Chemistry Education Research and Practice

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal

Cite this: DOI: 10.1039/d2rp00213b

Evaluative judgement – a practitioner's case in chemistry research projects†

Anna Bertram *[a] and Carmen Tomas [b]

Engaging students actively in assessment, using a range of formative activities, consistently over sustained periods is a common recommendation in the assessment for learning literature. Despite this, practice still lags behind. Our case study aims to bridge the widening gap between theory and practice by illustrating the application of recent theoretical concepts in practice. The literature contains many examples and research on isolated events to engage students in assessment. Cases that explore engagement in different formative practices, over extended periods of time, are scarce and challenging for practitioners to implement in the absence of examples and evidence. Consequently, whilst adoption of theory informed practices remains limited, research also remains limited. Our case study aims to bridge the widening gap between practice and theory by elaborating a case example for practitioners. The redesign of a third year laboratory module, in which students undertake research projects, is presented. Our case illustrates how practitioners can incorporate assessment for learning principles, outlined in contemporary frameworks (evaluative judgement) considering process and a learning sequence over an entire year. The learning design of a module, before and after, is fully described to exemplify how practitioners can implement theoretical principles in practice. Students' perception of the value of the new tasks were gathered to inform reflections for practitioners in the implementation of evaluative judgement.

Received 18th July 2022,
Accepted 28th October 2022

DOI: 10.1039/d2rp00213b

rsc.li/cerp

Introduction

Designing assessment for learning: theory and practice

The challenges to assessment practice are well documented. Heavy summative assessment load, lack of or poor alignment to learning outcomes and transparency of criteria, and the low engagement of students in assessment and feedback are well rehearsed in the literature (Evans, 2013; Winstone and Nash, 2016; Jessop and Tomas, 2017; Tomas and Jessop, 2019). Contemporary literature abounds with proposed theoretical frameworks providing a direction. Student engagement in assessment needs to be sustained over periods to have an impact on student learning. This proposition presents challenges for practitioners.

Assessment should be used strategically in the learning process. The literature lends support to the statement that '*students see assessment as the curriculum*' (Gibbs, 1999) and that whether a task is assessed or not shapes students' perceived value of the task. There is considerable discussion in the literature about the value of formative alongside summative assessment in the curriculum and how these can be successfully balanced to optimise learning. The proportion of formative assessment in relation to summative assessment remains low (Jessop and Tomas, 2017; Wu and Jessop, 2018). This is perhaps due to the evidence in the literature that students will focus most effort on what 'counts' (Harland *et al.*, 2015) and therefore course designers are reluctant to change assessment practices. For example, many frameworks propose more time is allocated to involve students actively in understanding quality but it is, at times, hard to see how these activities may be prioritised over other important practices or activities. Reducing summative assessment, increasing formative, is very challenging in practice.

Bridging the gap between theoretical models and practice requires careful consideration. Good practice in assessment is available but evolves slowly. The changes required, in line with models that we explore in sections below, demand a transformation that goes beyond single interventions. Practitioners often lack the confidence that many of the ideas proposed will be effective and deliver positive outcomes. Creating assessment for learning designs that incorporate multiple interventions and over a period of time poses multiple challenges for practitioners. There is a paucity of models or case examples situated

[a] School of Chemistry, University of Nottingham, University Park, Nottingham, UK. E-mail: anna.bertram@nottingham.ac.uk
[b] Honorary Assistant Professor, School of Education, University of Nottingham, Nottingham, UK. E-mail: carmen.tomas@cantab.net

† Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d2rp00213b

in real contexts for practitioners. Illustrating how theory can be implemented in practice is the first step.

Secondly, institutional cultures still do not promote the reduction of summative and the increase of formative assessments, over sustained periods. Recent research evidences that dominant discourses and culture do not facilitate these changes for individual practitioners (Winstone *et al.*, 2021). Whilst theoretical proposals suggest increasingly sophisticated ideas, institutional cultures do not promote the practices. The gap between practice and theory is widening. Consequently, whilst practice remains so far behind theoretical proposals, building up an evidence base at the scale and scope required, remains compromised.

Our study, in a third-year practical chemistry module, explores how to develop assessment for learning designs in practice. Designing learning and student engagement, over an entire year, is the primary focus of our study. Literature with case examples that illustrate the application of theory into practice is limited. Consequently, the effectiveness of some of the principles, widely recognised as central, is yet to be evidenced. Our case study aims to advance the field by proposing a theory-driven redesign of an existing module in practice. An example, situated in a real context, will show practitioners how they can design learning in practice taking account of the proposals in most recently formulated theoretical frameworks.

Below, we review the existing frameworks for assessment for learning, the evidence and recent reformulations of assessment for learning. In this context, we present the aims and relevance of our study.

Student engagement in assessment and self-regulation: theories and evidence

Engaging students actively is a fundamental principle for good practice in all contemporary assessment frameworks (Price *et al.*, 2012; Evans, 2013; Winstone and Nash, 2016; Boud *et al.*, 2018). Instructional environments play a fundamental part. Programme design and engaging students in learning are key for effective design of learning and assessment (O'Donovan *et al.*, 2007; Evans, 2013; Winstone *et al.*, 2015, 2017; Jessop and Tomas, 2017; Tomas and Jessop, 2019). Theoretical proposals, despite using different terms, all converge on the centrality of developing self-regulated learning as the underpinning theory of assessment for learning in practice.

Self-regulated learning defines learning as students' understanding of their own abilities and themselves (metacognition). This has replaced former thinking about learning based on constructs such as intelligence or individual differences (Zimmerman and Kitsantas, 2002). Self-regulation is central to success in life and achievement in academic settings alike (Dignath *et al.*, 2008; Zimmerman, 2008; Sitzmann and Ely, 2011; Clark, 2012).

Peer, self and co-assessment are methods of engaging students that can promote autonomy and self-regulation. Peer learning is defined by Boud (1991) as the "*use of teaching and learning strategies in which students learn with and from each other without the immediate intervention of a teacher*." Peer and self-assessment are very closely linked; Boud (1991) described self-assessment as "*the involvement of students in identifying standards and/or criteria to apply to their work and making judgements about the extent to which they have met these criteria and standards*." This was later reiterated by O'Donovan *et al.* (2007) who said that "*by involving students in the marking process, they can expand the assessment of learning into an effective learning tool and generate a technique of assessment for learning. Peer and self-assessment techniques thus enable a student to become more autonomous in their learning and help to develop a student's ability to identify their own learning needs*."

Evidence is available on various strategies that address aspects of self-regulation. For example, the formative use of rubrics to engage students in self and peer assessment promotes the development of self-evaluation and reflection skills (Dochy *et al.*, 1999; Jönsson and Svingby, 2007; Reddy and Andrade, 2010; Panadero and Jönsson, 2013; Brookhart, 2018). These are essential skills for self-regulation, autonomy and ultimately life-long learning (Struyven et al., 2005; Carless, 2007; Boud *et al.*, 2018). Comprehensive reviews of the literature (Winstone *et al.*, 2015, 2017; Evans, 2013) also reinforce this essential aspect of instructional environments and how students' active participation could be promoted. An essential element of active participation is to develop the students' ability to understand requirements and how their performance aligns with the criteria for assessment.

In chemistry education, the same principles have also been applied and similar types of activities implemented, for example peer review of reports (Jones and Seybold, 2016), peer review and digital badges for laboratory skills development (Seery *et al.*, 2017), engaging students in creating assessment questions using PeerWise (Ryan, 2013; Galloway and Burns, 2015), self-assessment and cooperative learning in mini-lab activities (Branan and Morgan, 2010).

Formative assessment has been shown to be effective in supporting the development of experimental techniques, note taking, data analysis and report writing in a second year chemistry practical module (Seery *et al.*, 2019). Significantly the formative feedback was integrated throughout the whole laboratory module, this was in contrast to previous models which involved the majority of feedback occurring towards the end of the laboratory sessions.

A common limitation, to both practice and research, is that much of the existing literature and evidence base has been drawn from single interventions. All existing frameworks emphasise the need for a sustained engagement in such activities or deploying different activities over time. Recent revisions of theoretical frameworks, to guide practice, have taken account of the integration of these various techniques, their sequencing and process over a period. Below, we present a recent theoretical proposal reframing self-regulation, practice, and the process of engaging students in assessment.

Bringing it all together: evaluative judgement

Evaluative judgement (Boud et al., 2018) provides a theory-based framework for practice. It emphasises process and longitudinal aspects of self-regulation. Students' evaluative

judgement is the ability to understand quality, to judge self and the work of others and learn from those judgements. This framework is significant for practitioners to integrate multiple known formative practices (*e.g.* rubrics, peer and self-assessment, use of exemplars) with self-regulation in learning. Evaluative judgement, providing a coherent framework for practitioners, promotes actively and explicitly the students' ability to judge their own work and that of others. Similar concepts exist and predate evaluative judgement in the literature (*e.g.* Sadler, 1989; Price et al., 2012). This recent theoretical reformulation advances important notions of sequence of a range of activities and process in practice.

Student self-regulation is an important driver in instructional design. Despite this, instructional designs and assessment practices do not fully embrace the tenets of student self-regulation (Winstone *et al.*, 2021). Evaluative judgement provides a framework to advance our understanding of how instruction can take account of self-regulated learning (Panadero and Broadbent, 2018). Self-regulation is not an individual trait or ability. Self-regulation is self-generated thoughts, feelings, and behaviours that are oriented to attain goals (Zimmerman, 2000; Zimmerman and Kitsantas, 2002).

The components of self-regulation are cognition, metacognition and motivation. Self-regulation involves self-awareness of one's own abilities, the ability to self-assess in relation to a task in order to self-correct and, finally, motivation to engage. Motivation is rooted in the belief and perceptions about one's own abilities (self-efficacy), (Bandura, 1982). Self-efficacy is a predictor of motivation. The beliefs one holds about one's own abilities can determine the time spent on a task and intrinsic motivation (Zimmerman and Kitsantas, 1997) which are, again, associated with success in tasks.

Self-regulation is not an ability but a process. Four levels of self-regulation exist (observation, emulation, self-control and self-regulation) (Zimmerman and Kitsantas, 2002). These levels relate to each other and progression between levels is non-linear. Panadero and Broadbent (2018) propose that various assessment-related interventions would be appropriate at different levels and stages of this process. The proposal remains theoretical but considers instruction and learning, drawing attention to the process elements. This proposal is summarised below (Panadero and Broadbent, 2018):

SRL level 1 observation. Students observe an expert performing the task. The emphasis is on understanding what is required. At this level, examples of relevant instructional strategies include use of assessment rubrics to engage students in generating and understanding criteria.

SRL level 2 emulation. Students perform the task following an example. Self, peer or co-assessment and practice observation of qualities found in exemplars and models.

SRL level 3 self-control. Students practice in the absence of a model and gain greater self-control. Self-assessment may become more automatic and understanding of standards would be expanded beyond the exemplars.

SRL level 4 self-regulation. Students can attempt similar and then unseen problems in the absence of experts or model answers.

Rationale for the study: theory into practice

The primary aim of our study is to help bridge the gap between theory and practice. Evaluative judgement integrates and reframes pre-existing theoretical frameworks (*e.g.* Sadler, 1989; Price et al., 2012). Evidence has also grown, on the effectiveness of many techniques to engage students in assessment (*e.g.* use of rubrics, peer and self-assessment), albeit as single events. Readers can consult literature reviews elsewhere on the evidence to date (Evans, 2013; Panadero and Jönsson, 2013; Winstone and Nash, 2016).

Despite theory and evidence advancements, adoption in practice lags behind. Recent research reminds us that culture and dominant notions in practice pose barriers to the adoption of the proposed concepts (Adachi *et al.*, 2018; Winstone *et al.*, 2021). Evaluative judgement, despite its advantages as a coherent framework, poses challenges to practitioners. Whilst theoretical developments become more sophisticated, these outpace the developments in practice.

Our study explores a recently formulated theoretical framework, and its implementation in practice. The aim is to provide an important direction and example for practitioners. The model proposed above involves integrating, in practice and over a period of time, a range of support mechanisms to foster evaluative judgement (Boud *et al.*, 2018). This adds questions in practice concerning design, timing, frequency, and sequencing of these activities. The framework requires practitioners to deploy multiple strategies (rubrics, self and peer assessment, reflection), most of which are new to many practitioners. Adding a consideration of process, presents new challenges to most practitioners to begin to visualise what this really means in practice. Moreover, making time for these additional activities requires reducing summative assessment which is another well-known challenge in practice.

Taking into consideration the challenges in practice, providing example cases and stimulating reflection with concrete examples is an important step when considering the widening gap between theory and practice. Translating theoretical frameworks into practice and real contexts requires cases and examples in the first instance. Enabling and inspiring others to embrace the model and principles is a requisite for growing these practices. Growth of cases and at scale, will pave the way for larger scale research into the impact of the theory described on student learning and self-regulation. Our primary aim is to stimulate the discussion amongst practitioners by providing an example.

We have referred to a growing body of evidence, on the effectiveness of proposed techniques for learning. In addition to impact on learning, previous research also shows that students' perceptions of different modes of engagement also matter for practitioners to understand how best to implement a given technique (Patton, 2012). Understanding students' experience is an important aspect of practice as, in addition to learning gains, the experience of different tasks poses important considerations for practitioners to implement the framework.

Our aim is to support the advancement of practice within the direction set in the aforementioned theoretical framework of evaluative judgement. Our focus, therefore, is the design of

student engagement in assessment. To advance this, as a first step towards exploring the theoretical models in practice, our case first explores the redesign from a traditional approach in chemistry laboratory projects. Secondly, we aim to understand student reactions and experience to the range of techniques deployed to help inform the development and implementation in practice of the initial model. Therefore, our aim is to explore student reactions to and perceptions of the multiple tasks throughout a year of study.

Method: case study

Our aim is to advance the implementation of theory-based models in practice, in line with a recently formulated framework. By illustration, we expect to grow the interest in the chemistry education community of practitioners. The study, placed in a real setting, blurs the boundaries between context and the focus of investigation (Yin, 2002).

Our study focusses on a theory-based redesign and provides an example for the chemistry education community of how to implement a series of challenging ideas in practice. We consult students on their experience of the initial redesign to gain insights and inform our thinking and revisions of the initial proposal. Therefore a flexible framework, allowing a plurality of perspectives and an emphasis on complex aspects (Stake, 2000; Yin, 2002), is required to capture both the redesign and student experience. The case study, whilst mainly framed within qualitative approaches of enquiry, allows for a variety of information sources and types to be collected (*e.g.*, quantitative and qualitative) throughout the process. The rich descriptions that case studies permit are necessary for other practitioners to learn from this case and extrapolate to their contexts.

A case study offers the possibility of using a range of diverse methods to bring to light aspects described above. This case study is exploratory, aiming to provide insights for chemistry education practitioners of how to approach the design of evaluative judgement in practice. The case study is comprised of two distinct parts to provide insights into the process of redesign by the practitioner and the students' experience of the new activities proposed.

- Part 1 provides a rich description of the application of principles from theory in a real practical context.

A module redesign, taking into account the theoretical principles and applying them in practice, is the focus. In line with the model, the stages of preparation, practice and self-regulation were included in the redesign. This section provides a rich description of the practitioner's thinking and reflection on theory in practice.

This rich and descriptive case of our module redesign can make a significant contribution for chemistry education practitioners.

- Part 2 explores students' perceptions of the value of different activities that were included over the course of a year.

The review identified that moving from theory to practice is challenging. One challenge is the additional factors that affect implementation. In addition to the value to learning, students' perceptions and reactions are also key to the success of embedding new practices. In our initial implementation, it was important to gauge students' reaction, their ability to engage with the tasks and the value they perceived the tasks having. The effective use of peer assessment in practice requires understanding students perceptions and reactions to the tasks (Patton, 2012). In our first implementation of the design we wanted to incorporate student feedback as a starting point to consider the effectiveness of exposing students to such a range of diverse strategies (rubrics, peer assessment, self-assessment, reflection sessions) most of which are not part of mainstream practice. Certainly, in our context, students might have occasionally encountered some activities to engage them actively. Understanding students' perceptions of the tasks presented was essential in the first implementation.

We distributed three questionnaires at different points during the course of the year. Each covered a range of activities and stages in line with the proposals of evaluative judgement (see Table 4). The questionnaires captured student perception of the value of the tasks. In order to avoid participant fatigue, three questionnaires were distributed:

- Before preparation workshops (Questionnaire 1, see Appendix 1, ESI†).
- After preparation workshops (Questionnaire 2, see Appendix 2, ESI†).
- At the end of the module (Questionnaire 3, see Appendix 3, ESI†).

Each questionnaire covered a range of the new activities in the module. The focus was on evaluating student perception of the range of tasks experienced throughout the module and as detailed in part 1.

The appendices contain the full questionnaires which were used. This initial case study reports on the questions aimed at capturing the student experience and comments on the activities. Future publications will present fully other exploratory questions aimed to capturing student learning and self-efficacy, which is still under investigation.

The questionnaires included Likert scale type questions to understand the student experience (overall satisfaction) of the sessions. Students' ratings of their perceived value of the activity using a five-point scale (from not valuable to extremely valuable). Additional questions to understand the value of the sessions in students' learning and whether they would take part again.

Students also provided open text comments to gain insight into the reasons for their ratings on value of different tasks.

Ethical considerations

Prior to filling in all questionnaires, students were briefed on the optional nature of taking part in the evaluation. Students provided their informed consent in writing prior to completion of the survey. Support researchers anonymised all questionnaires. This step was taken to ensure that the evaluation would not influence any other assessment activities. The ethical procedures applied in this study were approved by one of the University's ethics committee.

Sample

Out of a potential 169 students taking the module:

- 116 responded to Questionnaire 1 before taking part in the preparatory workshops.
- 89 responded to the Questionnaire 2, after preparatory workshops.
- 22 students completed Questionnaire 3 (end of the module).

The sample size varies depending on each question. The exact number of students answering each question is detailed where appropriate, as sample sizes vary. See Questionnaires 1–3 in the Appendix (ESI†).

Data analysis

The questions used varied in terms of the response formats.

Descriptive numerical summaries are offered in relation to questions where students were asked to declare preferences for future sessions and modes of engagement. Questions where we have quantified answers were either category (yes/no) responses or Likert scales. Percentages, frequency counts and medians or means are used for these.

Likert scale ratings generated a numeric type of response. For students' ratings of the perceived value of different activities in the preparation stage (see Table 4 workshops 1–3: co-construction of criteria and co-assessment of exemplar activities) we obtained a large sample in the first and second questionnaires. This allowed for us to test students' ratings of their perceived value of those initial activities during the preparation stage. A Wilcoxon signed ranks test was used given the data derived from Likert scale ratings in Questionnaire 2 (5 point Likert scale). Given the large sample (89 responses to Questionnaire 2) we report summary statistics using parametric measures (mean and standard deviation).

Evaluation at the end of the module (Questionnaire 3), covered activities: self-assessment, feedback review and action planning. This also included other elements like the rubric used throughout the module and tutor feedback to enable us to get a complete picture of the key elements of engaging students. The lower participation as it was the end of the year, resulted in a lower response rate. We could only match eight questionnaires to the responses provided in the previous questionnaire (2). This limits the analyses and we present summary descriptive statistics since sample size of eight responses would not allow for statistical or significance testing. Also, non-parametric measures are used given the low response rate. Medians are used as they are a more resilient measure of central tendency for a small sample.

The third and last questionnaire, also collected students' justifications for their ratings as open text comments. These open text comments have been quoted directly. The extracts are quoted verbatim to enable readers to interpret the student perception of the various activities. We have used classification of the comments according to positive and negative comments.

Also, we cluster comments in relation to key aspects of self-regulation described in the model: understanding expectations, goal setting, reflection. These are the key concepts referred to in the in the SRL model. Due to the small number of open text comments, we have included the full comments verbatim to allow readers to interpret the comments and views expressed.

Part 1 Description of the practical project module design before and after transformation

The third-year chemistry practical module involves students working on research projects in what is their first experience of working in teams and on open-ended experiments. In first- and second-year laboratory modules, students work independently on practicals which are well defined. The emphasis is on following procedures, developing experimental techniques and reproducing results which are well documented. The third-year practical module was designed in 2014 (Bertram et al., 2014) to bridge the gap between this approach and research. Much of the scaffolding present in the first- and second-year practical course was removed, the experiments were no longer well defined and the emphasis was very much on the development of transferable skills required for independent research in line with the requirements from the Royal Society of Chemistry (2017).

The third-year module was broken down into three five-week projects in which students did one project in each of the three areas: inorganic, organic and physical chemistry. In the first week of the module, a series of short workshops were delivered focussing on some of the key transferable skills students would require to successfully complete the projects (Table 1).

Many of the challenges described in the literature were visible in this module. For example, module evaluations suggested that students did not understand the assessment criteria and had a poor understanding of what was required in order to perform well. Additional feedback from students also highlighted an issue with the demanding timetable for each project. Very limited time to complete the required work and little time to digest feedback from the previous project was a challenge.

The assessment criteria, with descriptors for different attainment levels, were provided as a document on the Virtual Learning Environment (VLE) with the expectation that students would read and understand the document. When work was assessed, a numerical rubric was attached to the assessment, but these numerical marks did not link with the criteria and descriptors. Students were then expected to cross-reference this with the document containing the criteria and descriptors of attainment levels. However, it was clear from student feedback that this was not happening and the level of engagement with the documentation outlining the assessment criteria was low.

The assessment of the third-year practical module incorporated some laboratory-based elements, a final report and a presentation. Encouraging students to engage with learning outcomes and assessment criteria early in the assessment cycle was a consistent challenge. Many attempts had been made to increase the prominence and signposting of the information *via* the VLE, but this lack of engagement remained an issue. A lack of understanding of the assessment criteria is likely to diminish the learning gain and may be responsible for the lack

Table 1 Structure of the project module before transformation; students did three five-week projects

	Activities	Teaching weeks
Preparation	Transferable skills workshops Introductions to: searching the literature, report writing, presentations	2
Project 1	Searching the background literature Preparing a project plan (summative feedback) Practical project work Data analysis & report writing (summative feedback) Assessed presentations (summative feedback)	3 4–6 7
Project 2	Searching the background literature Preparing a project plan (summative feedback) Practical project work Data analysis & report writing (summative feedback) Assessed presentations (summative feedback)	8 9–11 12
Project 3	Searching the background literature Preparing a project plan (summative feedback) Practical project work Data analysis & report writing (summative feedback) Assessed presentations (summative feedback)	19 20–22 23

of engagement with feedback, and it was therefore critical that this issue was addressed.

There was an expectation that students would develop transferable skills, which are key to the success of student project work, throughout the module. However, the experience of running this module suggested that this was challenging as these transferable skills build on skills previously practiced and knowledge previously acquired, thus requiring students to work with a level of autonomy not previously required. In light of the initial observations and feedback after running the project module, we wanted to expand the transferable skills sessions at the start of the academic year and focus on supporting students in learning independently. In other words, developing self-regulation and autonomy as an important basis for project work.

A theory and evidence-based module redesign

In the context of learning for undergraduate chemistry projects, and in the light of the challenges experienced, our case echoed many of the concerns expressed in the literature. Therefore, in line with the suggested principles from the literature, the first objective was to explore the implementation of a number of the ideas to reduce load and increase the emphasis on learning, drawing from the evaluative judgement framework. As such, we focussed on key factors that are discussed in the literature in promoting a culture of learning and engaging students in assessment. In doing so, we wanted to address a number of known issues in our redesign with the aim to gain greater student engagement:

- Reduction of assessment load to allow for more opportunities to engage in learning.
- Enhanced transparency of assessment (*e.g.* use of rubrics in marking and feedback).
- Engaging students in understanding quality, criteria, self-assessing and planning.

Following the evaluative judgement as our theoretical model, as already discussed, we wanted to explore in the preparatory stage of the module how Self-Regulation Levels (SRL) 1 and 2 could be integrated. It is very important for the effective running of projects to enhance students' understanding of what is expected of them (SRL 1 observation) early in the process and foster self-evaluation (linking to SRL 2 Emulation), all of which are linked to motivation (self-efficacy). The steps taken in the redesign are detailed below.

Reduction in summative assessment load to increase student reflection and learning. We embarked on a complete redesign of the third-year laboratory module with an emphasis on student engagement with learning outcomes and assessment criteria from the very beginning. In order to do this, we redesigned the structure and timetabling of the entire module to allow more time to deliver workshops and support the feedback and reflection process. The number of projects students would take was reduced from three to two (one in each semester) to create time for skills development, feedback and reflection activities (Table 2). The rationale for this was that if we could increase the value and impact of each of the two projects, then two projects would be more beneficial to student learning than three projects in the original format.

Enhancing transparency for better student engagement – module learning outcomes and assessment criteria. The communication of learning outcomes, criteria, and descriptors needed improvement as noted in the previous years. A key driver to redesign the module was to increase transparency in order to engage students with module learning outcomes during organised sessions.

We redesigned the module learning outcomes, subsequently broke these down into assessment criteria (Table 3) and created a module assessment rubric with descriptors for four levels of attainment for each assessment criterion (Fig. 1). The literature contains a range of papers on the design of assessment rubrics, and the work of Mertler (2001); Dawson (2017) was very helpful. Overall, we had five overarching module learning outcomes: team working and time management; Safety and good laboratory practice; Technical competence; Knowledge, understanding and

Table 2 Timetable of redesigned module

	Workshop	Activities	Teaching weeks
Preparation	1–3	Transferable skills workshops (see details in Table 4)	2–3
Project 1		Preparing a project plan (formative feedback)	
		Peer assessed presentations on project aims (formative feedback)	4–5
		Practical project work	6–9
		Data analysis & report writing (summative feedback)	10–12
		Assessed presentations (summative feedback)	
		Student self-evaluation against assessment criteria for project	
Feedback and Reflection	4	Feedback reflection, comparison of assessor's feedback with self-assessment.	19–20
	5	Action planning for future development & Project 2.	
		Meeting with assessor (see details in Table 5)	
Project 2		Preparing a project plan (formative feedback)	21–22
		Peer assessed presentations on project aims (formative feedback)	
		Practical project work	23–26
		Data analysis & report writing (summative feedback)	27–29
		Assessed presentations (summative feedback)	
		Student self-evaluation against assessment criteria for project	
Feedback and reflection	6	Feedback reflection, comparison of assessor's feedback with self-assessment	30–31
	7	Action planning for future development	
		Meeting with assessor (see details in Table 5)	

critical thinking; Technical writing and presentation skills. These were then broken down into a total of 17 assessment criteria. To emphasise alignment and enhance transparency of assessment for students, the rubric represented all the criteria, indicated where they would be assessed and the weighting of each. A small, summarised extract of the rubric is shown in Fig. 1. The full assessment rubric has been attached in Appendix 4 (ESI†).

The rubric type chosen was analytic (*i.e.* matrix type of display), as this type of rubric is preferred for learning due to the greater level of detail and feedback they provide than holistic rubrics (Brookhart and Nitko, 2019). It was essential to have a comprehensive assessment rubric like this to enable us to design the previously described activities to engage students with the module learning outcomes. In contrast to previous practice, this same rubric was used for students to self-assess themselves and to communicate levels of performance by the assessor.

Preparing students from the start: workshops designed to engage students in understanding expectations. Self-regulated learning (SRL) level 1

In the first two weeks of the module, we introduced a series of workshops aimed at engaging students in understanding the

Table 3 Module learning outcomes and assessment criteria

Module learning outcomes	Assessment criteria
Team working & time management	Plan & devise experimental work for each laboratory session to make best use of the time available Communication & team working
Safety & good laboratory practice	Identification of hazards, relevant precautions & disposal to ensure safe experimental work Use of Good Chemistry Laboratory Practice (GCLP)
Technical competence	Accurately record experimental procedures & observations (laboratory notes) Quality experimental work
Knowledge, understanding & critical thinking	Breadth of knowledge of the background research area of the project Understanding of the methodologies/techniques/tools relevant to the project (*e.g.* synthetic routes, software, equipment, spectroscopic methods) Ability to describe relevant results
Technical writing & presentation skills	Report is well structured Visual elements (graphs, figures, tables): range of types of visual elements, quality and adequate presentation (labels, captions) Nomenclature, schemes, equations, figures. Style (use of technical vocabulary) Experimental data reported in journal style Degree of consistency of the formatting of the document (font size, references list, use of bold/italic, layout, spacing, *etc.*) Accuracy of the grammar, structure of sentences and division of text into paragraphs Use of in-text citations to support development of argument Presentation of experimental results to a small audience using PowerPoint (or similar)

	Achievement Levels					
	1	2	3	4		
	Adequate	Good	Excellent	Exceptional	Assessment point	Weighting
Assessment Criterion 1	Descriptors of each level provided				Laboratory (notebook & demonstrators)	5%
Assessment Criterion 2	Descriptors of each level provided				Post-lab - Written report	10%
Assessment Criterion 3	Descriptors of each level provided				Post-lab - oral presentation	10%

Fig. 1 Summarised extract from the assessment rubric outlining the information available to students for each criterion.

module learning outcomes and the assessment criteria associated with the individual pieces of work. Promoting this engagement has been proven to enhance students' motivation and learning. These workshops focussed on the first two levels of self-regulation, observation and emulation (Table 4). Our plan below combines a variety of activities to promote different aspects of self-regulated learning (SRL).

Workshop 1 – promoting understanding of learning outcomes and self-assessment. (SRL 1 understanding expectations). This workshop focussed on increasing students' awareness and understanding of the learning outcomes and assessment criteria for the module. Our aim was to support students to understand the rationale for the module design. To do this, we guided them through the process we had gone through ourselves in designing the module, as we proposed that this would improve their understanding of what was required of them to do well.

The workshop would rely on successfully engaging students in this reflective dialogue. As such, we created online boards to which students could anonymously submit contributions. Students were asked to:

- Identify skills already developed in practical modules in earlier years, which they would need for project work.
- Identify skills they still needed to develop to successfully complete project work.
- Propose the module learning outcomes.
- Suggest how the module learning outcomes could be assessed.

Students worked in groups to discuss these questions and then commented using their mobile phones or tablets on the electronic discussion board which received very good engagement. After each activity, students' comments were summarised in a reflective discussion by the facilitator.

Workshops 2 & 3 – students generating assessment criteria and students as assessors: co-construction of criteria and co-assessment (SRL 1 and 2). Our aim of workshop 2 was to enable students to better understand what constituted a good project report and presentation. We proposed that this would in turn lead to a better understanding of what the assessment criteria would be.

In this module, students are required to write a report on their project in the style of a Royal Society of Chemistry journal, the first activity in this session was for students to work in groups to look at exemplar reports. Students were asked to pick out elements which they thought were good and those which needed improvement. Following on from this, students were

Table 4 Summary of workshops 1–3 delivered before project work commences to support understanding expectations (SRL1)

Workshops	Outline	
1	Promoting understanding of learning outcomes and self-assessment.	Setting the context – highlighting the differences between approach in 2nd and 3rd year laboratory modules. Project description given, tasks requiring teams of students to: • Summarise skills required to successfully carry out project work, as those: ○ Already developed. ○ Requiring development. • Generate a list of proposed learning outcomes for the module. • Propose how these learning outcomes could be assessed. At the end of this series of exercises the actual learning outcomes for the module were shared and there was very close correlation.
2	Students generating assessment criteria for project reports and assessing exemplars.	An introduction was given to report writing, including discussion of some exemplars. Students worked in small groups to co-review three different reports. After each report was considered a class discussion took place led by the academic facilitator. A comparison was made between student 3rd year project reports and research papers, with the aim being to ensure students could see where their current learning fitted in. Students were asked to come up with a list of assessment criteria for the report. The assessment criteria for the report were shown, these were discussed and comparisons with those suggested by students made. We found that students were able to propose criteria which matched were very closely aligned.
3	Students generating assessment criteria for presentations	Colleagues delivered two contrasting presentations and students were asked to highlight positive and negative elements. This exercise culminated in students suggesting assessment criteria for the presentations. We found that students were able to propose criteria which matched were very closely aligned.

asked to suggest the assessment criteria that should be used for the reports.

Our rationale for this approach was that students, given the right support, would be able to identify for themselves what constituted a good report. Orsmond *et al.* (2002) describe the benefits of using exemplars in increasing students' understanding of assessment criteria. Interestingly, they also note that peer assessment can be more objective than self-assessment.

Workshop 3 was delivered using the same strategy as workshop 2 but focussing on presentations. Staff presented using two very different styles and students again suggested assessment criteria. Online discussion boards were used to facilitate both of these sessions.

Activities designed to engage students during the project development (on both project 1 and 2)

In line with the literature, our design of student engagement also took into account the consideration that to promote students' self-regulation a variety of tasks needed to be implemented which spanned the whole module. The section below explains the activities planned to address the need for continued engagement of students in learning.

Peer assessment. As already discussed, a major change was redesigning the module to allow more time for all activities (Table 2).

In the newly designed module, students have two weeks to search the literature and write a project plan before commencing practical work. During this period, we introduced peer-assessed presentations. The purpose of the peer-assessed presentations is three-fold: to ensure students have a clear understanding of their project area before commencing project work; to provide an opportunity to practice a presentation which is not formally assessed; to increase understanding of the assessment criteria for the presentations. The peer-assessed presentation gives an opportunity for students to operate in SRL 2 (emulation) following on from earlier sessions on presentations in which exemplars were given (SRL 1 observation).

Feedback and reflection stage – activities designed to engage students in self-assessment, feedback reflection, and action planning (self-regulated learning (SRL) levels 2 & 3)

After the practical component of each project was complete, workshops were delivered to promote reflective engagement with the feedback and action planning for future learning (Tables 2 and 5).

Students were asked to self-assess themselves against the marking criteria used to assess the projects and this self-assessment was then submitted alongside the project report. The aim of this is to encourage students to self-assess (SRL 2 emulation) more routinely and to work towards a situation where they are able to self-assess as a matter of routine and in the absence of exemplars (SRL 3 self-control).

Reviewing feedback (Workshops 4 & 6). After the project reports had been assessed, students attended workshops 4 and 6 (Table 5) designed to facilitate reflection on feedback. Students were asked to identify three positive and three negative elements of feedback and, for each of these, were asked to reflect on the feedback. Students were also asked to compare and contrast their self-assessment with the tutor assessment and reflect on any differences. A proforma was provided to take students through this reflective process, see Appendix 5 (ESI†). These activities culminated in students preparing three questions for discussion with their assessors. The following week a tutorial was timetabled for students to meet with their assessor to discuss the feedback and their questions (Table 5).

Action planning (workshops 5 & 7). Following on from the meeting with the assessor, there was an additional workshop (5 & 7) (Table 5) to facilitate students writing a development plan ahead of future project work. Students were provided with a template, see Appendix 6 (ESI†), to take them through this process, during which they are working towards developing skills needed for SRL 3 (self-control) & SRL4 (self-regulation).

Table 5 Summary of workshops 4–7 delivered after the project work to support engagement in feedback, reflection and action planning (SRL 2 & 3)

Workshops	Outline	
4 & 6	Feedback and reflection	These workshops are delivered after each practical project and are designed to encourage students to engage with the feedback provided by the assessors. The aims are to: • Facilitate student's understanding of what they're doing well in their project work & areas they could improve in. • To help students develop their evaluation skills to enable them to better judge the quality of their work. This workshop is focussed around the use of the proforma in Appendix 5 (ESI).
5 & 7	Action plan for future learning	These workshops are designed to follow on from the workshop in the previous week. Having reflected on their feedback and considered the significant feedback they've received, students are guided through a process of writing an action plan for their future development. This process involves students considering the following: • What can you do to build on the positive feedback you've received? • How can you develop your work in the areas where you received some criticisms? Students are also asked to develop their own evaluation by considering the following: • What can you add to those above? • What do you want to keep doing in the future? • What do you want to change or improve in the future? This workshop is focussed around the use of the proforma in Appendix 6 (ESI).

Part 2 student experience of the evaluative judgement activities throughout the year

The objective of the evaluation in this exploratory stage was to gain insights into the students' perceptions of the various activities described in part one. As described in the methods section, and to avoid evaluation fatigue, three questionnaires were distributed to students. Whilst we did not evaluate every single activity, we collected students' reactions to the various elements introduced in our redesign.

Increasing student engagement in the preparation stage. Prior to the workshops (Questionnaire 1), participants were asked if they had read the assessment rubric. Out of the 115 respondents, 21 (18 per cent) had read the rubric, 26 (23 per cent) had not read the rubric and 68 (59 per cent) were unaware it was available.

Students' perceptions of the value of activities in the preparation stage: co-assessment of exemplars and discussions. During the workshops to prepare students for the year, two different types of co-assessment activities took place: co-construction of criteria and co-assessment of exemplars. In the post-workshop questionnaire (Questionnaire 2), students were asked to rate both activities. Both activities were rated on a five-point scale, from not valuable (1) to very valuable (5) (Table 6). Just under 80 per cent of students said that the co-assessment of exemplars was valuable or very valuable. There was more of a spread of opinion on the co-construction activities aimed at increasing their awareness and understanding of module learning outcomes and assessment criteria.

A Wilcoxon signed rank test was performed to explore whether students rated the two activities differently. Out of the total 89 responses, only 87 were complete so that we could match and compare responses on the two activities. The results show that there was a significant difference in participants' ratings of the two activities ($W = -3.766$, $n = 87$, $p < 0.001$). The co-construction of criteria having a lower rating ($M = 3.46$, SD = 1.06) than the co-assessment of exemplars activity ($M = 4.00$, SD = 0.90). The scale was a five point Likert scale (1 not valuable, 5 extremely valuable). This difference is associated with a medium effect size, $r = 0.285$. In terms of student preferences, 35 participants rated the activities the same, 12 ascribed higher value to the co-construction of criteria and 40 thought the co-assessment exemplar activity was more valuable (Table 6).

An additional breakdown of students' ratings is provided in Fig. 2 below. It shows the frequency of students' ratings of the value of the task engaging them in co-constructing criteria.

Fig. 3 below provides a graph showing the frequency counts for different ratings by 87 students of the co-assessment of exemplars. This activity involved students assessing a range of example reports in class and discussing assessment criteria.

Table 6 Students' perceived value of co-construction of criteria and co-assessment activities in preparation stage

Activity	N	Mean	SD
Co-construction of criteria	87	3.46	1.06
Exemplar co-assessment	87	4.00	0.90

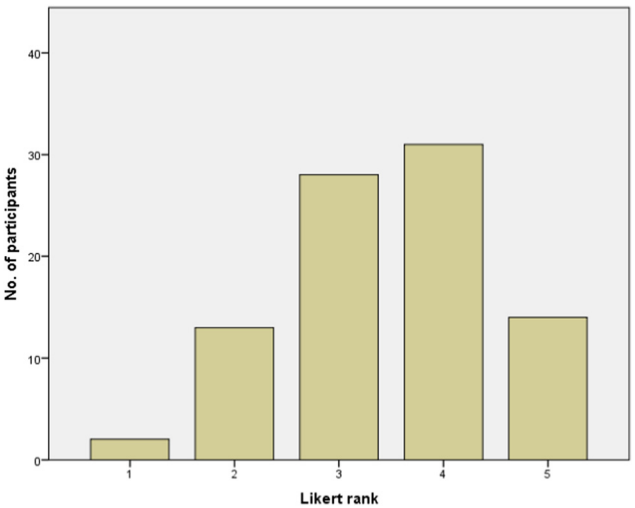

Fig. 2 Frequency count of students' ratings of the value of the co-construction of criteria task (1 not valuable – 5 extremely valuable).

Lastly, in addition to understanding differences, we wanted to capture the student experience of the full set of workshops designed to engage them in this very early stage, overall. When asked *'would you like to take part in future sessions similar to these ones?'*, out of 87 respondents, 79 (90 per cent) indicated they would take part in future similar sessions.

Student engagement after summative submissions to review performance (self-assessment), feedback and action planning. At the end of the module, students were asked to rate the value of the activities aimed to engage them with their own learning. These activities included the self-assessment, the tutor assessment, feedback reflection, action planning and the usefulness of the assessment rubric (Table 7). A total of 22 students provided their ratings to these different aspects of engagement that had taken place over the second part of the module. Detail in Table 7 shows the number of ratings per activity type. Given the low response rate, in

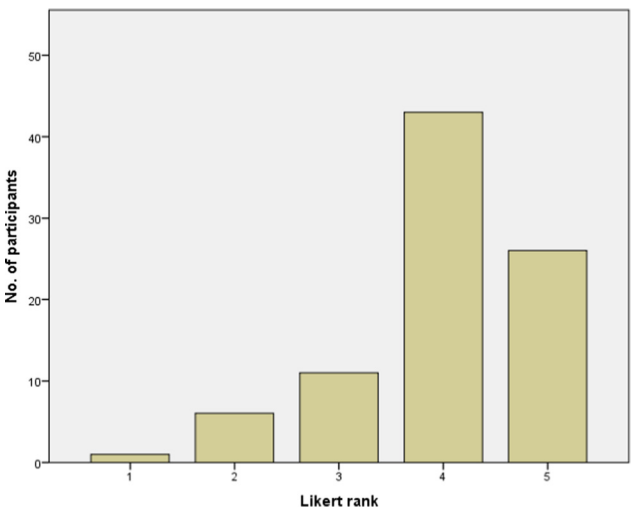

Fig. 3 Frequency count of students' ratings of the value of the co-assessment of exemplars task (1 not valuable – 5 extremely valuable).

Table 7 Students' median ratings of the value of different activities

Activity	Median	N
Self-assessment	3	21
Tutor assessment	4	15
Reviewing feedback	4	18
Assessment rubric	4	20
Action planning	3	5

this final questionnaire, we report the results using non-parametric descriptive statistics. Statistical or significance testing was not deemed appropriate given the small sample and missing data.

The ratings above indicate that students perceived some activities to be more valuable than others. Students were also asked to justify their ratings. A total of twelve students (of 22) provided justifications for their ratings which are summarised below.

Assessment rubric. Pivotal to the success of the module redesign was the development of a comprehensive assessment rubric that linked module learning outcomes and criteria. When asked what they thought of the rubric, a range of comments were made in relation to improved ability to understand expectations which supports SRL 1. A total of seven comments were provided by students justifying the ratings. The complete set of comments is included showing positive impact:

"helpful to have a breakdown of the marks"

"excellent guideline"

"excellent for seeing what is required of me"

Other comments showed enhanced ability to identify actions to improve, which is a higher level of self-regulation:

"clear to find how to improve"

"good to know where to improve"

Lastly, important insights about how understanding rubrics is not always straightforward were made:

"useful but sometimes difficult to interpret"

"vague an non-precise categories"

Self-assessment activity (during projects). Students self-assessed themselves against the assessment criteria for the project and submitted this with the coursework submission. When asked to rate the value of this activity, the median of the responses was 3 (out of 5) indicating students were ambivalent about its value. Students' ratings were justified by comments that indicate that this activity, performed individually, was perhaps challenging, in the main, whilst some students expressed a positive experience. A total of twelve (of 22) students provided a justification.

Self-assessment posed challenges for students. We quote justifications that provide insights into how the task was not a valuable experience:

"I would not submit a report I didn't think was excellent, therefore self-evaluation seemed a little pointless,"

"everyone wants to give themselves high marks"

"I feel the rating itself did not benefit me much"

Two of the less positive comments referred to the rubric and limited understanding of standards as the reason for the limited value of this task: *"categories vague"; "difficult to judge own work. Lacking understanding of what a good report consists of"*

Other comments related to the timing of the self-assessment being submitted prior to receiving feedback.

"do not feel it is useful before feedback is given"

However, on a number of cases, students referred to important thinking processes like reflection and clarifying expectations that were positively impacted.

"good to think about my role in the project"

"the self-assessment helped me to understand the requirements of the lab report"

"I would have never thought back about the progress"

"good to really take in improvements required"

"helpful to plan where marks were coming from"

"useful to know where marks are obtained and distributed"

Tutor assessment (feedback). Students valued the tutor assessment highly (4 out of 5). Seven of the 22 responses contained a written justification. Students still see the tutor (assessor) as the expert. This is illustrated in comments such as:

"getting ratings from a professional is always good"

"they know the project in great detail and can add really detailed comments"

"good for specific questions"

Other positive comments were generic (*"very useful"*, *"good to see others' opinions of the work"*)

Students do not value their own evaluations or peer assessment as highly.

The remaining comments on tutor feedback also revealed common problems of the transparency and fairness of written feedback and marking:

"inconsistent and harsh"

"some marks are not clearly justified so it would be more helpful to know why these marks were awarded"

Reviewing feedback workshops 4 & 6. Reviewing feedback was rated overall positively by the students (4 out of 5). Students' comments reveal that they engaged in reflection of self and standards (SRL 2 practicing, and SRL 3 gaining greater self-control). The full set of comments provided by students are quoted verbatim for the reader to interpret. There were a total of eleven written comments justifying the ratings. Whilst some of the comments are broadly positive, some include references to reflection (analyse, reflect, compare):

"helpful to analyse where I went wrong"

"important to understand positive/negative aspects of work"

"give time to reflect and break down feedback"

"very informative"

"good to compare own opinion with assessors"

"have time to reflect on performance and can discuss with people on how to improve"

"good for actually making me self-reflect"

"being told to learn what we've done bad and accept it"

"useful to know where marks were lots, and how to improve"

One comment referred to situations where the activity may not be valuable:

"only positive feedback meant not much to add or improve on"

One comment (of the 11) was not positive: *"less useful"*

A minority of students did not really understand what the benefit of the feedback review was. Comments were mostly positive.

Action planning

These workshops were delivered to allow students to action plan ahead of either their next third year project (project 2) or their fourth-year project. When asked about the perceived value of this workshop the median response was 3, the same as for the self-assessment. Students' comments show further insight and positively support that higher levels of self-regulation were enabled. Two participants provided an explanation:

"helped with making a plan of what to do on the second project"

"good to think about how to implement self-evaluation"

Discussion of findings

Our case study aimed to advance contemporary theoretical proposals to drive assessment for learning in practice. There is considerable evidence in the literature on the challenges and barriers to promoting and implementing the principles of assessment for learning (Winstone *et al.*, 2021). Contemporary theoretical frameworks seek to guide instructors on how to support learning. The evaluative judgement framework (Boud *et al.*, 2018) draws on an extensive evidence base and links practice with models of self-regulated learning. This recent framework emphasises the longitudinal aspects of the design of assessment for learning and proposes a process to frame practice linking to theories of self-regulated learning.

There are many examples in the literature on how various elements (*e.g.*, use of rubrics, peer/co/self-assessment, action planning) in isolation, as single interventions for students, impact on learning (Panadero and Jönsson, 2013; Winstone et al., 2017; Brookhart, 2018). Supporting learning over a sustained period, by engaging students in the known elements of self-regulation (understanding the task, expectations, practice and own performance), is less well evidenced given the challenges for practitioners (Winstone *et al.*, 2021) to implement learning design. The aim of our case study was, first, to illustrate how we can inform our practice in line with the evaluative judgement framework. An evaluation of student perceptions of the value of a range of tasks, over the course of a whole year, provides an exploratory insight into what students may value most or perceive more accessible in the context of a year-long exposure to different ways of being engaged in their learning as part of assessment.

In this context, our case contributes to chemistry higher education, but also to other disciplines, in two ways. Our redesign case may inspire module leaders to consider greater alignment of criteria and engagement of students in learning. More importantly, our redesign case illustrates how contemporary theories of student autonomy and self-regulation can be articulated in our instructional design over a sustained period, and this could extend to any discipline. Our project module redesign demonstrates how to address key challenges in the literature regarding assessment over-load and student engagement. The redesign has involved reducing the number of summative assessment points, increasing the number of formative activities with the aim of familiarising students with criteria and descriptors, training them as markers by marking and discussing quality of exemplars in the early stage of the module.

Sustained engagement

The engagement of students was sustained by following up with further peer assessment activities, self-assessment and reflection. All these activities built on each other in line with the timing of summative assessments. This substantial contribution illustrates how, as instructors, we can transform assessment *of* learning to be *for* learning in a variety of ways and in a sustained manner. In order to enable both theory and practice to evolve, cases such as this are necessary. The full set of self-regulation levels has been considered in a year-long design of student learning and engagement. Cases in the literature, such as this one, are scarce and necessary to bridge the gap between theory and practice.

Reduction in assessment load

The reduction in the number of projects, and, in turn, the assessment load (summative), enabled us to design this module with the overarching aim of increasing student engagement and autonomy. With the extra time we freed up from reducing the number of projects, we were able to provide more workshops to support learning throughout the year.

Regarding the student experience, based on their views we gathered, the initial results, overall, are positive and encouraging. These initial findings also present variable levels of student acceptance and perceived value which are an important element to reflect on for future implementations.

Workshops as a tool to engage students

Approximately two thirds of the group had not engaged with the rubrics prior to the preparation workshops in the year. This seemed to reflect what we had found in previous years when, despite all the assessment information being available on the virtual learning environment, few students engaged with it before completing the assessed work. Our workshops aimed at increasing awareness and engagement with learning to change this behaviour. This echoes messages from the literature that there is a need to build in-class activities to engage students in understanding criteria and, on a more elementary level, to highlight the availability of an important guide for learning. The rubric was one of the aspects valued positively by students.

The evaluative judgement framework proposes that a range of activities is implemented, as illustrated in our theoretical discussion. Models, such as the one by Panadero and Broadbent (2018), include a whole range of activities (*e.g.*, self-assessment, reflection, evaluation of the work of others). Students' perceived value was variable for different elements. Students attributed a greater value to the rubric and activities that took place at the preparation stage. In particular, assessing the quality of exemplars, using the rubric and then discussing this with the tutor. Co-construction of criteria was perceived to have a lesser value in comparison with the active engagement in the co-assessment of

exemplars. Ninety per cent of students declaring they would like similar sessions to run in the future.

Tutor feedback, and reflection on tutor feedback in the second part of the year, were valued positively by the students. Self-assessment and action planning activities, which took place in the second half of the year, were the least valuable in comparison to the whole suite of activities according to students' ratings. Many factors may have influenced these outcomes. Students completed self-assessment and action planning individually in a workshop setting. In this initial iteration, we have not yet included opportunities for discussion, debate or input from an expert or in a peer group.

Our initial evaluation suggests that self-assessment and action planning are activities that we need to reconsider. In the broader context, these activities were encountered only once in the programme of study. These new, and more demanding, activities may have played a part in how students were able to appreciate their value. Self-assessment and reflection are advanced skills in self-regulation and considering the level of mastery of other aspects (*e.g.*, understanding quality, SRL level 1) may influence students' ability to perform these higher-level skills in an explicit manner. This is an important consideration for practice and future research.

Students seemed to rate activities in the preparation stage of the year more highly (*e.g.*, reviewing report exemplars). Some of the other activities like self-assessing, identifying skills and synthesising module learning outcomes were less highly valued. Theoretical models suggest that these activities are all part of promoting self-regulation and consequently, part of a whole. Yet, in practice, the initial feedback provided by third-year students on their experience, reveals that not all the activities are easily accessible.

In our future implementations we will seek to find better ways of implementing self-assessment and action planning. Also, to promote familiarity integrating into a programme-level assessment strategy across all years of the degree would also be important. It is hardly surprising that when applied in a standalone third year module students are not easily able to change their approaches and views on assessment. Introducing more self-assessment activities throughout the degree course may support students to understand their value and be able to honestly self-assess themselves, and in turn develop skills of self-regulation. Theoretical proposals for practice (Panadero and Broadbent, 2018) may incorporate these considerations in their models to continue to support and guide practitioners.

Self-assessment and action planning are challenging to students. We made an assumption that students would be able to self-assess and action plan as routine. In our experience, students need guidance and support. We know from the evaluative judgement framework that self-assessment and action planning require students to operate at SRL levels 3 and 4. The 'Developing Engagement with Feedback Toolkit' (DEFT) framework (Winstone and Nash, 2016) suggests that students can do these activities on their own. In our experience, however, we have gained insight into the higher level of demand for students and the need to support students in this process. Self-assessment and planning actions might require either scaffolding with an expert and, possibly, a much greater level of preparation of students than we had anticipated. We have taken on board these initial insights and are introducing changes to these activities to incorporate peer support in the future.

Limitations

This is an initial exploratory study illustrating a design of learning. Our case makes an important contribution illustrating how practitioners can apply the principles of evaluative judgement into practice. Our initial review focussed on capturing students' reactions to the broad range of activities proposed in the literature.

This initial study has focussed on how students perceived the value of a range of tasks during a year. The study provides initial insights about tasks perceived to be more valuable. Evaluations of self-regulation are still necessary to provide additional insights and continue developing the theoretical models and practice guidelines. Our study has drawn primarily from student self-reported perceptions of value. Additional methods should be used to evaluate and verify the perceived value and the impact on learning.

In addition, larger samples are also needed. Our exploratory study shows some differences in how students valued different tasks. Due to limitations of sample size (Questionnaire 3), we have not been able to conduct statistical tests on the full set of student ratings. However, descriptive statistics suggest that this line of enquiry would be worth pursuing further.

Future publications will provide additional insights drawn from parts of our evaluation looking at the impact on student learning.

Conclusion and future work

Our case study provides a model for practitioners and direction to design courses which focus on developing students' autonomy and evaluative judgement. These are essential skills and attributes for chemistry graduates. Our case provides an important illustration of how to design student learning.

Reducing the number of projects and, in turn, the amount of summative assessment is quite a bold move in an environment where the culture is to set regular summative assessments. As already discussed, moving towards increasing the number of formative activities is still seen as high risk. A common experience in practice is that students will focus their efforts on activities that 'count.' We need to change student perceptions about higher education, assessment, feedback and their role as partners with assessors in the process. Our redesign case study can provide inspiration for practitioners by helping to think of an alternative concept of how to design and support assessment in practice.

In this initial exploration, student perceptions and experiences, show a higher value attributed to the activities in the earlier part of the year. These activities aimed at increasing their understanding of what was required of them to do well.

Students commented on the benefit of having a comprehensive assessment rubric and on using exemplar reports in co-assessment exercises. These were relatively small interventions with a high value for students. Practitioners may be encouraged to introduce these types of developments which are of high impact for students.

Students' perceived value of the self-assessment and action planning activities was lower than we expected. In future implementations we will consider how best to prepare students for these tasks.

To conclude, we recommend the introduction of activities aimed to develop students' evaluative judgement skills and their understanding of assessment criteria, primarily for student learning and secondly for assessment transparency. Decreasing the summative assessment load may seem challenging, for all the reasons highlighted earlier, but is necessary to create time to introduce more valuable activities aimed to:

- Enhance the transparency of assessment.
- Engage students in understanding quality and criteria.
- Develop students' self-assessment skills and action planning for future development.

Overall, our case study demonstrates the design of a chemistry module to promote student learning and autonomy. The theoretical proposals are challenging for practitioners, since implementing each of these activities requires preparation for both staff and students. In particular, staff have to learn about each of the proposed techniques and adapt them to their context. Our findings provide an indication for practitioners about which types of activities they could work on introducing first. We have evidence of activities that may be of greater value and more accessible to students.

Colleagues might prioritise the elements that are more accessible and valuable to students. Rubrics and co-assessment of exemplars are good starting points. Introducing these right at the start we also recommend was effective. Reflecting on tutor feedback was another valuable activity for students. Other elements recommended in the literature (*e.g.* self-assessment and action planning) require more preparation and practice for students to understand and obtain the desired value from them.

Conflicts of interest

There are no conflicts to declare.

References

Adachi A., Tai J. and Dawson P., (2018), Academics' perceptions of the benefits and challenges of self and peer assessment in higher education, *Assess. Eval. High. Educ.*, **43**(2), 294–306.

Bandura A., (1982), Self-efficacy mechanism in human agency, *Am. Psychol.*, **37**, 122–147.

Bertram A., Davies E. S., Denton R., Fray M. J., Galloway K. W., George M. G., Reid K. L., Thomas N. R. and Wright R. R., (2014), From cook to chef: Facilitating the transition from recipe-driven to open-ended research-based undergraduate chemistry lab activities, *New Dir. Teach. Phys. Sci.*, **10**, 26–31.

Boud D., (1991), *Implementing student self-assessment, HERDSA Green Guide*, 2nd edn, Sydney: Higher Education Research and Development Society of Australasia.

Boud, D., Ajjawi R., Dawson P. and Tai J., (2018), *Developing evaluative judgement in Higher Education: Assessment for knowing and producing quality work*, London: Routledge.

Branan D. and Morgan M., (2010), Mini-Lab Activities: Inquiry-Based Lab Activities for Formative Assessment, *J. Chem. Educ.*, **87** (1), 69–72.

Brookhart S. M., (2018), Appropriate criteria: Key to effective rubrics. *Front. Education: Assessment, Testing and Applied Measurement*, **3**, 22.

Brookhart S. M. and Nitko A. J., (2019), *Educational Assessment of Students*, 8th edn Boston, MA: Pearson Education.

Carless D., (2007), Learning-oriented assessment: conceptual bases and practical implications, *Innov. Educ. Teach. Int.*, **44**, 57–66.

Clark I., (2012), Formative assessment: assessment is for self-regulated learning, *Educ. Psychol. Rev.*, **24**, 205–249.

Dawson P., (2017), Assessment rubrics: towards clearer and more replicable design, research and practice, *Assess. Eval. High. Educ.*, **42**, 347–360.

Dignath C., Buettner G. and Langfeldt H. P., (2008), How can primary school students learn self-regulated learning strategies most effectively? a meta-analysis on self-regulation training programmes, *Educ. Res. Rev.*, **3**, 101–129.

Dochy F., Segers M. and Sluijsmans D., (1999), The use of self-, peer and co-assessment in higher education: a review, *Studies High. Educ.*, **24**, 331–350.

Evans C., (2013), Making sense of assessment feedback in higher education, *Rev. Educ. Res.*, **83**, 70–120.

Galloway K. W. and Burn S., (2015), Doing it for themselves: students creating a high-quality peer-learning environment, *Chem. Educ. Res. Pract.*, **16**, 82–92.

Gibbs G., (1999), Using assessment strategically to change the way students learn, In: Brown S. and Glasner A. (ed.) *Assessment Matters in Higher Education*. Buckingham: Society for Research into Higher Education and Open University Press.

Harland T., McLean A., Wass R. Miller E. and Sim K. N., (2015), An assessment arms race and its fallout: high-stakes grading and the case for slow scholarship, *Assess. Eval. High. Educ.*, **40** (4), 528–541.

Jessop T. and Tomas C., (2017), The implications of programme assessment patterns for student learning, *Assess. Eval. High. Educ.*, **42** (6), 990–999.

Jönsson A. and Svingby G., (2007), The use of scoring rubrics: reliability, validity and educational consequences, *Educ. Res. Rev.*, **22**, 130–144.

Jones M. L. B. and Seybold P. G., (2016), Combining chemical information literacy, communication skills, career preparation, ethics, and peer review in a team-taught chemistry course, *J. Chem. Educ.*, **93**, 439–443.

Mertler C. A., (2001), Designing scoring rubrics for your classroom, *Pract. Assess., Res. Eval.*, **7**, 1–10.

O'Donovan B., Price M. and Rust C., (2007), Know what I mean? Enhancing student understanding of assessment standards and criteria, *Teach. High. Educ.*, **9**, 325–335.

Orsmond P., Merry S. and Reiling K., (2002), The Use of Exemplars and Formative Feedback when Using Student Derived Marking Criteria in Peer and Self-assessment, *Assess. Eval. High. Educ.*, **27**, 309–323.

Panadero E. and Jönsson A., (2013), The use of scoring rubrics for formative assessment purposes revisited: a review, *Educ. Res. Rev.*, **9**, 129–144.

Panadero E. and Broadbent J., (2018), Developing evaluative judgement: a self-regulated learning perspective, in D. Boud, R. Ajjawi, P. Dawson and J. Tai (ed.), *Developing Evaluative Judgement: Assessment for Knowing and Producing Quality Work*, Abingdon: Routledge.

Patton, C., (2012) 'Some kind of weird, evil experiment': student perceptions of peer assessment, *Assess. Eval. High. Educ.*, **37** (6), 719–731.

Price M., Rust C., O'Donovan B., Handley K. and Bryant R., (2012), *Assessment Literacy: The Foundation For Improving Student Learning*, Oxford: Oxford Centre for Staff and Learning Development.

Reddy Y. M. and Andrade H., (2010), A review of rubric use in higher education, *Assess. Eval. High. Educ.*, **35**, 435–488.

Royal Society of Chemistry, (2017), Accreditation of Degree Programmes.

Ryan B. J., (2013), Line up, line up: using technology to align and enhance peer learning and assessment in a student-centred foundation organic chemistry module. *Chem. Educ. Res. Pract.*, **14**, 229–238.

Sadler D. R., (1989), Formative assessment and the design of instructional systems, *Instructional Sci.*, **18**, 119–144.

Seery M. K., Agustian H. Y., Doidge E. D., Kucharski M. M., O'Connor H. M. and Price A., (2017), Developing laboratory skills by incorporating peer-review and digital badges, *Chem. Educ. Res. Pract.*, **18**, 403–419.

Seery M. K., Jones A. B., Kew W. and Mein T., (2019), Unfinished Recipes: Structuring Upper-Division Laboratory Work To Scaffold Experimental Design Skills, *J. Chem. Educ.*, **96**, 53–59.

Sitzmann T. and Ely K., (2011), A Meta-Analysis of Self-Regulated Learning in Work-Related Training and Educational Attainment: What We Know and Where We Need to Go, *Psychological Bulletin*, **137**, 421–442.

Stake R., (2000), Case Studies, in Denzin, N. K. and Lincoln, Y. S. (ed.) *Handbook of Qualitative Research*, 2nd edn, London: Sage.

Struyven K., Dochy F. and Janssens S., (2005), Students' perceptions about evaluation and assessment in higher education: a review, *Assess. Eval. High. Educ.*, **30**, 325–341.

Tomas C. and Jessop T., (2019), Struggling and juggling: a comparison of student assessment loads across research and teaching-intensive universities, *Assess. Eval. High. Educ.*, **44** (1), 1–10.

Winstone N. E., Nash R. A., Rowntree J. and Menezes R., (2015), What do students want most from written feedback information? Distinguishing necessities from luxuries using a budgeting methodology, *Assess. Eval. High. Educ.*, **41**, 1237–1253.

Winstone N. E. and Nash R. A., (2016), *The Developing Engagement with Feedback Toolkit*, York, UK: Higher Education Academy.

Winstone N. E., Nash R. A., Parker M. and Rowntree J., (2017), Supporting Learners' Agentic Engagement with Feedback: A Systematic Review and a Taxonomy of Recipience Processes, *Educ. Psychologist*, **52**, 17–37.

Winstone N., Pitt E. and Nash R., (2021), Educators' perceptions of responsibility-sharing in feedback processes, *Assess. Eval. High. Educ.*, **46** (1) 118–131, DOI: **10.1080/02602938.2020.1748569**.

Wu Q. and Jessop T., (2018), Formative Assessment: Missing in Action in Both Research-Intensive and Teaching Focused Universities? *Assess. Eval. High. Educ.*, **43**, 1–13.

Yin R. K. (2002), *Case Study Research: Design and Methods*. Thousand Oaks, CA: SAGE Publications.

Zimmerman B. J. and Kitsantas A., (1997), Developmental phases in self-regulation: shifting from process to outcome goals, *J. Educ. Psychol.*, **89**, 29–36.

Zimmerman B. J. (2000), Attaining Self-Regulation: A Social Cognitive Perspective, in M. Boekaerts, P. R. Pintrich and M. Zeidner (ed.), *Handbook of Self-Regulation*, San Diego, CA: Academic Press, pp. 13–39.

Zimmerman B. J. and Kitsantas A., (2002), Acquiring writing revision and self-regulatory skill through observation and emulation, *J. Educ. Psychology*, **94**, 660–668.

Zimmerman, B. J., (2008), Investigating Self-Regulation and Motivation: Historical Background, Methodological Developments, and Future Prospects, *Am. Educ. Res. J.*, **45**, 166–183.