RESEARCH **Open Access**

Student perceptions of an open-source smart worksheet for automated student assessment and feedback built using python

Samuel C. Perry[1*]

*Correspondence:
Samuel C. Perry
s.c.perry@soton.ac.uk
[1]School of Chemistry and Chemical Engineering, University of Southampton, Southampton, UK

Abstract
Smart worksheets are an emerging educational technology that supports self-directed learning and provides targeted feedback on student assessment, while also reducing staff marking workload. However, their adoption may be limited by high cost, limited flexibility, and proprietary ownership. This study presents the design, implementation, and evaluation of an open-source smart worksheet developed using Python. This worksheet provides a flexible, interactive platform for performing calculations, analysing data, and receiving real-time feedback. It supports automated grading, dynamic content updates and graphical data analysis. Student feedback indicates high satisfaction with the worksheet's ability to enhance conceptual understanding and streamline the analysis process, demonstrating that students perceive similar learning benefits to commercial alternatives. Importantly, its open-source nature ensures adaptability across academic disciplines, enabling its use in diverse educational contexts to support assessment, feedback, and self-directed learning.

Keywords Smart worksheet, Interactive learning, Self-directed learning, Feedback, Open-source, Python, Accessible learning, Automated assessment

1 Introduction

Modern education in science, technology, engineering and mathematics (STEM) increasingly demands tools that bridge theoretical concepts with practical application, particularly in quantitatively intensive disciplines like chemistry [1]. Students frequently encounter difficulties applying mathematical skills to specific tasks, such as interpreting graphical data or completing complex calculations [2]. These challenges are compounded by the diversity of student backgrounds, particularly in practical STEM education [3]. Variability in prior hands-on experience and unequal access to STEM learning opportunities make it challenging for educators to provide consistent support across entire cohorts [4, 5].

An additional challenge in practical chemistry comes with an often-asynchronous nature of classroom and laboratory modules. Students typically rotate through different experiments due to limitations in equipment and resources. Consequently, those in earlier rotations may experience a higher cognitive load [6] and struggle to contextualise results, which can negatively impact learning outcomes and confidence [7].

To address these issues, educators have increasingly turned to digital learning tools that offer structured guidance, formative assessment, and timely feedback. Research into interactive e-worksheets demonstrates that embedding such tools within problem-based or inquiry-based pedagogies can significantly enhance problem-solving abilities, creative thinking, and engagement [8–11]. Key features, such as rapid feedback and opportunities for repeated practice, contribute to improved learning efficiency and support inclusive learning environments [12].

1.1 The importance of feedback

Providing clear, constructive and actionable feedback is a powerful mechanism for supporting student learning [13, 14]. Effective feedback should be timely and offer students the opportunity to reflect and improve their performance [15, 16]. High-information feedback, that is specific, explanatory and aligned with learning goals, has been shown to positively influence student learning and development [17].

In STEM disciplines, feedback is critical due to the complexity of translating abstract concepts into experimental practice [18]. Studies have demonstrated that timely, high-quality feedback can enhance student confidence, promote self-regulated learning, and improve problem-solving skills [19]. However, delivering personal, high-frequency feedback in large laboratory cohorts remains challenging. The demands of equipment preparation, data validation, and safety oversight often limit the time available for individualised instruction and feedback [20]. Digital feedback systems have emerged as a promising solution to these constraints [12]. A recent example is the RATsApp automated feedback system, which provides students with feedback on formative assessments to support learning [21].

1.2 Smart worksheets in pedagogy: strengths and limitations

Recent advancements in educational technology have introduced smart worksheets as an emerging pedagogical tool in science education [22]. Smart worksheets provide real-time targeted feedback to students, integrating feedback with instructions and automated assessment to facilitate self-directed learning in a supported environment [23]. These systems can deliver summative assessment and adaptive problem sets tailored to individual learning trajectories. Feedback may be scaffolded with introductory courses or inclusive learning environments in mind, or designed to promote conceptual thinking and metacognitive reflection [24].

This approach is particularly useful in practical STEM education, where concurrent hands-on data acquisition and analysis can enhance learning [25]. Commercial platforms such as LearnSci, Aktiv Chemistry and Tecquipment have demonstrated effectiveness in improving examination performance, fostering higher-order thinking, and reducing instructor grading burdens [26–28]. Empirical studies show that smart worksheets can identify specific skill gaps, support critical thinking, and improve engagement, confidence and performance among undergraduates in chemistry [29, 30] and pharmacy [31] courses.

Smart worksheets are typically embedded within inclusive curriculum frameworks, guiding students through assessed tasks with specific feedback and automated grading [24]. They may also be used to support flipped classroom models [32] or as progress checks within broader assessment strategies [33]. Although current literature has focused primarily on applications in chemistry, the underlying principles are applicable to any discipline where assessment is measured by defined correct answers, making smart worksheets a flexible tool across higher education and beyond.

Despite their benefits, reliance on proprietary systems presents notable challenges. High licensing costs limit accessibility to underfunded institutions and closed-source architectures restrict pedagogical flexibility and content customization. Additionally, updates to commercial smart worksheet content often depend on vendor timelines, hindering rapid adaptation to student feedback or curriculum changes. These limitations highlight the need for an open-source, flexible alternative that retains the pedagogical strengths of smart worksheets, while enhancing accessibility and adaptability.

1.3 Open-Source solutions for digital learning tools

Open-source frameworks offer a compelling alternative to develop digital learning tools, providing cost-effective, customizable platforms for the development of a wide range of digital educational resources, including smart worksheets. This study focuses on Python, a widely used open-source programming language with robust libraries for numerical calculations, data visualisation and statistical analysis [34]. By integrating Python with the tkinter module for graphical user interface (GUI) design, educators can create tailored smart worksheets based on specific pedagogical goals.

Unlike commercial systems, a Python-based open-source smart worksheet al.lows educators to build modular, adaptable assessments with full control over content, grading logic, and data management. Key advantages include the ability to make real-time edits, maintain local gradebooks, and incorporate discipline-specific datasets. This flexibility supports creation of a tailored educational experience based on specific learning needs, which aligns with growing demand for accessible and adaptable educational technologies [35, 36].

2 Study aims

The pedagogical effectiveness of smart worksheets in supporting learning has been well documented, particularly in STEM education [29–31]. This study presents a new method for building an open-source smart worksheet using Python, which we have designated OSPREY (Open-Source Python Responsive Evaluation sYstem). This study evaluates the student perspective on its efficacy to support an undergraduate electrochemistry experiment. The open-source smart worksheet guides students through data analysis, while providing real time feedback and automated assessment. We assess whether this system is perceived by students to replicate the benefits of commercial smart worksheets, including enhancing conceptual clarity and reducing cognitive load, while overcoming cost and customization limitations. In this way, we demonstrate that our open-source approach offers a viable, cost-effective and accessible alternative for implementing smart worksheet technologies in higher education.

3 Methodology

This study employed a mixed-methods approach to evaluate the effectiveness and usability of an open-source smart worksheet in an undergraduate chemistry laboratory module. The primary aim of the data collection was to assess: (i) students' perception of the worksheet's usefulness, fairness, and feedback quality; (ii) the impact on students' confidence and understanding of the experiment; and (iii) overall user experience and preferences for integrating such tools into laboratory teaching. The questionnaire was given to 55 first year undergraduate students after completion of the smart worksheet-supported laboratory session. The response rate (n) is given in the caption of all presented data. Values of n lower than 55 are due to students choosing not to submit a questionnaire at the end of the session. No data were excluded from the study.

Quantitative data were collected using a structured questionnaire with Likert-scale items (1–10). Likert data are treated as ordinal, so median values are reported. Questions were designed to measure student perceptions of the laboratory experiment and the smart worksheet. To minimise response bias and enhance clarity, each question was framed in straightforward language with clear endpoints.

To help students conceptually separate their experience of the practical component versus their experience using the smart worksheet, the questionnaire included targeted questions addressing each aspect separately. The questionnaire also included three open-ended questions that allowed students to describe what they liked or disliked about the worksheet and to provide additional comments.

To evaluate the internal consistency of the questionnaire, Cronbach's alpha (α) was calculated for the seven questions evaluating perceptions of the smart worksheet. Cronbach's alpha is a widely used metric for determining the reliability of survey instruments in educational research, with values above 0.7 generally considered acceptable [37]. The calculated value of $\alpha = 0.89$ indicates strong internal consistency, suggesting that these items reliably measure a single construct of student perceptions of the worksheet.

3.1 Open-source smart worksheet design

The open-source smart worksheet design is built using Python's tkinter module [38], a lightweight and flexible cross-platform GUI library. The tkinter module provides an intuitive interface for developing interactive applications, offering a wide range of widgets including buttons, text fields, images, canvases and labels. These elements allow for dynamic content updates and pop-up windows to contribute to an active learning environment [39], making the platform well-suited for delivering interactive real-time feedback. The worksheet code can be packaged as an executable file using the PyInstaller module [40], facilitating straightforward dissemination without requiring proprietary software (Fig. 1).

Designed as a self-contained assessment tool, the smart worksheet presents questions alongside input fields or checkboxes. Students submit response via a "submit" button, which triggers automated evaluation against predefined correct answers. Correct responses earn full marks, while incorrect response prompt a tailored feedback pop-up window. Students may then revise their answer based on the feedback for a reduced mark, promoting self-regulated learning [41] and reducing staff workload. A "solve" button is also available, which provides the correct answer for zero marks, and a more detailed feedback response that explains the correct answer.

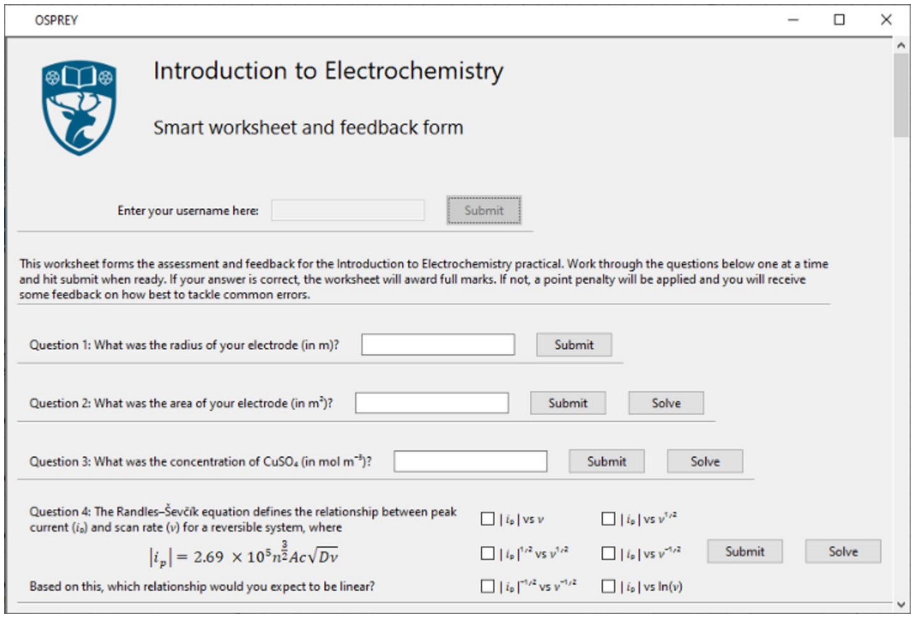

Fig. 1 Screen shot of the open-source smart worksheet, OSPREY, that has been designed to support the "Introduction to Electrochemistry" practical. This smart worksheet is coded using the tkinter module within Python and is fully open-source. The full commented code for this worksheet is freely available (see declarations for access instructions)

3.2 Questions and feedback mechanism

Questions within the smart worksheet are based on the same structured design. The educator defines the question, the correct answer, the maximum score, the penalty for an incorrect answer, and some general feedback for incorrect responses. Additionally, specific feedback can be programmed for common errors, allowing the worksheet to address misconceptions directly. This is especially useful in STEM education, where common mistreatment of equations or misunderstanding of unit conversions can lead to significant errors in student responses.

The open-source smart worksheet supports two types of correct answers:

- *Given answers*: Predefined responses, either as text or checkbox selections, that are directly compared to the student input.
- *Calculated answers*: Dynamically generated answers based on the same calculation as is expected from the student, ensuring consistency and fairness.

For complex, multi-step calculations, assessments can be broken down into stages, each with its own feedback mechanism. This scaffolded approach helps to reduce cognitive load and guides students through the problem-solving process [42].

Calculated answers require attention to precision. Students who round intermediate values may produce slightly different final answers, which can lead to frustrations if they perceive they have been penalised despite performing the correct calculations. From a pedagogical standpoint this can be an advantage, as learning the importance of precision in data handling is a useful skill. To support precision, there are two built-in automated feedback checks:

- *Significant figures check*: Ensures students have entered a sufficient amount of significant figures, as defined by the user.

- *Tolerance check*: Flags answers within 5% of the correct value, suggesting a rounding error is a likely cause.

Both of these checks come with custom feedback that tells students the nature of the error and highlights the importance of precision in quantitative analysis (Fig. 2).

3.3 Integration of graphical analysis

Graphical analysis capability is integrated using Python's Matplotlib module [43]. Students can input x and y data, which the worksheet dynamically plots within the GUI. Data may be entered as comma separated numbers or as a column of data directly copied from Excel, enhancing usability with larger datasets. The worksheet exports an image of the graph, and then reloads it into the smart worksheet window, allowing for a secondary analysis of graphical presentation on top of the automated grading.

The dynamic update function is well suited for graphical analysis. A "Preview" button allows students to adjust parameters and observe changes on the graph in real time. In our implementation, students label axes using Python syntax, reinforcing coding skills alongside chemistry learning. This approach could be applied to any of the wide range of formatting controls available through Matplotlib, from simple graph formatting to complex baseline corrections and data selection.

3.4 Scatter graphs and related analysis

The smart worksheet design supports both scatter or linear graphs (Fig. 3). A key analytical technique for a linear scatter graph is to extract values from its gradient. In this study, students calculate the diffusion coefficient of copper ions in solution by analysing linear trends derived from cyclic voltammetry data [44].

To support this, a linear regression is performed on student data to determine the gradient and associated error. The worksheet then guides students through step-by-step

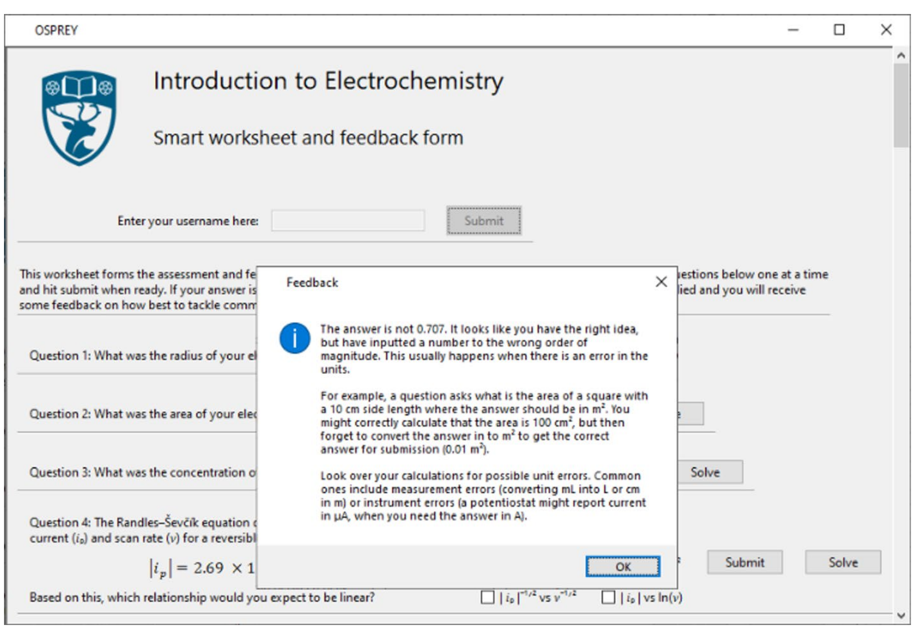

Fig. 2 Screen shot of a built-in error message in response to an incorrect answer from a student. Here, the student has made an error in their unit handling when asked to report the area of an electrode in square metres. The smart worksheet has spotted this and provides advice on checking the conversion of units during calculations

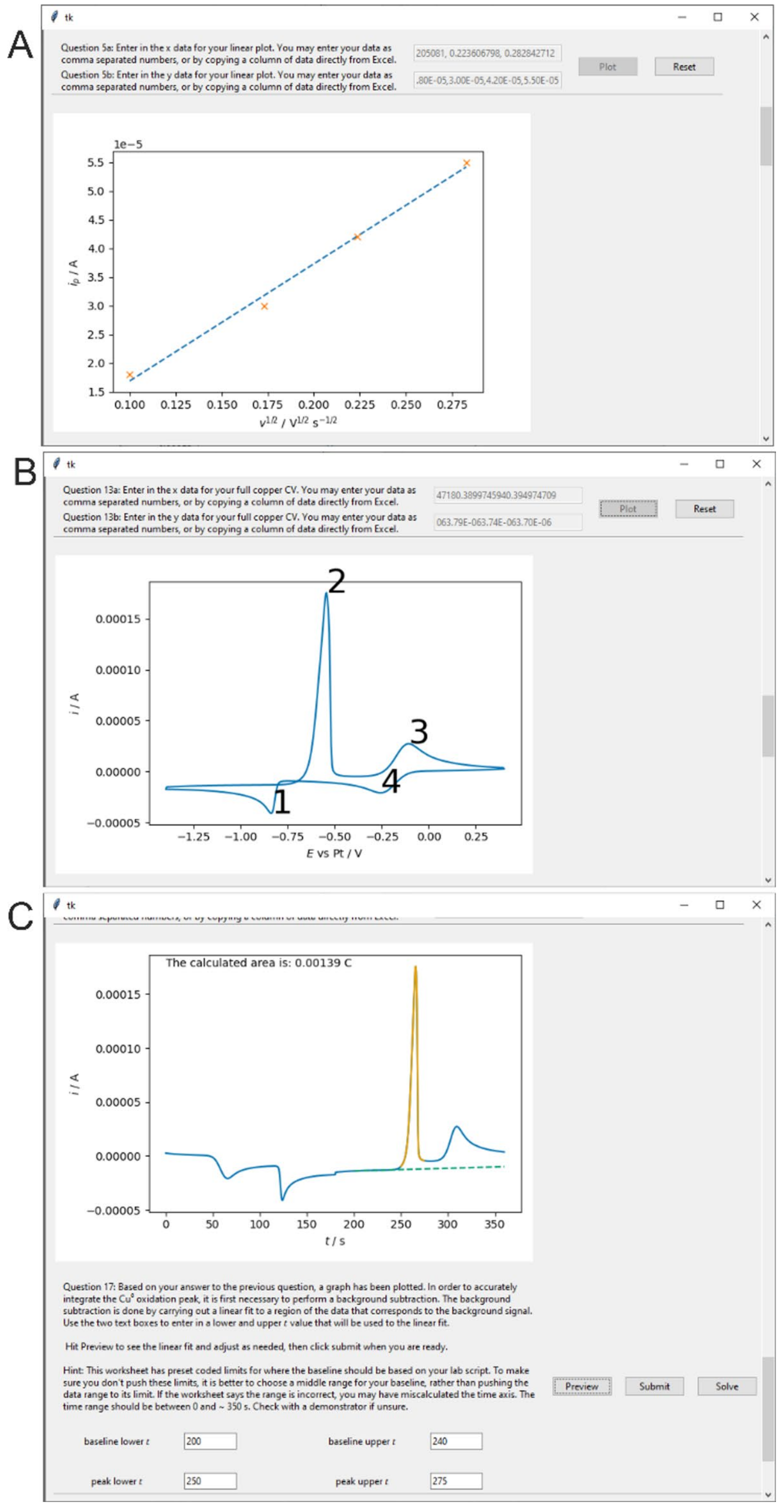

Fig. 3 (See legend on next page.)

calculations to determine both the diffusion coefficient and its uncertainty. Feedback is tailored to the specific values entered, allowing the system to identify and address common mistakes such as incorrect unit conversion or rounding errors.

(See figure on previous page.)

Fig. 3 Demonstrations of the graphing functions available within the smart worksheet. **A** A scatter graph displaying a linear trend. Students enter the data, and the worksheet plots the graph and displays a linear trendline. The worksheet performs a linear regression analysis, then guides the student through that same analysis. Student inputs are compared to calculated answers, then automated assessment and feedback are provided. **B** A line graph for a cyclic voltammogram of copper sulfate. The worksheet searches for peaks in specific regions of the graph and numbers them to correspond with further questions. **C** An alternative question format for a line graph. Students calculate the area under a peak on the graph. Students can select a range in their data to choose the peak they want to analyse (orange) and the baseline they want to use for background correction (green). These features directly address the research aim of evaluating whether an open-source smart worksheet can guide students through complex data analysis while providing targeted feedback

This is a challenging analysis for students, as it combines a new experimental technique (cyclic voltammetry) with a complex calculation. Student cognitive load is often increased by a lack of familiarity with calculations using spreadsheet software such as Excel [45], and asynchronous learning of electrochemistry theory. The smart worksheet is therefore a valuable tool in bridging these gaps by providing in-the-moment support and personalised feedback.

3.5 Line graphs and related analysis

There are two built-in modes of linear graph analysis, tailored to different assessment objectives:

- *Peak identification*: The worksheet displays the graph and automatically detects peaks within a predefined region (Fig. 3B). Students correlate the labelled peaks with a specific chemical reaction. Here, this supports electrochemical interpretation, although the same structure could also reinforce spectroscopic analysis.
- *Peak integration*: Students select a peak and define a baseline region using input fields. A "Preview" button allows dynamic adjustments before the worksheet performs a background correction and integration. The worksheet guides students to use the calculated area to extract analytical information, helping develop the analytical skill while receiving step-by-step support.

As with the linear regression example, the worksheet performs the integration and subsequent calculations. It then evaluates student inputs against the correct answers, providing tailored feedback as needed.

3.6 Grading and assessment

The structured grading system is designed to support formative assessment and encourage self-regulated learning. Full marks are awarded for submission of a correct answer at the first attempt, with reduced marks for subsequent correct submissions. The grading logic follows a structured decision tree:

- *Correct answer*: The worksheet calculates the final mark based on attempt count, applying penalties accordingly.
- *Incorrect answer*: The worksheet determines the most relevant feedback based on the type or error. The worksheet then adds to a running tally of incorrect to appropriately score the eventual correct answer.

Educators can define both the max score and the penalty per incorrect answer in the Python code. Student progress throughout the worksheet is tracked with the addition of labels that display individual scores (Fig. 4). Submission of a correct answer will disable

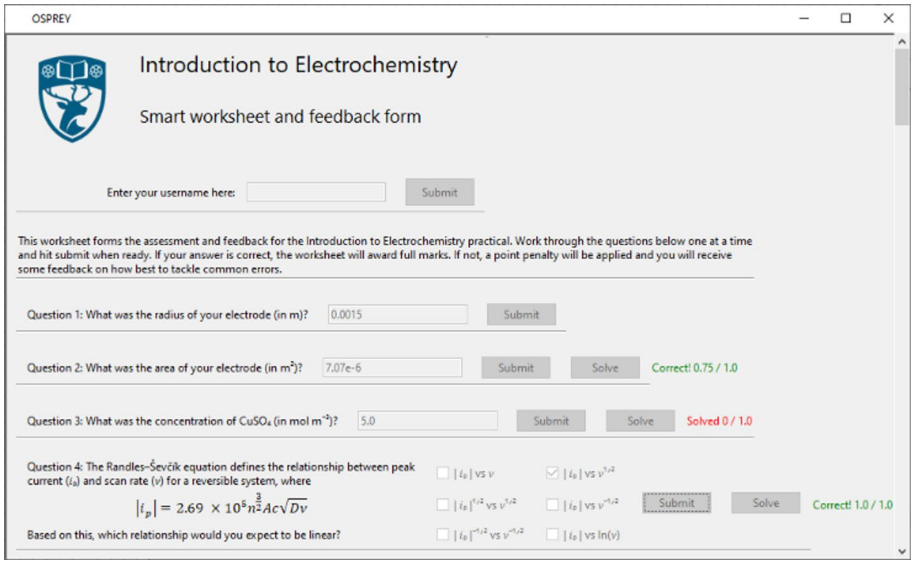

Fig. 4 A scored section of the open-source smart worksheet. The student first entered in the radius of the electrode they used. This was not scored, but is used for reference in the rest of the worksheet. The student then answered the second question correctly on the second attempt and received a reduced score. They then could not answer question 3, so used the "Solve" button to get the correct answer, along with some feedback. Finally, they correctly answered the checkbox question on the first attempt for a full mark. This demonstrates the grading logic and feedback mechanisms in action, facilitating formative assessment and self-regulated learning

the submit and solve buttons and reveal the awarded score next to the relevant question. A final submit button at the end of the worksheet then totals up the final score.

3.7 Data storage

During development, three methods were investigated for recording and reporting student scores (i) storing data directly within the code; (ii) saving data as a local text file; or (iii) exporting data remotely to GitHub. The choice between these was defined by three key parameters: security, submission speed, reloadability (Table 1).

Although storing data within the code is computationally fastest, this is not viable due to the loss of data when the worksheet is closed. Exporting to GitHub using the PyGithub module [46] offers the greatest security, since all progress is stored separately from the student. However, a delay of five seconds between submitting an answer and receiving a response would lead to frustration for the student user. It has been demonstrated that a poor user experience with digital learning tools can increase anxiety [47], decrease engagement [48] and negatively impact learning outcomes [49].

Saving student progress to a local text file offers a practical compromise, balancing data security with responsiveness. Each submission writes the student score a master text file, and the number of incorrect attempts to a secondary file. While students could theoretically manipulate these files, this risk can be mitigated by hiding or encrypting them using the subprocesses [50] and cryptography [51] modules. In practice, hiding the text file has proven sufficient to prevent any student tampering.

4 Student experience

This new open-source smart worksheet design was used to support an introductory electrochemistry experiment for first-year undergraduate students in their second semester. These students had limited prior experience with electrochemistry theory or practical

Table 1 Evaluation of data storage methods for the smart worksheet

Data storage method	Secure to student tampering?	Speed of submission / s (correct answer)	Speed of submission / s (incorrect answer)	Reloadable?
Store within the code	✗	0.002 ± 0.001	0.002 ± 0.0004	✗
Save to text file	✓	0.25 ± 0.01	0.11 ± 0.01	✓
Export to GitHub	✓	5.18 ± 0.14	2.95 ± 0.13	✓

Three options compared are storing data within the code, exporting to a text file, or exporting to github. Secure to student tampering indicates that a student would not be able to manipulate the score from the smart worksheet. Speed of submission is the time delay between pressing submit and the correct score being calculated, as recorded on a Dell latitude 5440 laptop (13th gen Intel® core™ i5-1345U 1.6 ghz processor, 16 GB RAM). Values given are the mean average of five replicates along with an error of one standard deviation. Reloadable indicates if student progress is retained if the worksheet is closed and then reopened. A red cross (✗) indicates the condition is not met, a green tick (✓) means the condition is met. A yellow tick (✓) means the condition is met, although there are some caveats that should be considered. This evaluation of the trade-offs between speed, security, and reloadability informed the technical design of the open-source smart worksheet, supporting the broader goal of creating a responsive and user-friendly educational tool

techniques within higher education. Prior to the practical session, students accessed the laboratory script (available in the supplementary information), which provided an overview of theory and the instructions for the practical session. They also watched a brief pre-lab video, which introduced the experimental technique, cyclic voltammetry, and its applications. Notably, the video did not include a walkthrough of the calculations featured in the smart worksheet, preserving the worksheet's role in guiding independent analysis.

During the session, students worked in pairs to complete the experimental component and then independently used the smart worksheet to complete their analysis and assessment. All students had prior experience with a commercial smart worksheet used in a first-semester titration practical, providing a useful point of comparison. Students were encouraged to engage with the worksheet's feedback mechanisms and work through the assessment autonomously.

After completing the session, students were given a short paper questionnaire (see supplementary information) to evaluate their experience with the practical and the smart worksheet. All questions were rated on a 1–10 Likert scale, with 10 indicating the most positive response (Fig. 5). Feedback on the practical component was collected but not analysed to help students to conceptually separate their experience of the practical component versus their experience using the smart worksheet.

The study aimed to determine if an open-source smart worksheet could replicate the positive student perceptions of learning environment and outcomes of commercial alternatives. Questions for student feedback were chosen to investigate student perceptions of positive features that were previously reported for commercial smart worksheets [29–31]. We interpret the study aim as supported when the majority of students provide a strong positive response (Likert score ≥ 7) to relevant questions (Fig. 5A).

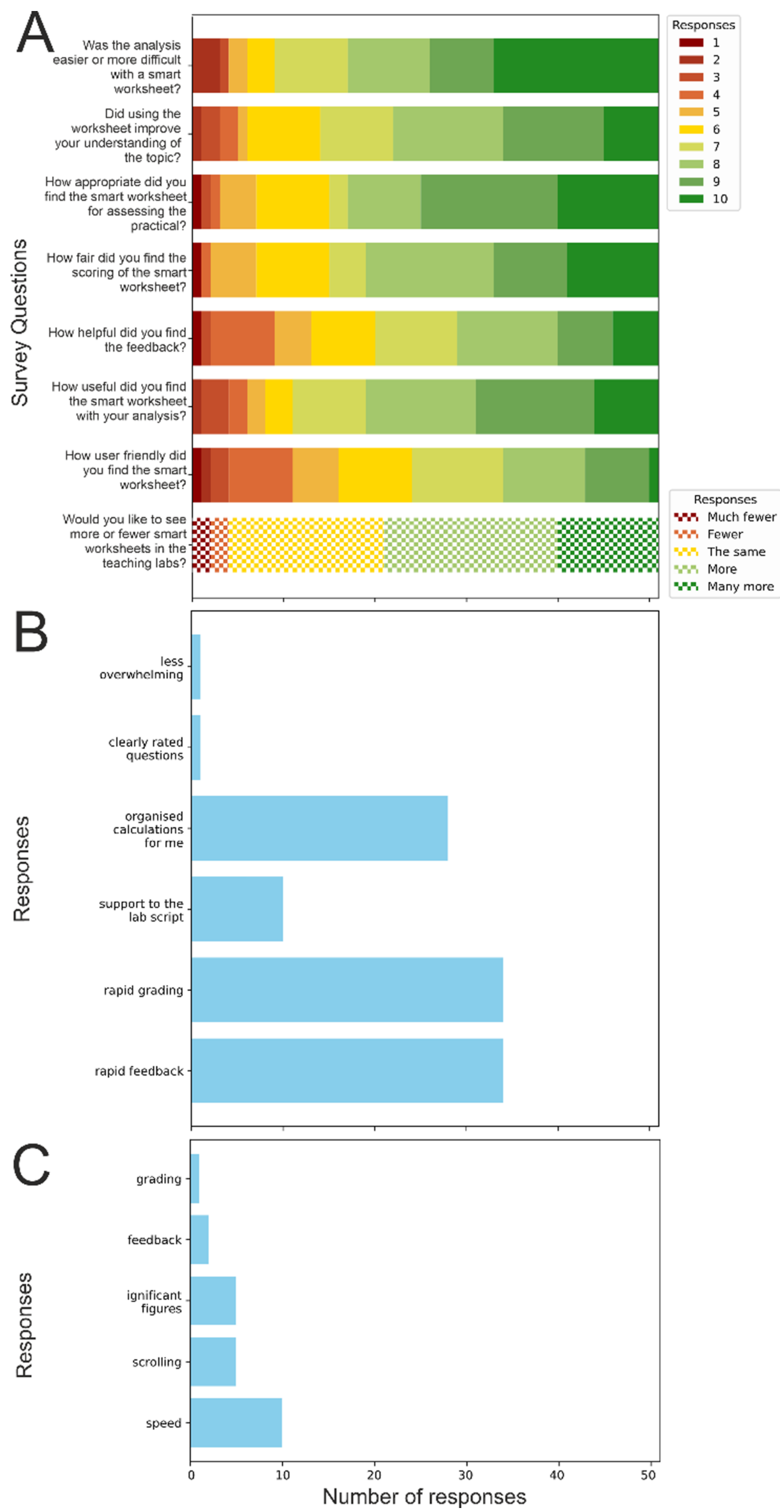

Fig. 5 (See legend on next page.)

Student feedback was generally positive. Most students reported that they felt the worksheet made their analysis easier and improved their understanding of the topic. 78% (40/51) of students gave a strong positive response for how useful they found the smart worksheet for their analysis, 82% (42/51) of students reported that they found it

(See figure on previous page.)

Fig. 5 Student responses to their experience using the open-source smart worksheet to support an electrochemistry practical, $n=51$. Panel A shows Likert-scale ratings of the worksheet's effectiveness in supporting learning and feedback. Panels B and C summarise student preferences and criticisms. These data provide evidence for the worksheet's impact on student experience in support of the study's aim of evaluating its pedagogical value. **A** Responses to fixed questions asking students to rate how the smart worksheet performed in terms of supporting learning and providing feedback. Responses were on a 1–10 Likert scale, where 10 indicates the most positive responses (strongly agree, very helpful, very useful, etc.). Colour bar is selected so that green indicates a strong positive response (7–10), yellow indicates an intermediate response (5–6) and red indicates a poor response (1–4). **B** Student responses to what features they especially liked about the smart worksheet. Checkboxes were provided for the "rapid feedback", "rapid grading", "organised calculations for me", "support to the lab script" and "nothing at all", as well as space for a free text response. **C** Student responses to what features they especially did not like about the smart worksheet. These were based on free text responses, so were thematically grouped based on the speed of loading, scrolling through the worksheet, use of significant figures in entering answers, quality of feedback provided, and the grading provided

made the analysis easier, and 73% (37/51) indicated they felt it improved their understanding of the topic.

These outcomes suggest that the worksheet not only supports technical skill development but also reduces cognitive load, a key barrier in practical STEM education [52]. Open-ended responses further highlighted appreciation for the rapid grading and feedback mechanisms (Fig. 5B). This aligns with positive perceptions that were reported for commercial smart worksheets, which further supports the study aim being met.

From a pedagogical perspective, smart worksheets support constructivist learning, where students actively build knowledge through interaction with feedback and iterative problem-solving [24]. The worksheet's design encourages self-regulated learning, allowing students to reflect on errors, revise responses, and engage with tailored feedback at their own pace [53]. The results from this study are consistent with prior research on commercial smart worksheets, which have been shown to enhance engagement, problem-solving skills, and academic performance in STEM disciplines [29–31]. Importantly, students perceive that this open-source worksheet replicates these benefits while offering greater flexibility and accessibility due to its open-source nature.

Students also had the opportunity to identify features they disliked. Thematic analysis revealed that the most common criticism concerned the slow loading speed of the worksheet (Fig. 5C). This issue was traced to the packaging of the worksheet as a standalone executable file using PyInstaller, which triggered antivirus scans on university-managed laptops. These scans are part of institutional IT policies and are not unique to this smart worksheet design, but they nonetheless impacted the user experience.

This limitation can be negated by running the software on a virtual machine, which removes the requirement for a protective virus scan and allows for immediate launch. Another solution could be to run the worksheet directly as a Python code within Anaconda or a similar Python package manager, although this would allow students to interact with the code directly, which is undesirable. It may also be possible to mitigate this delay by producing a web-hosted version using frameworks such as Flask [54]. However, this would require a sizeable rewrite of the Python codebase, whereas running the executable file on a virtual machine is a simple and immediate fix.

Another mentioned issue was difficulty scrolling, likely due to touchpad incompatibility requiring use of the scrollbar. These usability concerns highlight areas where commercial smart worksheets offer an advantage, since they tend to be hosted within a web browser, which students are already comfortable navigating. A single line instruction to use the scroll bar in place of the track pad can mitigate any student confusion

Table 2 Likert scale responses to questionnaires asking students to rate their experience using an open-source smart worksheet and a commercial smart worksheet

	Open-source smart worksheet ($n=51$)		**Commercial smart worksheet ($n=49$)**	
Question	Median	Interquartile range	Median	Interquartile range
How user friendly did you find the smart worksheet?	7	3	8	2
How useful did you find the smart worksheet in conducting your analysis?	8	2	8	3
How helpful did you find the feedback provided by the smart worksheet?	7	2.5	7	2
How fair did you find the scoring of the smart worksheet?	8	3	7	3
How appropriate did you think the smart worksheet is for assessing this practical?	9	3	8	2
Did using the smart worksheet with the experiment improve your understanding of the topic?	8	3	8	4
Was the analysis easier or more difficult with the smart worksheet compared to how you would normally carry out your analysis?	8	3	8	3
Would you like to see more or fewer smart worksheets in the teaching labs?	More		More	

Questionnaires for both worksheets were shared to 55 first year undergraduate students. The open-source worksheet was shared along with a complex electrochemistry practical to students in the second semester of their first undergraduate year. The commercial smart worksheet was used to support a simple Titration experiment as students first experience in an undergraduate teaching laboratory, in the first semester of their first year. No formal statistical comparison has been done due to significant difference in student experience and assessment complexity between the two worksheets. The number of responses to the questionnaire (n) is given for separately for each worksheet

on this point. Encouragingly, very few students expressed dissatisfaction with feedback and grading. This suggests that the open-source worksheet provides a similarly positive learning experience to commercial systems.

Students who participated in this study had been previously asked the same questions after using a commercial smart worksheet for a different practical (see supplementary information). Direct comparison between the two datasets is limited by differences in experimental complexity and student experience. The commercial worksheet had been employed to support students through their first experience in the teaching laboratory at the beginning of their first undergraduate year, and so student experience and assessment complexity were both significantly lower relative to this electrochemistry session. However, an exploratory comparison between datasets shows a relatively small variance in median Likert scores (Table 2). Although we do not claim that the open-source worksheet directly competes with commercial alternatives, the similarity in student responses provides confidence that educators can adopt open-source solutions to achieve a comparable student experience at significantly reduced cost.

4.1 Limitations in the study

It is important to mention that this study is subject to some limitations. The possibility of desirability bias cannot be ruled out, as students may have provided favourable responses due to social or contextual factors. Efforts were made to minimise the impact of desirability bias. Students completed the questionnaire individually and anonymously to minimise perceived evaluation pressure or alignment with peer norms. Students were also not told that the worksheet was created by their instructor to remove any desire to please the instructor in their responses. Despite best efforts, the possibility of desirability bias cannot be ruled out and should be considered when interpreted the results.

Additionally, all findings are based on self-reported data collected through Likert-scale ratings and open-ended questionnaire responses. While these data provide valuable insights into student perceptions, they do not offer objective evidence of learning gains. As such, claims regarding improved understanding or analytical ability should be interpreted as perceived outcomes rather than measured achievement. Educators looking to incorporate an open-source smart worksheet into their own teaching may wish to consider incorporating objective performance metrics, such as pre/post testing or comparison with control groups, to more rigorously evaluate the impact of open-source smart worksheets on their students' learning. This is planned for future developments of the open-source worksheet design.

4.2 Further scope

While this study focuses on the use of an open-source smart worksheet to support STEM learning, its potential extends far beyond these fields. Any subject involving structured question-answer formats, data analysis or conceptual feedback could benefit from the smart worksheet model. Moreover, any analytical task that can be conducted in Python can be adapted for assessment using this open-source smart worksheet structure.

Smart worksheets can break large-scale calculations into more manageable steps, coupling with targeted feedback to provide a scaffolded environment for self-directed learning. This gives these worksheets a broad reach, such as supporting epidemiology students through disease spread modelling, environmental science students through population ecology or psychology students through cognitive text analysis.

The variety of answer formats (e.g. free text, numerical input, checkboxes) make this open-source smart worksheet adaptable to support learning outside of STEM fields. Free text entry could be used in language courses to provide feedback on grammar and syntax. Numerical entry could be used in history courses to define chronology of events. Checkboxes could be applied to literature courses to identify poetic devices or narrative structures. Graphing tools could be used to support business and economics assignments through profit/loss or supply/demand datasets. These features demonstrate that this open-source smart worksheet is not limited to STEM education but can serve as a flexible, cross-disciplinary tool for enhancing student engagement, feedback, and assessment.

5 Future developments in automated feedback and assessment

As digital learning environments continue to evolve, the integration of automated assessment and feedback systems is well-placed to play a growing role in education. Smart worksheets exemplify how emerging technologies can support scalable, personalised learning experiences. These tools offer real-time feedback, adaptive assessment, and dynamic content delivery, aligning with broader pedagogical goals of inclusivity, engagement, and self-regulated learning.

Open-source software options are particularly desirable. By removing licensing barriers and enabling customisation, open-source tools promote equitable access to high-quality educational resources. This directly supports United Nations Sustainable Development Goal 4, which aims to ensure inclusive and equitable quality education

[55]. In under-resourced institutions or regions, open-source smart worksheets offer a cost-effective alternative to proprietary platforms, empowering educators to tailor content to local curricula and learner needs.

Looking ahead, the integration of artificial intelligence (AI) with smart worksheet technologies presents exciting opportunities. In the design phase, educators can engage with AI to aide in question content and assessment design [56]. AI also has the potential to enhance the functionality of the worksheet itself. AI-enhanced automated feedback could analyse the student responses to identify patterns of misunderstanding, generate personalised hints, and adapt question difficulty in real time [57]. This feedback could also support students in prompt design itself, a skill increasingly relevant in AI-assisted learning environments [58]. Educators could also use a custom built large language model (LLM) trained on carefully selected course materials to generate feedback, ensuring that responses are accurate and course relevant [59].

Smart worksheets also have the capacity to act as a bridge between traditional assessment and AI-enhanced learning. Their structured format and modular design make them ideal platforms for embedding AI agents that provide context-aware support, simulate peer collaboration, or offer multilingual feedback. This hybrid approach supports diverse educational contexts, from STEM laboratories to humanities classrooms, and fosters a more inclusive and responsive digital pedagogy.

The future of automated assessment lies in the convergence of open-source innovation, AI integration, and pedagogical adaptability. Tools like this open-source worksheet demonstrate how educators can harness these technologies to create accessible, engaging, and future-ready learning environments. As institutions seek to adapt to the changing landscape of education, embracing open-source and AI-enhanced solutions will be key to delivering meaningful, equitable, and scalable learning experiences.

6 Conclusions

This study highlights the versatility of a new open-source smart worksheet designed to support active learning and formative assessment. Built using Python, its design offers a user-friendly interface, real-time feedback, dynamic content updates, and a range of customisable features that collectively enhance its value as an educational tool. By integrating features such as rapid grading, targeted feedback, error detection, and graphical analysis, the worksheet is well-suited to support students through complex scientific calculations and data interpretation. The ability to provide tailored, step-by-step feedback based on student inputs ensures that common misconceptions are addressed promptly, fostering deeper understanding and promoting self-regulated learning.

Student feedback from the study indicates a high level of satisfaction, with the majority of students reporting improved understanding and a streamlined analysis process. Specifically, 82% (42/51) of students agreed that they found the smart worksheet made the analysis easier, while 73% (37/51) reported that they felt using the worksheet improved their understanding of the topic. These results reflect strong positive responses and align with prior findings on the effectiveness of smart worksheets in promoting conceptual clarity and analytical confidence. This supports the viability of an open-source smart worksheet as a cost-effective solution for educational institutions seeking to enhance digital learning without incurring licensing fees.

Importantly, the flexibility of the Python architecture allows for straightforward adaptation for use in a variety of academic disciplines. The core features of question-answer formats, dynamic feedback, data handling, and graphical analysis, can be applied to any subject requiring analytical thinking or conceptual reinforcement. Its open-source nature offers customization, scalability, and widespread access, without the constraints of proprietary systems. These features support formative assessment, self-directed learning, and inclusive pedagogy, making this open-source smart worksheet design suitable for diverse educational contexts. Moreover, by automating grading and feedback, the smart worksheet reduces the administrative burden on educators, enabling more time for student interaction and instructional design.

This work finds that students perceive the open-source smart worksheet as providing a similarly beneficial learning experience as a commercial equivalent. In this way, this open-source smart worksheet has the potential to become a valuable educational tool to support learning, provide real-time feedback, foster engagement and reduce staff workload.

Abbreviations

OSPREY	Open-Source Python Responsive Evaluation sYstem
STEM	Science technology engineering and mathematics
GUI	Graphical User Interface
AI	Artificial Intelligence
LLM	Large language model

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1007/s44217-025-00835-6.

Supplementary Material 1

Acknowledgements
The author would like to thank Jake Lewis, Kat Swindells-Ridley and Jess Silvester, the outstanding technician team at our University, for testing the early releases of OSPREY and without whom our teaching laboratory program would not be such a success.

Author contributions
This work has a single author, who was responsible for all aspects of this work.

Funding
The authors have no relevant financial or non-financial interests to disclose. No funds, grants or other support was received.

Data availability
The datasets supporting the conclusions of this article are available in their original comma separated value file format from [https://doi.org/10.5281/zenodo.15225943](https:/doi.org/10.5281/zenodo.15225943). The source Python code for the OSPREY smart worksheet is available through the following source: Project name: OSPREY Project homepage: [https://github.com/Perry-SC/OSPREY_echem](https:/github.com/Perry-SC/OSPREY_echem) Archived version: Version 1.2.0 available online at [https://doi.org/10.5281/zenodo.15125391](https:/doi.org/10.5281/zenodo.15125391) Operating system: Platform independent Programming language: Python Other requirements: Minimum requirement Python version 3, with (at least) the following modules: tkinter, numpy, matplotlib, statsmodels.api, os, csv, subprocess. License: GPL v3.

Declarations

Ethics approval and consent to participate
This study involving human participants was performed in accordance with the 1964 Declaration of Helsinki and its later amendments. Approval was granted by the Ethics Committee of University of Southampton (Approval Number: ERGO-FEPS-101545). Informed consent was obtained from all individual participants included in the study.

Consent to publish
The participants have consented to the submission of the article to the journal.

Competing interests
The authors declare no competing interests.

Received: 16 May 2025 / Accepted: 4 September 2025
Published online: 02 October 2025

References

1. Bretz SL. Evidence for the importance of laboratory courses. J Chem Educ. 2019;96(2):193–5. https://doi.org/10.1021/acs.jchemed.8b00874.
2. Bain K, Rodriguez J-MG, Towns MH. Chemistry and mathematics: research and frameworks to explore student reasoning. J Chem Educ. 2019;96(10):2086–96. https://doi.org/10.1021/acs.jchemed.9b00523.
3. Ejiwale JA. Barriers to successful implementation of STEM education. J Educ Learn (EduLearn). 2013;7(2):63–74.
4. Batty L, Reilly K. Understanding barriers to participation within undergraduate STEM laboratories: towards development of an inclusive curriculum. J Biol Educ. 2023;57(5):1147–69. https://doi.org/10.1080/00219266.2021.2012227.
5. Cech EA. The intersectional privilege of white able-bodied heterosexual men in STEM. Sci Adv. 2022;8(24):eabo1558. https://doi.org/10.1126/sciadv.abo1558.
6. Hung CT, Wu SE, Chen YH, Soong CY, Chiang CP, Wang WM. The evaluation of synchronous and asynchronous online learning: student experience, learning outcomes, and cognitive load. BMC Med Educ. 2024;24(1):326. https://doi.org/10.1186/s12909-024-05311-7.
7. Jensen J, Smith CM, Bowers R, Kaloi M, Ogden TH, Parry KA, et al. Asynchronous online instruction leads to learning gaps when compared to a flipped classroom. J Sci Educ Technol. 2022;31(6):718–29. https://doi.org/10.1007/s10956-022-09988-7.
8. Ramadhan FA, Sumarni W, editors. Development of interactive E-Worksheet based on STEM-Integrated Problem-Based learning to improve student's Problem-Solving Abilities2025 2025-04-01T23:00:00.000Z: Atlantis.
9. Kurniawan LC, Wahyuni I. Development of e-worksheet based on engineering design process for composition function material. Res Dev Educ (RaDEn). 2024;4(1):162–75.
10. Eriana S, Harini E, Kusumaningrum B, Sulistyowati F, Kuncoro KS, Irfan M et al. Effectiveness of e-worksheet on mathematical problem-solving ability based on student learning motivation. AIP Conference Proceedings. 2024;3046(1). https://doi.org/10.1063/5.0194630
11. Fadila A, Dasari R, Setiyaningsih, Septiana R, Sari RM, Rosyid A. The Development of Electronic Flash Worksheet Based on Adobe Flash Cs6 on Fraction Numbers in the Seventh Grade of Junior High School. Journal of Physics: Conference Series. 2019;1155(1):012019. https://doi.org/10.1088/1742-6596/1155/1/012019
12. Mahanan MS, Talib CA, Ibrahim NH. Online formative assessment in higher STEM education; a systematic literature review. Asian J Assess Teach Learn. 2021;11(1):47–62.
13. Carless D. Feedback for student learning in higher education. In: Tierney RJ, Rizvi F, Ercikan K, editors. International encyclopedia of education (Fourth Edition). Oxford: Elsevier; 2023. pp. 623–9.
14. Park LE, Ward DE, Moore-Russo D, Rickard B, Vessels V, Hundley J. Positive feedback as a lever to boost students' STEM outcomes. Pers Soc Psychol Bull.0(0):01461672241265954. https://doi.org/10.1177/01461672241265954
15. Williams A. Delivering effective student feedback in higher education: an evaluation of the challenges and best practice. Int J Res Educ Sci. 2024;10:473–501. https://doi.org/10.46328/ijres.3404.
16. He Y, Olumayegun O, Zandi M. Investigating chemical engineering students' perceptions of feedback: a glimpse into current problems and a platform for improvement. Educ Chem Eng. 2024;46:54–61. https://doi.org/10.1016/j.ece.2023.12.001.
17. Wisniewski B, Zierer K, Hattie J. The power of feedback revisited: a Meta-Analysis of educational feedback research. Front Psych. 2020;10–2019. https://doi.org/10.3389/fpsyg.2019.03087.
18. Mapplebeck A, Dunlop L. Student learning in STEM: feedback practices. In: Peters MA, editor. Encyclopedia of teacher education. Singapore: Springer Nature Singapore; 2022. pp. 1673–9.
19. Hattie J, Timperley H. The power of feedback. Rev Educ Res. 2007;77(1):81–112. https://doi.org/10.3102/003465430298487.
20. Beaumont C, O'Doherty M, Shannon L. Reconceptualising assessment feedback: a key to improving student learning? Stud High Educ. 2011;36(6):671–87. https://doi.org/10.1080/03075071003731135.
21. Steinert S, Krupp L, Avila KE, Janssen AS, Ruf V, Dzsotjan D, et al. Lessons learned from designing an open-source automated feedback system for STEM education. Educ Inf Technol. 2025;30(5):6019–60. https://doi.org/10.1007/s10639-024-13025-y.
22. Coyte E, Lowry RB. Paired online laboratory assessments: formative engagement and summative attainment. J Chem Educ. 2024;101(8):3514–21. https://doi.org/10.1021/acs.jchemed.4c00308.
23. Harmer NJ, Hill AM. J Chem Educ. 2021;98(12):4094–100. https://doi.org/10.1021/acs.jchemed.1c00853. Unique Data Sets and Bespoke Laboratory Videos: Teaching and Assessing of Experimental Methods and Data Analysis in a Pandemic.
24. Williams NA. Using an inclusive curriculum framework to address an awarding gap in a First-Year chemistry module. J Chem Educ. 2022;99(1):171–6. https://doi.org/10.1021/acs.jchemed.1c00362.
25. Chuang S. The applications of constructivist learning theory and social learning theory on adult continuous development. Perform Improv. 2021;60(3):6–14. https://doi.org/10.1002/pfi.21963.
26. Aktiv. Aktiv Learning: Aktiv Chemistry—Teaching College Chemistry [Internet]. 2019 [cited 2025 March 25]. https://aktiv.com/
27. LearnSci. Smart Worksheets - Dynamic Scientific Worksheets [Internet]. 2018 [cited 2025 March 25]. https://www.learnsci.com/products/smart-worksheets
28. Techquipment. Smart Worksheets [Internet]. 2010 [cited 2025 March 25]. https://www.tecquipment.com/engineering-science/smart-worksheets
29. Shallcross D, Davies-Coleman M, Lloyd C, Heslop B, Eastman J, Baldwin T, et al. Smart worksheets and their positive impact on a Second-Year chemical kinetics course. J Chem Educ. 2024;101(4):1735–40. https://doi.org/10.1021/acs.jchemed.3c00934.
30. Shallcross D, Davies-Coleman M, Lloyd C, Dennis F, McCarthy-Torrens A, Heslop B, et al. Smart worksheets and their positive impact on a First-Year quantitative chemistry course. J Chem Educ. 2025;102(3):1062–70. https://doi.org/10.1021/acs.jchemed.4c00705.

31. McCartney J, Egieyeh S, Ebrahim N, Braaf E, Beukes D. Innovations in pharmacy education during the pandemic. SA Pharma J. 2021;88(4):31–3. https://doi.org/10.10520/ejc-mp_sapj_v88_n4_a9.
32. Logothetis TA. The times of COVID-19 and beyond: how laboratory teaching evolved through the pandemic. Chem Teacher Int. 2024;6(1):59–64. https://doi.org/10.1515/cti-2022-0044.
33. Hill AM, Harmer NJ. A murder mystery gamification session to consolidate analytical biochemical techniques learning. J Chem Educ. 2023;100(11):4514–24. https://doi.org/10.1021/acs.jchemed.3c00279.
34. Weiss CJ. A creative commons textbook for teaching scientific computing to chemistry students with python and Jupyter notebooks. J Chem Educ. 2021;98(2):489–94. https://doi.org/10.1021/acs.jchemed.0c01071.
35. Major L, Francis GA, Tsapali M. The effectiveness of technology-supported personalised learning in low- and middle-income countries: A meta-analysis. Br J Educ Technol. 2021;52(5):1935–64. https://doi.org/10.1111/bjet.13116.
36. Haleem A, Javaid M, Qadri MA, Suman R. Understanding the role of digital technologies in education: a review. Sustain Oper Comp. 2022;3:275–85. https://doi.org/10.1016/j.susoc.2022.05.004.
37. Taber KS. The use of cronbach's alpha when developing and reporting research instruments in science education. Res Sci Educ. 2018;48(6):1273–96. https://doi.org/10.1007/s11165-016-9602-2.
38. Python docs. tkinter — Python interface to Tcl/Tk 2014 [25th March 2025]. https://docs.python.org/3/library/tkinter.html
39. Freeman S, Eddy SL, McDonough M, Smith MK, Okoroafor N, Jordt H, et al. Active learning increases student performance in science, engineering, and mathematics. Proc Nat Acad Sci. 2014;111(23):8410–5. https://doi.org/10.1073/pnas.1319030111.
40. PyInstaller. PyInstaller Manual 2022 [25th March 2025]. https://pyinstaller.org/en/stable/
41. Zimmerman BJ. Becoming a Self-Regulated learner: an overview. Theory Pract. 2002;41(2):64–70. https://doi.org/10.1207/s15430421tip4102_2.
42. Sweller J. CHAPTER TWO - Cognitive load theory. In: Mestre JP, Ross BH, editors. Psychology of learning and motivation. Volume 55. Academic; 2011. pp. 37–76.
43. Matplotlib. Matplotlib 3.10.1 documentation 2024 [25th March 2025]. https://matplotlib.org/
44. Pletcher D. Techniques for the study of electrode reactions. A first course in electrode processes. Royal Society of Chemistry; 2009. pp. 154–221.
45. Chen M-Y, Tang JT. Developing a digital game for excel skills learning in higher education - a comparative study analyzing differences in learning between digital games and textbook learning. Educ Inf Technol. 2023;28(4):4143–72. https://doi.org/10.1007/s10639-022-11335-7.
46. PyGithub. Introduction - PyGithub 2014 [25th March 2025]. Available from: https://pygithub.readthedocs.io/en/latest/introduction.html
47. Lee J-C, Xiong LN. Investigation of the relationships among educational application (APP) quality, computer anxiety and student engagement. Online Inf Rev. 2021;46(1):182–203. https://doi.org/10.1108/oir-08-2020-0348.
48. James TL, Zhang J, Li H, Ziegelmayer JL, Villacis-Calderon ED. The moderating effect of technology overload on the ability of online learning to Meet students' basic psychological needs. Inform Technol People. 2021;35(4):1364–82. https://doi.org/10.1108/itp-03-2021-0225.
49. Valverde-Berrocoso J, Acevedo-Borrega J, Cerezo-Pizarro M. Educational technology and student performance: A systematic review. Front Educ. 2022;7–2022. https://doi.org/10.3389/feduc.2022.916502.
50. Python docs. Subprocess management — Python 3.13.2 documentation 2023 [25th March 2025]. https://docs.python.org/3/library/subprocess.html
51. PyPI. Cryptography 38.0.1 2016 [25th March 2025]. https://pypi.org/project/cryptography/
52. Rathnayaka CM, Ganapathi J, Kickbusch S, Dawes L, Brown R. Preparative pre-laboratory online resources for effectively managing cognitive load of engineering students. Eur J Eng Educ. 2024;49(1):113–38. https://doi.org/10.1080/03043797.2023.2204813.
53. Khalil M, Wong J, Wasson B, Paas F. Adaptive support for self-regulated learning in digital learning environments. Br J Educ Technol. 2024;55(4):1281–9. https://doi.org/10.1111/bjet.13479.
54. Gaspar D, Stouffer J. Mastering Flask Web Development: Build enterprise-grade, scalable Python web applications, 2nd Edition: Packt Publishing; 2018.
55. United Nations. Sustainable Development Goals: 4 Quality Education 2022 [27th August 2025]. https://www.un.org/sustainabledevelopment/education/
56. Fernández AA, López-Torres M, Fernández JJ, Vázquez-García D. ChatGPT as an instructor's assistant for generating and scoring exams. J Chem Educ. 2024;101(9):3780–8. https://doi.org/10.1021/acs.jchemed.4c00231.
57. Erümit AK, Sarıalioğlu RÖ. Artificial intelligence in science and chemistry education: a systematic review. Discover Educ. 2025;4(1):178. https://doi.org/10.1007/s44217-025-00622-3.
58. Araújo JL, Saúde I. Can ChatGPT enhance chemistry laboratory teaching?? Using prompt engineering to enable AI in generating laboratory activities. J Chem Educ. 2024;101(5):1858–64. https://doi.org/10.1021/acs.jchemed.3c00745.
59. Maughan AE, Toberer ES, Zevalkink A. Integrating large Language models into the chemistry and materials science laboratory curricula. Chem Mater. 2025;37(7):2389–94. https://doi.org/10.1021/acs.chemmater.5c00111.

Publisher's note