learning designed explicitly to accommodate this diversity of learners—including differences in disciplinary

background, learning preference, learning speed, and access to resources. Inclusivity of the framework is achieved through the integration of the three core UDL principles (CAST, 2018)—multiple means of representation, engagement, and action and expression—during the design and implementation stages, as well as a follow-up improvement cycle during and after delivery, serving the interests of these diversified learners. Modules built on this model inherit the general calibration challenge identified above in an intensified form: a single set of formative assessment resources must now speak to students drawn from many different programmes, faculties, and levels of prior data-handling experience simultaneously.

Students as Partners (SaP) and co-creation with students offer one route through this challenge by treating students not as the object of assessment design but as contributors to it. In their foundational systematic review, Mercer-Mapstone et al. (2017) synthesised the empirical SaP literature and found that partnership initiatives were associated with predominantly positive outcomes for students and staff, but that the practice was dominated by small-scale, extracurricular projects involving undergraduates already engaged with their institutions. Bovill (2020) sharpened this critique, arguing that the prevailing focus on small and already-privileged students has left co-creation comparatively under-explored within the mainstream, whole-class curriculum. Focusing specifically on assessment, Chan and Chen (2023) conducted a systematic review of 43 empirical studies and identified four roles students can occupy in assessment partnerships: co-designers, assessors, consultants, and decision-makers, spanning assessment design, execution, quality assurance, and policy, while also confirming that sustained staff support is a consistent precondition for successful partnership.

Read together, this body of work leaves a specific gap that motivates the present study. Formative assessment scholarship establishes that calibration to the learner is central to effectiveness but offers little guidance for modules operating at cross-disciplinary and institutional scale. Student co-creation shows that involving students as assessment co-designers is both feasible and generally beneficial, but has been examined almost exclusively in small, elective, single-discipline settings rather than in large, compulsory, cross-faculty microcurricula particularly focusing on STEM subjects. Universal Design for Learning offers a theoretical bridge between these two literatures: its principle of multiple means of representation and engagement implies that co-created resources can function as an additional, optional pathway through module content, rather than a replacement for instructor-authored materials (CAST, 2018), which is precisely how FACCtS was implemented in this study.

Background

The context for this study is Data Literacy and Analytics for the 21st Century, an online, asynchronous module built on a microcurricula design: a modular, cross-faculty structure that allows a single set of learning materials to serve students from many different programmes at once (Abgaz et al., 2023; Abgaz & Dunne, 2023; O'Donoghue et al., 2014). The module has grown from 10 to 15 programmes across all five faculties of Dublin City University (DCU) and now enrols more than 2,000 students annually (Abgaz, 2026a). Students consequently arrive with markedly different academic backgrounds, prior data-handling experience, and motivations for taking the module, which makes it difficult to calibrate a single instructor-authored approach to formative assessment for the whole cohort.

The module is organised into seven Data Literacy and Analytics Topics (DLATs), each combining conceptual content with applied, hands-on tasks. Two of these topics, Introduction to Spreadsheets (DLAT4) and Intermediate Spreadsheets (DLAT5), are practice-heavy and technology-oriented, and are also delivered using Python and R for students who prefer a

programming-based route. These two topics form the focus of the present study, as their applied, skills-based nature makes them a demanding test case for scaffolding at scale.

Formative Assessment Design

To address the limitations of instructor-only assessment design, the module introduced FACCtS. Six high-performing students from the previous cohort were recruited to co-design formative assessment items alongside the teaching team. The first four students are drawn from three disciplines (Bioprocessing, Social Science, and Psychology) while the other two students come from Computer Science background. Working from their recent experience with the module, these student co-creators drew on the strategies, misconceptions, and study patterns that had helped or hindered their own learning to shape questions that would speak directly to future students. Over the course of the collaboration (600+ hours), the team produced approximately 350 formative assessment questions across the seven DLAT topics plus an additional 150 for three programming-based routes.

Because the co-creators had recently completed the module themselves, they were able to anticipate points of difficulty that instructors, working from an expert perspective, might not readily identify. This positions FACCtS as a form of near-peer scaffolding: prospective students learn not only from the content of the questions but also implicitly from the feedback, strategies, and experience of students who had recently faced the same material.

Implementation

FACCtS items were implemented as optional scaffolding resources, made available at the end of each topic and explicitly positioned as preparation for the corresponding summative assessment. Engagement was entirely voluntary: students could choose to complete some, all, or none of the co-created formative items before attempting the summative assessment for that topic. This optional, low-stakes design was intended to preserve the diagnostic and confidence-building function of formative assessment, rather than adding a further graded requirement to an already large and diverse cohort.

Methods

This research draws on quantitative data collected from the learning platform (Loop) and qualitative data obtained through an interview. Participation in the interview was voluntary and based on participants' explicit informed consent, which was obtained after they had received an information sheet outlining the study. Ethical approval for all procedures was granted by the Dublin City University Research Ethics Committee (Approval Nos. DCUREC/2024/181 and DCUREC/2025/007).

Of all the DLATs, this study focuses on two practice-heavy, technology-oriented DLATs: DLAT4 with a population of 1,432 students, and DLAT5 with a population of 1,387 students. For each topic, students were divided into two comparison groups: those who engaged with the FACCtS, and those who did not. A mixed-methods design was adopted to address the study's objective of understanding the impact of FACCtS on both student engagement and student performance, combining a quantitative statistical comparison of summative outcomes with a qualitative analysis of student feedback on their experience of the co-created resources.

The research question guiding the quantitative strand was: Do FACCtS improve the summative results of students? Student performance was operationalised as the score achieved on the summative assessment for each topic. To answer this question, independent-samples t-

tests were used to statistically compare the summative scores of students who participated in FACCtS with those of students who did not participate, separately for each topic.

The qualitative strand drew on data collected from students about their experience of participating in answering FACCtS. Detailed interviews of five students each spending 20-45 minutes answering 11 questions were conducted. The interview scripts were analysed using Braun & Clark's (2006; 2017) thematic analysis.

Results

Quantitative Findings

Independent-samples t-tests were conducted to evaluate the impact of the formative assessment intervention across two sequential cohorts, DLAT4 and DLAT5. For the DLAT4 cohort, students who participated in the FACCtS program (M = 88.67, SD = 10.05, N = 392) demonstrated statistically significant, higher summative results than those in the Non-FACCtS group (M = 86.54, SD = 10.66, N = 1040), $t(1430) = 3.44$, $p < .001$. The standardised effect size for this difference was Cohen's d = 0.20, indicating a small but practically meaningful educational effect. I acknowledge that cohort-level participation patterns could influence the aggregate results, and this represents a limitation of the current analysis; a stratified analysis by programme/module stream is planned for future work.

A matching trend was observed in the subsequent DLAT5 cohort. Students who attempted the FACCtS (M = 83.52, SD = 11.64, N = 216) scored significantly higher on their summative metrics compared to the Non-FACCtS baseline control group (M = 81.21, SD = 11.93, n = 1171), $t(1385) = 2.65$, $p = .008$. The calculated effect size was Cohen's d = 0.19, mirroring the small yet reliable positive impact identified in the previous cohort. Across both timeframes, the implementation of the FACCtS protocol consistently resulted in an absolute performance lift exceeding 2.1 percentage points in the two practice-heavy topics.

Table 1: Statistics Obtained from DLAT4 and DLAT5 summative result

Topic	Group	N	M	Min Result	Max Result	SD	SE	Skewness
DLAT4	FACCtS	392	88.67	70	100	10.05	0.50	-0.41
	Non-FACCtS	1040	86.53	30	100	10.66	0.33	-0.41
DLAT5	FACCtS	216	83.51	40	100	11.64	0.792	-0.21
	Non-FACCtS	1171	81.21	30	100	11.93	0.349	-0.59

The DLAT4 distributions for both groups are mildly negatively skewed (FACCtS = -0.41, Non-FACCtS = -0.41), indicating a near-symmetric distribution with a slight concentration of high scorers. For DLAT5, Non-FACCtS's distribution is more moderately negatively skewed (-0.592) and leptokurtic (kurtosis = 0.91), indicating a more peaked, less spread distribution with a comparatively heavier tail of lower-scoring students. Across both DLATs, the negative skew combined with modes at or near the score ceiling suggests these distributions are shaped by a concentration of high-achieving students. Furthermore, the mean result improvement in both cases indicates that FACCtS helps low-performing students achieve a better mark by pushing their marks in DLAT4 to the passing mark while resulting in a 10-point improvement in DLAT5 (see Min Result in Table 1). This is consistent with the aim of

FACCtS introduction to provide multiple means of engagement and representation for those students who struggle in achieving the passing mark.

Qualitative Findings

The interview scripts were coded using multiple codes resulting in six initial themes showing consistent support across multiple interview responses. Four strong, but highly interdependent themes emerged and presented below.

Metacognitive Awareness and Self-Regulation (five respondents)

Students reported that evaluating their comprehension via the co-created questions allowed them to actively gauge their conceptual understanding prior to attempting practical applications. This foundational baseline supported proactive self-regulation and systematically reduced performance anxiety ahead of the high-stakes summative assessments. Illustrating this transition, one student shared that the items allowed them "to gauge how much I understood, refresh my memory and be able to make the final assessment less stressful and quicker." (Anonymous participant 1, 2025)

Multiple Means of Representation (five respondents)

The systemic integration of the FACCtS modules offered students alternative pathways to engage with and synthesise core course content. Rather than merely duplicating existing instructional resources, FACCtS served as a distinct, complementary pedagogical tool. The strategic value of this varied exposure was highlighted by a participant who noted that "FACCtS helped me in the actual assessment when I had the time to do them." (Anonymous participant 2, 2025)

Enhanced Active Learning Dynamics (four respondents)

The design infrastructure of the co-created assessments successfully shifted student behaviour from passive review to active application. By prompting immediate feedback, the interactive elements fostered deeper cognitive engagement with the materials. This experiential benefit was echoed in student feedback, which indicated that "FACCtS' interactive element made the overall completion of the DLAT more academically stimulating." (Anonymous participant 3, 2025)

Applied Learning and Theoretical Integration (four respondents)

Participants consistently described the FACCtS items as an effective cognitive bridge linking theoretical principles to hands-on practical execution. This scaffolding was particularly vital for the complex, technology-oriented tasks central to the DLAT4 and DLAT5 curricula. Underscoring this integration, one student concluded, "I believe they were the main factor in actual learning of the module components." (Anonymous participant 4, 2025)

Discussion

Taken together, the quantitative and qualitative findings suggest that FACCtS is best suited to practical, technology-heavy modules, where the significant improvement in summative results aligns with students' own accounts of enhanced metacognitive awareness, active learning, and integration of theory and practice. The consistency of these findings across DLAT4 and DLAT5, two topics with practice-heavy content, points to a broader mechanism: co-created formative items appear to support students in rehearsing and self-assessing applied skills in a way that generic instructor-authored items may not.

These findings also align with core principles of Universal Design for Learning (UDL) for STEM subjects, particularly multiple means of representation and multiple means of engagement, since FACCtS gave students an additional, optional route through the material rather than a single fixed pathway. This positions student co-created formative assessment as a scalable, equity-aligned strategy for cross-disciplinary STEM education, consistent with calls in the literature for formative assessment that is responsive to diverse learner populations (Andersson & Palm, 2017; King, 2022; Reynolds et al., 2020).

At the same time, two considerations temper these findings. First, engagement with FACCtS was voluntary, and uptake was moderate (roughly a fifth of students in DLAT4 and a sixth in DLAT5), so the comparison between groups may partly reflect pre-existing differences in motivation or preparedness rather than the effect of the intervention alone. Second, designing FACCtS well requires careful alignment between assessment items and learning outcomes, and implementation at this scale requires substantial resourcing, both in recruiting and supporting student co-creators and in maintaining the resulting question and feedback bank over time. Future work should examine strategies to encourage broader participation and sustain co-creation processes with fewer resources.

Conclusion

This paper presented FACCtS, a student co-created formative assessment intervention embedded within a large, cross-disciplinary, microcurricula Data Literacy and Analytics module. Across two practice-heavy, technology-oriented topics, students who engaged with the co-created formative items achieved significantly higher summative results than those who did not, and described the resources as supporting metacognitive awareness, active learning, and the integration of theory and practice. These findings position student co-created formative assessment as a scalable, equity-aligned, and near-peer scaffolding strategy for cross-disciplinary STEM education, one that repositions students as active agents and co-creators of the learning process rather than passive recipients of assessment. The results carry implications for the design of equitable, inclusive curricula at scale, while also highlighting the careful alignment and resourcing that such co-creation initiatives require. Future work will include detailed and longitudinal analysis of the impact of FACCtS and periodic co-creation exercise to keep FACCtS relevant for new student cohorts.

Acknowledgment

The author thanks the six student co-creators who contributed their time, insight, and experience to the design of the FACCtS items, and the students of the Data Literacy and Analytics module who engaged with the resources and generously shared their feedback. The author also thanks the DCU Teaching Enhancement Unit (TEU) for providing two rounds of SATLE funding to recruit students for the implementation of FACCtS and the National Forum for the generous support via the SATLE Impact Award.

References

Abgaz, Y. (2026a). An Integrated Framework for Interactive and Inclusive Asynchronous Online Learning at Scale: Data Literacy in Higher Education. Education Sciences, 16(4). https://doi.org/10.3390/educsci16040639

Abgaz, Y. (2026b). A Framework for Formative Assessments Co-Created with Students for Large Classes. 7th PHELC International Symposium (PHELC26). Pedagogy for Higher Education Large Classes. https://doi.org/10.5281/zenodo.20606718

Abgaz, Y., & Dunne, C. (2023). Designing Microcurricula-as-a-service: The case of large class, cross programme, and online asynchronous module. Pedagogy for Higher Education Large Classes. PHELC 2023. https://doi.org/10.5281/zenodo.8014148

Abgaz, Y., McCarren, A., Elger, P., Solan, D., Lapuz, N., Bivol, M., Jackson, G., Yilmaz, M., Buckley, J., & Clarke, P. (2023). Decomposition of Monolith Applications Into Microservices Architectures: A Systematic Review. IEEE Transactions on Software Engineering, 49(8), 4213–4242. https://doi.org/10.1109/TSE.2023.3287297

Andersson, C., & Palm, T. (2017). The impact of formative assessment on student achievement: A study of the effects of changes to classroom practice after a comprehensive professional development programme. Learning and Instruction, 49, 92–102. https://doi.org/https://doi.org/10.1016/j.learninstruc.2016.12.006

Boud, D., & Soler, R. (2016). Sustainable assessment revisited. Assessment & Evaluation in Higher Education, 41(3), 400–413. https://doi.org/10.1080/02602938.2015.1018133

Bovill, C. (2020). Co-creation in learning and teaching: The case for a whole-class approach in higher education. Higher Education, 79(6), 1023–1037. https://doi.org/10.1007/s10734-019-00453-w

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Broadbent, J., Panadero, E., & Boud, D. (2018). Implementing summative assessment with a formative flavour: A case study in a large class. Assessment & Evaluation in Higher Education, 43(2), 307–322. https://doi.org/10.1080/02602938.2017.1343455

CAST. (2018). Universal Design for Learning Guidelines version 2.2. https://udlguidelines.cast.org/

Chan, C. K. Y., & Chen, S. W. (2023). Student partnership in assessment in higher education: A systematic review. Assessment & Evaluation in Higher Education, 48(8), 1402–1414. https://doi.org/10.1080/02602938.2023.2224948

Clarke, V., & Braun, V. (2017). Thematic analysis. The Journal of Positive Psychology, 12(3), 297–298. https://doi.org/10.1080/17439760.2016.1262613

King, D. (2022). The Use of Online Formative Assessments and the Impact on Student Performance. Irish Journal of Academic Practice, 10(2). https://doi.org/https://doi.org/10.21427/cxd5-dz42

Mercer-Mapstone, L., Dvorakova, S. L., Matthews, K. E., Abbot, S., Cheng, B., Felten, P., Knorr, K., Marquis, E., Shammas, R., & Swaim, K. (2017). A Systematic Literature Review of Students as Partners in Higher Education. International Journal for Students as Partners, 1(1), 15–37. https://doi.org/10.15173/ijsap.v1i1.3119

Morris, R., Perry, T., & Wardle, L. (2021). Formative assessment and feedback for learning in higher education: A systematic review. Review of Education, 9(3), e3292.

O'Donoghue, D., Saggion, H., Dong, F., Hurley, D., Abgaz, Y., Zheng, X., Corcho, O., Zhang, J. J., Careil, J.-M., Mahdian, B., & Zhao, X. (2014). Towards Dr Inventor: A Tool for Promoting Scientific Creativity. In Proceedings of the Fifth International Conference on Computational Creativity ICCC-2014 (pp. 268–271). ICCC. https://mural.maynoothuniversity.ie/id/eprint/6465/

Reynolds, K., O'Leary, M., Brown, M., & Costello, E. (2020). Digital formative assessment of transversal skills in STEM: a review of underlying principals and best practice. https://eprints.teachingandlearning.ie/id/eprint/5792/