Formative feedback on engineering self-study: Towards 1 million times per year per cohort

Peter B. Johnson
Mechanical Engineering
Imperial College London
peter.johnson@imperial.ac.uk

Jon Fenton
Physics
Imperial College London
j.fenton@imperial.ac.uk

Phil Ramsden
Mathematics
Imperial College London
p.ramsden@imperial.ac.uk

Robert Chatley
Computing
Imperial College London
rbc@imperial.ac.uk

Maria Ribera-Vicent
Aeronautics
Imperial College London
m.ribera-vicent@imperial.ac.uk

Karl Lundengård
Mathematics
Imperial College London
k.lundengard@imperial.ac.uk

***Abstract*—Timely and specific formative feedback is essential for quality education but challenging to scale for large cohorts. This paper presents Lambda Feedback, a digital platform providing automated real-time feedback to students on self-study. Deployed across multiple disciplines, it supports diverse question types, including symbolic mathematics. Teachers configure feedback without 'coding'. Feedback is provided by an external 'microservice', enabling anyone in the community to develop feedback technology that can plugin to the platform. User surveys show a positive effect on learning experiences. Currently a work-in-progress, the platform provides feedback a million times annually in total. Future work will expand feedback functionality and scale to one million feedback events per cohort per year, i.e. billions per institution per year.**

***Index Terms*—formative feedback, automation, software**

I. Introduction

Formative feedback is key to effective education ([1], [2]), especially when it is timely, specific, constructive, actionable, and frequent [3]. Feedback on self-study is challenging due to the range of times and locations of study; for large cohorts the challenge is amplified. We share work-in-progress on a digital platform providing automated feedback on self-study, currently delivering one million real-time feedback events annually.

We focus here on informative and corrective feedback [4]. The lack of consensus defining what constitutes 'good' feedback ([5], [6]) means feedback design is local and contextual. Software for automating feedback therefore needs to facilitate teachers configuring their feedback.

We present a software platform, called Lambda Feedback, that has been deployed to staff and students in the last three years. Teachers curate content and configure automated feedback. Students access exercises and submit responses for real-time formative feedback. We summarise design, deployment, feedback technology, and evaluation.

II. Conceptual Design

Lambda Feedback emphasises 'Feedback on my homework', i.e. on the activities central to learning and not an extraneous load (such as an additional assessment). The goal is to study in a natural and productive way, with 'technology' in the background. An important step towards this goal, especially when considering mathematics, is the option to input handwritten symbolic expressions as illustrated in Figure 1.

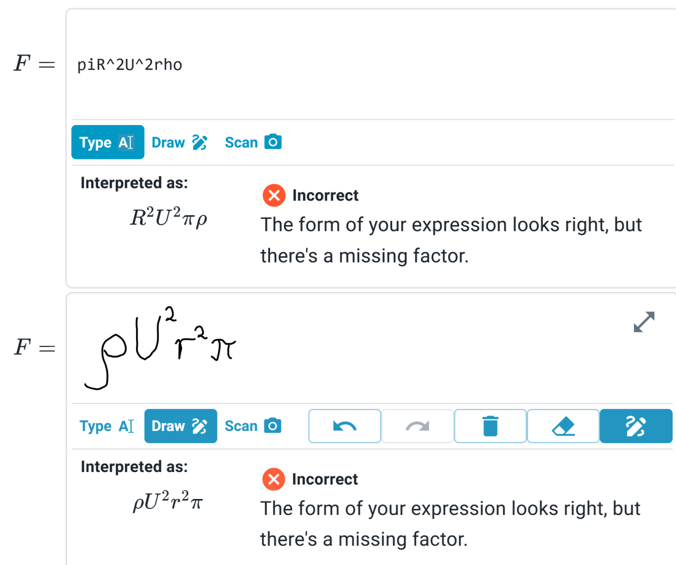

Fig. 1. Example student input of symbolic mathematics and real-time feedback. Above: typed input. Below: hand-written input (can be direct on screen or scanned). Here the student response is missing a factor of (1/6).

The web can be used to make content more accessible to some users, but backward compatibility with paper 'tutorial sheets' is also important for many. We therefore manage a single content source to be published both on the web (Figure 1) and to a LaTeX-rendered PDF.

III. Implementation

At its base, Lambda Feedback is a content management system where teachers curate exercises, with worked solutions, and publish them to students. Additionally, interactive

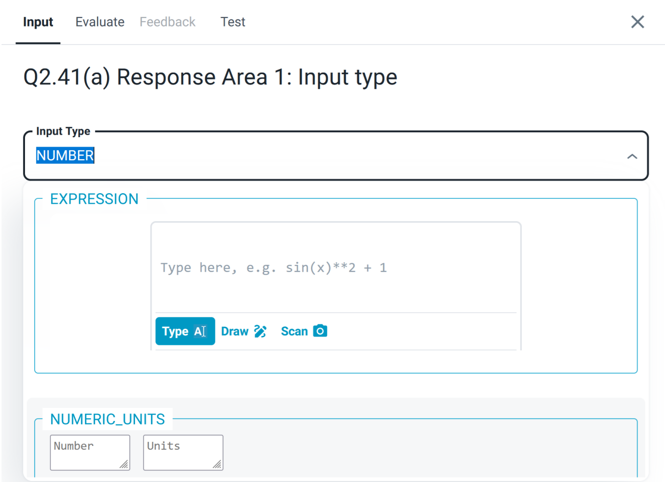

Fig. 2. Feedback configuration step 1: select an interactive component.

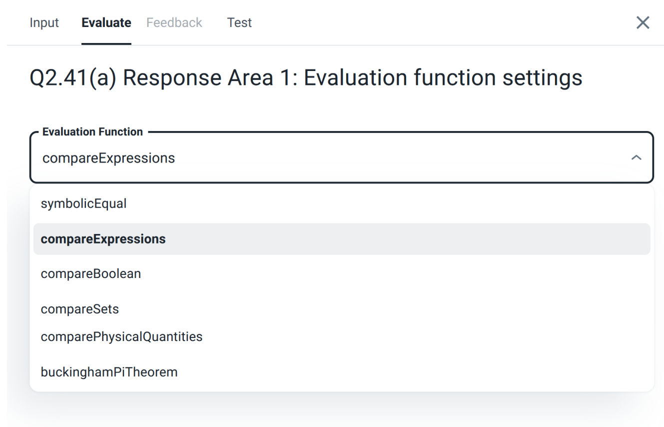

Fig. 3. Feedback configuration step 2: select an 'evaluation function' and configure its parameters (typically the default is adequate). No computer code is required by the teacher.

'Response Areas' are provided for students to express their responses and receive real-time feedback; for example Figure 1.

Important concepts in the development of the software platform included:

- *Microservice feedback*: any technologies (agnostic to language, infrastructure, or development environment) can be connected to provide automated feedback.
- *No-code configuration*: teachers focus on teaching, not on technology. Feedback is curated and configured without writing any code (Figures 2–3).
- *Data analytics* as a primary purpose of the platform: for direct value to users (actionable information), and long-term benefit informing feedback technology.

The software was first built and deployed in 2021/22 and is now used across the authors' institution in over 60 courses covering engineering, sciences, business, and medicine. Applied mathematics is the most common subject. Trials with other institutions in secondary and higher education are now underway, including for English Language.

At the time of writing there were 6,000 questions on the system, with 100+ teachers, and approximately 3,000 active students. The dominant question type (approximately half of questions) processes responses with symbolic expressions, which are more complex to check automatically than multiple choice or simple numeric answers. In total the platform provides automated feedback over 1 million times per year.

IV. Feedback Technology

Feedback is generated by comparing a student response to reference criteria provided by the teacher, using an 'Evaluation Function' . These functions are separated from the core of the platform and can be developed by anyone in the community (see Table I for examples). Evaluation Functions can analyse student answers in different ways, for example using computer algebra systems (CAS), or more recently Large Language Models (LLMs). Evaluation Functions can be written in any programming language, with recent examples in Python and Wolfram language.

Evaluation functions all provide feedback 'out of the box' when allocated to a new question. The level of detail of feedback depends on the algorithm. For a specialist algorithm, such as `buckinghamPiTheorem`, feedback is always (including on first use) detailed and specific, even given the infinitely many possible student responses [8]. For a more general algorithm, such as `compareExpressions`, the initial behaviour is to validate responses as correct/incorrect. More granular feedback on specific mistakes can be provided through additional teacher input after reviewing initial data, such as in Figure 1.

Deeper analysis of collected data over time can also be used to improve the function performance 'out of the box'. For example, our data analysis shows that it is common for the response and reference to differ by a ratio of a real number. More advanced learning from the data is an ongoing project.

V. Evaluation

A user survey is summarised in Table II showing high overall satisfaction. More detailed questions showed users found

TABLE I
A sample of open-source 'evaluation functions' from the community and available on Lambda Feedback (details in [7]).

Function	Technical summary
`compareExpressions`	Parameterised, advanced parsing of symbolic expressions. Uses `symPy`. Compares teacher and student expressions.
`buckinghamPiTheorem`	Validates groups in dimensional analysis. Applies to fluid mechanics. Based on [8].
`compareBoolean`	Parses logical expressions to compare truth tables. Uses `symPy` and a custom parser. Applies to basic electronics and computing.
`evaluateProof`	Feedback on mathematical proofs written in natural language. Uses LLMs.
`GCSEenglish`	Feedback on essays for GCSE English (a UK qualification at age 16).

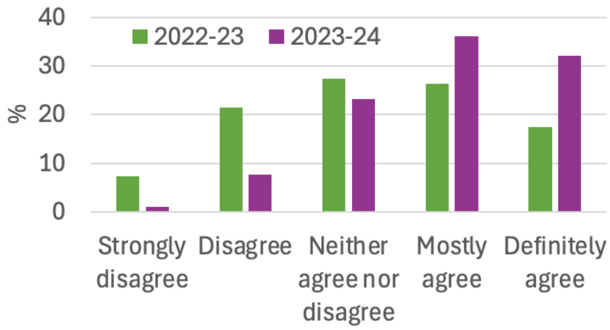

Fig. 4. A general survey run by the central institution, covering all aspects of learning, not just Lambda Feedback. Statement: 'Feedback throughout the module helped me to develop and improve my learning'. Responses are for first year undergraduate Physics cohorts (n = 418). Lambda Feedback was deployed in the year 2023-24, while other teaching was relatively unchanged.

TABLE II
A SURVEY OF PLATFORM USERS, ACROSS SUBJECTS (N=91)

Statement	Positive	Neutral	Negative
"I would recommend Lambda Feedback to other students"	92%	4%	3%
"More modules should use Lambda Feedback for the problems"	92%	2%	6%

the software to be accessible and enjoyable to use, with an emphasis on the value of step-by-step worked solutions. Users said that using Lambda Feedback helped them to continue with their study where otherwise they would have been 'stuck'. Over 90% of users found the analytics helped choose the next task, and seeing progress was motivating.

To evaluate the wider impact of using Lambda Feedback, we can use institution-wide surveys that cover whole modules. An example of data on the topic of feedback is given in Figure 3, representing a general trend that students perceived feedback to be better when Lambda Feedback was used.

VI. Discussion and Future Work

Evaluation of Lambda Feedback has been positive. We attribute this to building requirements from a community of staff and students, ensuring we meet user needs. We will continue to prioritise new features based on community input.

Feedback by microservice encourages contributions from across the education community, the feasibility of which is demonstrated by example: all of the functions in Table I were developed by people outside of the software development team at Lambda Feedback; indeed some were in different organisations. Our focus in the near future will be to foster a wider community of developers of evaluation functions to cover more academic specialisms.

Large Language Models (LLMs) represent a big opportunity, and challenge, to improve real-time feedback. Feedback on questions that previously wasn't possible, such as on proofs or essays, is now possible. We can also contemplate feedback on *method*, as opposed to answers; and dialogic (conversational) feedback. These are areas currently under development.

Data analytics was claimed here to be central to the value-proposition of Lambda Feedback, however currently most of the value of the data the system collects — a fingerprint of students studying throughout the year — is yet to be realised. Current efforts are on identifying the most valuable data to display to users. For example, for students showing study behaviours that indicate good/poor practice, like looking at the answer before making a genuine attempt. For teachers, the focus is on identifying student progress, which may become the basis of class timetables in a 'just-in-time' approach to classroom activities.

Feedback quality depends on robust evaluation functions, which require significant effort to develop. Large scale deployment is needed to justify such investment, but also to meet demand. Deployment so far, at a *total* of 1 million feedback events per year demonstrates initial feasibility. We anticipate a need for 1 million events *per cohort* per year[1], leading to billions of events per institution. Future engineering will be focussed on this increase of scale by orders of magnitude.

Acknowledgment

The work presented here was funded by Imperial College London Digital Innovation Fund. We thank Jarka Baker, Tim Alby, Max Bull, Pierre Thareau, and Louis Manestar for the technical work building the software. We also thank the many students and colleagues whose input was invaluable to building software that meets the needs of its users.

References

[1] J. Hattie and H. Timperley, "The power of feedback," *Review of Educational Research*, vol. 77, no. 1, pp. 81–112, 2007. [Online]. Available: https://doi.org/10.3102/003465430298487

[2] V. J. Shute, "Focus on formative feedback," *Review of educational research*, vol. 78, no. 1, pp. 153–189, 2008.

[3] B. Wiśniewski, K. Zierer, and J. A. C. Hattie, "The power of feedback revisited: A meta-analysis of educational feedback research," *Frontiers in Psychology*, vol. 10, 2020.

[4] G. Deeva, D. Bogdanova, E. Serral, M. Snoeck, and J. De Weerdt, "A review of automated feedback systems for learners: Classification framework, challenges and opportunities," *Computers & Education*, vol. 162, p. 104094, 2021.

[5] A. A. Lipnevich and E. Panadero, "A review of feedback models and theories: Descriptions, definitions, and conclusions," in *Frontiers in Education*, vol. 6. Frontiers, 2021, p. 720195.

[6] E. Panadero and A. A. Lipnevich, "A review of feedback models and typologies: Towards an integrative model of feedback elements," *Educational Research Review*, vol. 35, p. 100416, 2022.

[7] "Lambda feedback documentation." [Online]. Available: https://lambda-feedback.github.io/user-documentation/teacher/reference/evaluation_functions/

[8] K. Lundengård, P. Johnson, and P. Ramsden, "Automated feedback on student attempts to produce a set of dimensionless power products from a set of physical quantities that describe a physical problem," *International Journal for Technology in Mathematics Education*, vol. 31, no. 3, pp. 117–124, 2024.

[1] $\sim$ 1k multi-part exercises per student per a year, with multiple attempts