AdvanceHE

MEMBER BENEFIT

Impacts of higher education assessment and feedback policy and practice on students: a review of the literature 2016-2021

Edd Pitt, PhD SFHEA, and Kathleen M Quinlan, PhD PFHEA

Centre for the Study of Higher Education,

University of Kent

Contents

Foreword		3
Executive summary		4
1	Purpose and focus of this review	6
2	Methods	7
2.1	Procedures	7
2.2	Characteristics of the studies reviewed	9
3	Assessment policies and practices	13
3.1	High-impact practices (HIPs)	14
3.1.1	Capstone courses	14
3.1.2	Work-integrated learning	15
3.1.3	Undergraduate research	18
3.2	Authentic assessment	19
3.2.1	What is authentic assessment?	19
3.2.2	How effective is authentic assessment overall?	20
3.2.3	Impacts of simulations	22
3.2.4	Impacts of problem-based learning	26
3.2.5	Impacts of groupwork assessment	27
3.2.6	Impacts of portfolios and e-portfolios	29
3.3	Timing: competency-based learning	31
3.4	Timing: continuous assessment	32
3.4.1	Discussion boards, blogs and micro-blogging	35
3.5	Purposes: formative assessment	35
3.5.1	Automated assessments	38
3.5.2	Gamified learning: adjusting traditional assessments	40
3.6	Purposes: summative assessment	42
3.6.1	Constructive alignment and equivalencies across formats	42
3.6.2	Online exams	43
3.6.3	Beyond grades: other ways of representing student achievement	45
3.7	Design process: co-assessment	45
4	Feedback policies and practices	47
4.1	Introduction	47
4.2	Feedback expectations	48
4.3	The role of grades in feedback	51

4.4	Building relationships through feedback	52
4.5	Early and often: formative feedback for learning	53
4.6	Students' emotions about feedback	56
4.7	Technology-enhanced feedback	59
4.7.1	Immediate, automated feedback	59
4.7.2	Written versus audio/video feedback	60
4.8	Types of feedback offered	62
4.8.1	Elaborative versus corrective feedback	62
4.9	Exemplars: helping students understand quality standards	63
5	Peer assessment and peer feedback	66
5.1	Peer assessment	66
5.1.1	Reliability of peer judgments	68
5.1.2	Technology-enhanced peer assessment	69
5.2	Peer feedback	70
5.2.1	Training and peer feedback	71
5.2.2	Anonymity in peer feedback	72
5.2.3	Dialogue in peer feedback on and offline	72
5.3	Benefits to providers versus receivers of peer assessment/feedback	73
6	Conclusions and recommendations	76
6.1	Conclusions	76
6.2	Recommendations	76
6.2.1	Recommendations for practitioners	76
6.2.2	Recommendations for policymakers	79
6.2.3	Recommendations for assessment and feedback researchers	79
Acknowledgements		81

Foreword

Educators know that assessment and feedback practices are among the most effective levers for improving student learning in higher education (HE). Amidst current uncertainty through a variety of disruptors and with the shifting pedagogic landscape in global HE there is a pressing need to consider how we can better use the significant resources devoted to assessment to support students' learning, academic staff development, and to improve assessment's fitness for purpose in general.

This systematic literature review provides a timely update that will influence the Advance HE Transforming assessment in higher education framework. Focusing on evidence-based policy and practice that has a demonstrable impact on student outcomes from the period 2016 – 2021 the review identified 481 relevant empirical research articles, from 71 different countries and six continents, in the areas of assessment (n=216), feedback practices (n=201), and peer assessment/peer feedback (n=64). This context provides an opportunity for researchers to tend to a much wider range of literature that has been published in different locations, highlighting a variety of research areas and addressing a more globalised understanding of assessment and feedback.

This review is a comprehensive and useful resource for colleagues considering assessment and feedback. In particular, when evaluating the points of intersection with other fundamental areas of student success such as retention, attainment, inclusivity, flexibility, partnership, and outcomes, these core areas of assessment and feedback are crucial. The authors offer a series of recommendations based on a wealth of evidence, each of which is well considered and designed to assist higher education institutions, policymakers, and researchers in identifying aspects each group may wish to prioritise, consider, or review in relation to current assessment and feedback policy and practise. These 'high level' recommendations are typically nuanced and context is crucial; the authors have done an excellent job of providing additional details and explanation to each of these, ultimately with a view to rethinking and repositioning assessment strategies, processes and practices across the sector.

Ensuring assessment is fit for purpose, now and in the future, is crucial: clarifying the purpose and relevance of assessment to meet 21st century requirements within HE will help prepare the sector for future challenges and disruptors, ensuring we continue to develop the knowledge, skills, and expertise in students and support their positive trajectory through higher education. What is evident from the review is that assessment practice remains an integral part of curriculum design and it is a vehicle that can facilitate change in order to better support students to manage the learning requirements of 21st century environments, both within and beyond HE.

It leaves me to thank the authors, Dr Edd Pitt and Professor Kathleen M Quinlan, for their thoughtful and considered synthesis and analysis. Their review is an important addition to the ongoing interrelated work being undertaken by Advance HE through the review and update of the Essential Frameworks for Enhancing Student Success, ensuring that you, our members, have access to current, relevant, and research informed practice.

Stuart Norton
Senior Adviser Learning and Teaching
Advance HE

Executive summary

This literature review systematically explores evidence from peer-reviewed journal articles relating to assessment and feedback in higher education (HE) published from 2016 to 2021. Our aim is to highlight evidence-based assessment and feedback policies or practices that have had a demonstrable impact on key student outcomes including performance, engagement and satisfaction.

Assessment and feedback practices are among the most powerful levers educators have for improving student learning. Assessment is much more than testing students. Feedback is also much more than teachers telling students how well they have done. We advocate broad conceptions of assessment and feedback to centre student learning and prompt a rethinking of traditional views of assessment and feedback.

Assessment is a curricular procedure in which students do something that generates information about what or how much they have learned to date. The performance-related information resulting from assessment procedures – whenever they occur in the learning process – is feedback information. Students must make sense of this feedback information to promote their further learning. While feedback is traditionally assumed to come from teachers, we also showcase the role of peers in assessment and feedback. Although these three concepts of assessment, feedback and peer assessment/feedback are related and work together, we address each separately in this review, consistent with the Advance HE Transforming Assessment in Higher Education Framework (Advance HE, 2015).

We identified 481 relevant empirical research articles in the areas of: a) assessment (n=216); b) feedback practices (n=201); c) peer assessment/peer feedback (n=64), which we supplemented with related literature reviews. Section 2 provides more details on the methods. All the empirical references are provided in an accompanying database searchable by sub-category, discipline, country, author and date. We focus on the highest quality research in the narrative, citing evidence from studies from around the world.

In terms of assessment, we argue in section 3 that effective assessment practices should integrate as many powerful learning principles as possible. In section 3.1, we describe these principles and how they come together in high-impact educational activities such as work-integrated learning and undergraduate research. Authentic assessments also promote student learning insofar as they apply many of those principles. Authentic assessments reflect real-world demands, are cognitively (and, potentially, emotionally and ethically) challenging, and help students learn how to judge their own performance. They are designed to take place in typical HE curricular structures.

In general, authentic assessments promote meaningful student learning and growth, student engagement and student satisfaction. There are many types of authentic assessment, and their effectiveness depends on how well they are implemented. We discuss and offer recommendations on four types in section 3.2. Preparing students for novel assessments and clarifying assessment criteria is particularly important. We specifically urge the sector to create more opportunities for group work assessments, as teamwork is so integral to most professional roles. Authentic assessments can also generate products that showcase students' capabilities more richly than grades alone, making them potentially useful for rethinking how student performance is communicated to HE stakeholders (section 3.6.2).

We also review innovations related to traditional assessment practices such as quizzes and exams that tend to focus on content knowledge, rather than the complex application and integration of knowledge and skills in novel situations emphasised in authentic assessments. Traditional knowledge assessments can be improved by focusing on conceptual support, increasing student interaction and gamification, and enriching the feedback design associated with them. The design of feedback opportunities should be built into assessment processes to offer an integrated and coordinated approach to improve student learning outcomes.

In section 4 on feedback, our key argument is that for feedback to be effective, students need to make use of it. Thus, students need feedback designs that increase their understanding of performance criteria and standards, the role of feedback, how to respond to it constructively, and how to enact it in their subsequent work. Toward that end, we highlight that lecturer feedback is more effective when it is part of an ongoing relationship with students, offers opportunities for dialogue between lecturers and students and among students over time. Feedback also has a greater effect on students when there are opportunities to enact it prior to summative submission. Too often, students receive feedback on only one or two occasions in a module with no opportunity to apply feedback to a revision or to a similar task. Thus, feedback interventions such as discussions of exemplars can be helpful in building students' understanding of assessment standards if they are done repeatedly. Video and screencast feedback from teachers seems to promote more learning than written feedback because feedback is often more personal, expansive and elaborative in that format. Elaborative feedback that explains how and why seems to help students improve their work more than corrective feedback that only points out or corrects problems.

Peer assessment and peer feedback (section 5) can also play an important role in improving student learning outcomes. Web-based platforms can facilitate dialogue between peers about quality and standards. Technology affords anonymity in peer exchanges that helps students feel more comfortable offering critical comments, especially if they are new to giving feedback. Students benefit from guidance from lecturers, often in the form of assessment criteria or rubrics, training or at least cycles that include both peer and lecturer feedback. Peer feedback generally benefits the provider of feedback more than the receiver because the power of these interventions lies in engaging students actively with criteria and standards and building their capacity to recognise quality. They then bring this critical eye back to their own work, seeing where it could be improved. Thus, there was emerging evidence that the use of peer assessment and peer feedback may act by developing students' capacity for self-assessment.

Based on our review of evidence about a wide range of assessment, feedback and peer feedback practices, section 6 presents 12 recommendations for practitioners, four recommendations for policymakers and five recommendations for researchers.

1 Purpose and focus of this review

This review is one of a series updating evidence related to Advance HE's Essential Frameworks for Enhancing Student Success, which address seven core themes, each with its own framework (Advance HE, 2019). Although each of the core themes/frameworks are being reviewed separately, assessment and feedback touches on many of the other themes, including employability; retention, attainment and progression; and flexible learning. Ultimately, student success requires thinking holistically about students' experiences of higher education (HE).

This review identifies recent (2016-2021) literature on assessment and feedback in HE, focusing on evidence-based policy and practice that has a demonstrable impact on student outcomes, including student performance, progression, engagement, satisfaction, skill acquisition and/or self-confidence. Its aims were practical; we intended to help practitioners, policymakers and researchers to focus more effectively on relevant questions, issues or sources of evidence to inform their own research or practice. The review has been guided by the Advance HE Transforming Assessment in Higher Education Framework (Advance HE, 2015) and the associated guide to the Framework, which has been designed to help institutions and educators to "review current policy and practice in assessment and feedback, with a view to rethinking and repositioning their assessment strategy, processes and practices" (Elkington, 2020, 4).

The main research questions were:

1. Based on a review of literature since 2016, what is the state of the field of assessment and feedback in HE? What are the emerging trends and open questions?
2. What demonstrable impact have specific assessment and feedback policies and practices had on student outcomes (eg student performance, progression, engagement, satisfaction, skill acquisition and/or self-confidence)?
3. How do the identified areas relate to and move forward HE assessment and feedback practice, including the Advance HE Transforming Assessment in Higher Education Framework (Advance HE, 2015)?

2 Methods

In this section, we describe the review process and characterise the articles included before we discuss limitations of the methods.

2.1 Procedures

We systematically searched peer-reviewed journal article publications relating to assessment and feedback within the predefined timeframe of 2016 to 2021. The literature search was conducted using EBSCOhost, which includes: PsychInfo, British Education Index and ERIC – key databases of educational literature in English. We then carried out a search using Web of Science. We limited our searches by date, peer reviewed articles, and written in English. To identify articles focused on higher education, we used the search terms 'higher education' or 'college' or 'university' or 'post-secondary' or 'postsecondary'. To limit articles to those focused on assessment and feedback, we used the terms 'assessment' or 'evaluation' or 'feedback' or 'feedforward'. We used the terms; 'student outcomes' or 'student performance' or 'student achievement' or 'student engagement' or 'student satisfaction' or 'skill acquisition' or 'student self-confidence', to find articles that fit addressed demonstrable outcomes.

Importantly, the search focused on abstracts rather than keywords. Most articles are limited to just five keywords, which can be chosen idiosyncratically, while abstracts provide a better representation of the content. In a wide-ranging field where keywords are not well-standardised, this approach was more thorough in uncovering relevant articles. This choice, though, meant that the automated search results were over-inclusive, which required screening out on reading the abstracts and/or the full texts.

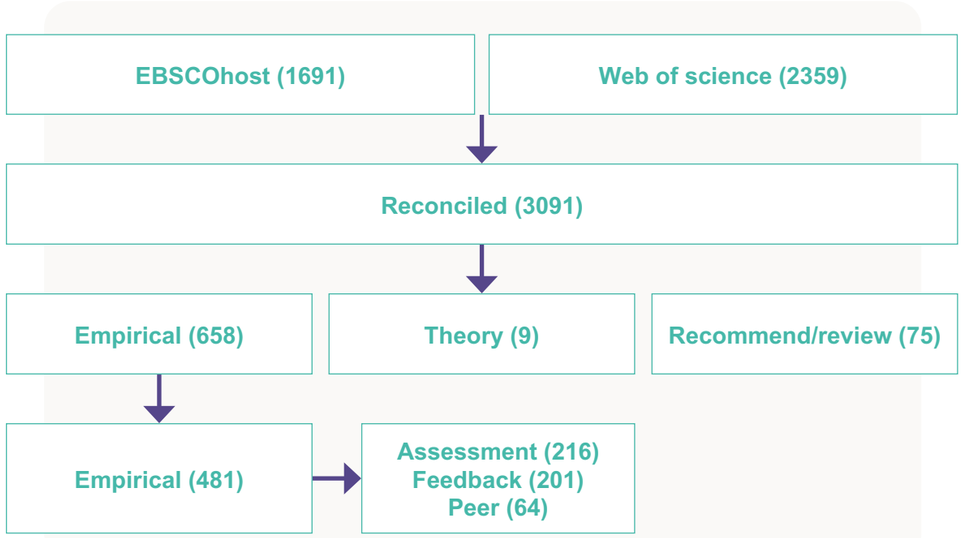
The EBSCOhost search revealed 1,691 articles and the Web of Science search produced 2,359 articles. We imported both lists and reconciled the close matches to produce an initial database of 3,091 articles (see Figure 1). We screened all 3,091 abstracts to ensure they addressed the research questions.

As our emphasis was on practice by teachers within higher education, consistent with the Advance HE Transforming Assessment in Higher Education Framework and its associated guide (Advance E, 2015; Elkington, 2020), we excluded multi-institutional learning outcomes assessments associated with, for example, institutional accreditation, professional licensing, qualifications frameworks and overall comparability of learning outcomes across institutions and countries. For an international review of those issues, see Coates (2016) and a special issue edited by Coates and Zlatkin-Troitschanskaia (2019). Kinzie (2019) summarised this assessment movement in the US and open issues, while a major study in the US suggested little impact of accountability systems related to assessment of learning outcomes on student experiences (Cox et al, 2017).

Through our first reading of the abstracts, we categorised them as: empirical studies related to the research questions (n=658), theory (n=9), or recommendation/review (n=75). We then reviewed the full text of the empirical articles, which reduced the sample to 481 articles. The empirical articles formed the centrepiece of our review, though we also drew on existing literature reviews to supplement the primary evidence base. We used the theory papers to identify key debates and concepts in the field. Using the Advance HE Transforming Assessment in Higher Education Framework's focal areas

(hereafter 'The Framework') (Advance HE, 2015), we coded the selected empirical articles into three categories: a) assessment (n=216); b) feedback practices (n=201); c) peer assessment/peer feedback (n=64) (Figure 1).

Figure 1. A summary of the process refining the reviewed articles



Through our second reading of abstracts and full texts, we identified sub-categories within each of the major sections of assessment (Section 3), feedback (Section 4) and peers (section 5), which were used as an initial guide to organising the text. Where available, we highlighted review articles within each of these sub-headings.

In the narrative, we showcase studies with more rigorous evidence standards or that offered unique insights that might guide future research and practice. The review has been conducted within an educational and cultural environment of accountability that privileges some forms of knowledge and knowledge production processes over others. In addition to any standards used by Anglophone peer reviewers, we also relied on published evidence standards in selecting studies to feature in our narrative.

We reviewed evaluation evidence standards recommended by the Office for Students in England. The Office for Students distinguishes between studies that offer narrative, empirical, or causal evidence that a policy or practice has an impact on desired outcomes (Office for Students, 2021). By definition, our review focused, at least, on empirical evidence. Demonstrating causal links between

an intervention and observed outcomes is more challenging. We sought to showcase articles that used study designs that allowed causal claims to be made. To help us in critiquing study designs, we considered the evidence standards used in reviews for the United States Department of Education's What Works Clearinghouse in Education (Institute of Education Sciences, 2018). In drawing conclusions, we have relied more heavily on studies that, according to those UK and US evidence standards, are of higher quality. Throughout the review, we provide commentary on our observations about study design, the evidence available, and limitations in methods and approaches that characterise the research reviewed, along with recommendations for future research.

2.2 Characteristics of the studies reviewed

We illustrate the scope of the empirical articles in a variety of ways. Table 1 indicates the year of publication and Table 2 shows the regionality of the studies included. The articles represented a wide range of disciplines and were coded according to the Becher-Biglan (Becher, 1989) typology (see Table 3 below). This typology classifies disciplines into 'hard pure' (eg maths, sciences), 'hard applied' (eg computing, engineering, medicine), 'soft pure' (eg social sciences, economics, history) and 'soft applied' (eg education, languages and business, nursing). For more detail, the accompanying database is searchable by sub-category, discipline, country, author and date.

Table 1. Articles by year of publication

Year of publication	2016	2017	2018	2019	2020	2021	Total
Assessment	31	39	35	39	44	28	216
Feedback	29	31	35	45	35	26	201
Peer	5	9	8	12	16	14	64
Total	65	79	78	96	95	68	481

Table 2. Distribution of articles by regions of the world

Region	Assessment	Feedback	Peer
Africa	10	3	0
Asia	28	33	19
Europe	70	87	24
Middle East	21	10	4
North America	41	38	13
Oceania	41	24	4
South America	5	6	0
Total	214	201	64

Figure 2. Distribution of articles by regions of the world

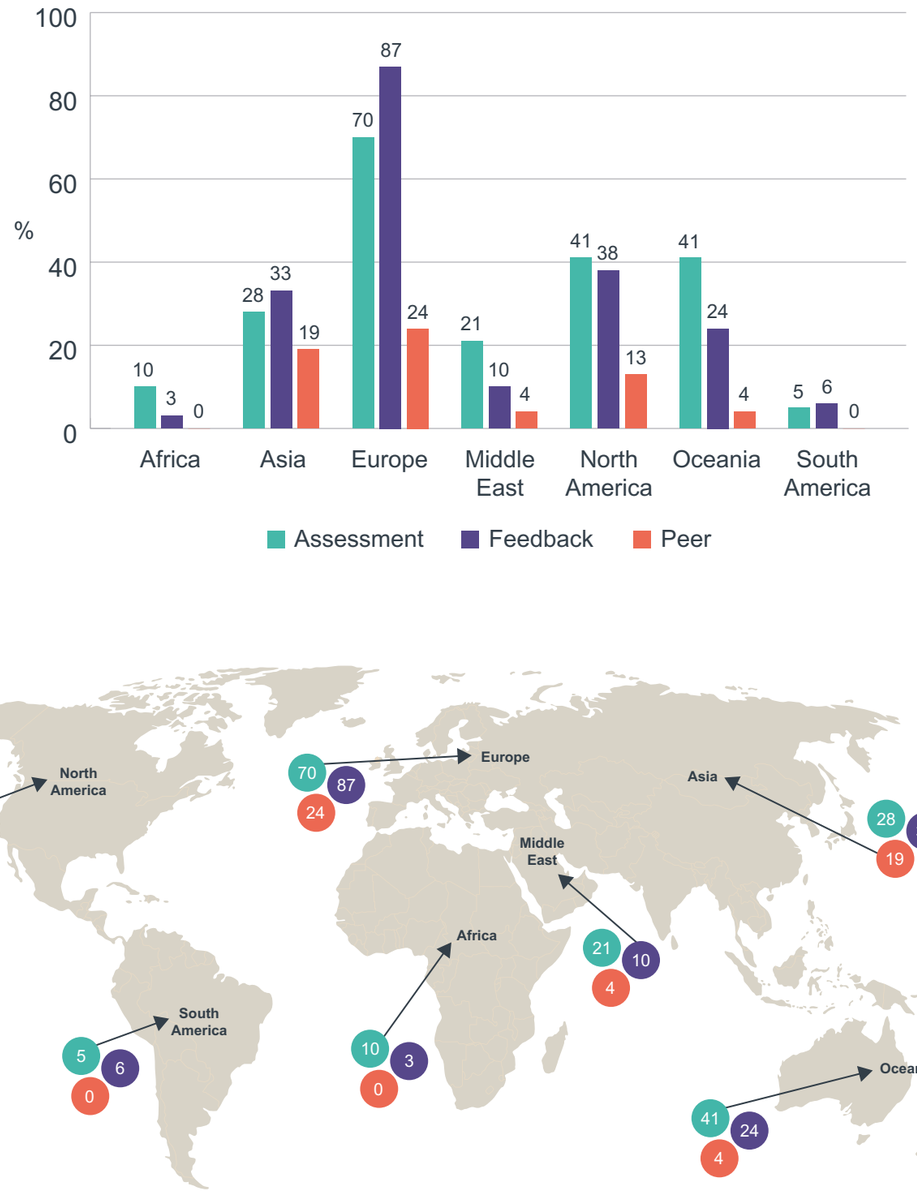
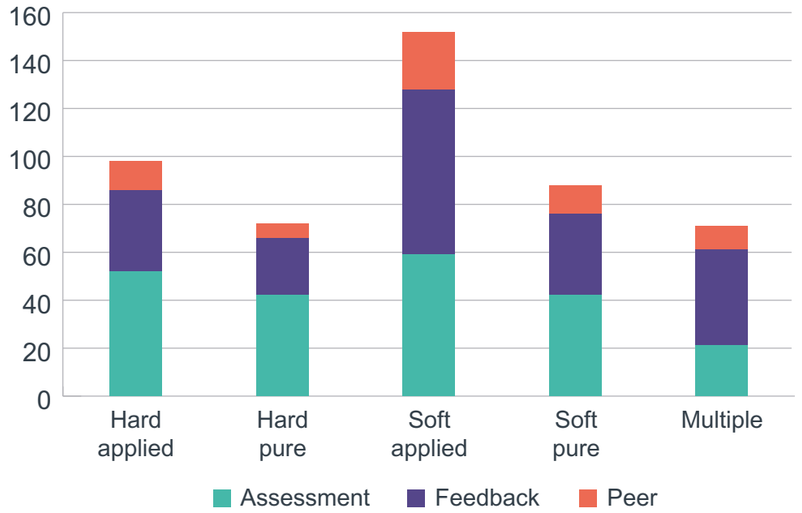


Table 3. Distribution of articles by discipline according to the Becher Biglan Typology

Discipline code	Assessment	Feedback	Peer
Hard Applied	52	34	12
Hard Pure	42	24	6
Soft Applied	59	69	24
Soft Pure	42	34	12
Multiple	21	40	10
Total	216	201	64

Figure 3. Distribution of articles by discipline according to the Becher-Biglan Typology



Limitations

The review takes the form of a narrative review. Given the nature of the evidence frequently used in this field, the wide variety of different practices reviewed, and the time constraints of the study, quantitative meta-analyses were not feasible. Our narrative review brings together a large body of new research into a structured report related to the Framework (Advance HE, 2015).

One limitation of all systematic reviews is bias in the selection or emphasis of certain primary studies. To address this issue, we set a threshold that outputs had to have been peer reviewed. Our search yielded studies that were diverse in their design, methodological quality, specific interventions used, types of students studied and outcome measures.

A key challenge and limitation of summarising "what works" in education is that effectiveness often depends on context. The effectiveness of any given practice is shaped by many layers of context: broader policies, systems, disciplines, cultural practices, teacher experience, resources available, and students' backgrounds and expectations. As we searched a wide range of literature from around the world (albeit it only published in English) and wrote this review with an international audience in mind, there is a danger that we have decontextualised or universalised the findings. We have, where possible, indicated which aspects of the findings we highlight may be more context dependent. At times, we have identified a system, intervention, approach, policy or practice that may have differential impacts on various student groups (eg mature/adult learners, commuting students, underrepresented or minoritised ethnicities, undergraduates versus postgraduates).

Our review has focused on research about HE only, though we recognise the value of research on similar practices in primary and secondary education or other educational settings. Given the time constraints of this review, it was impractical to include a broader search of potentially valuable findings from wider educational research that could have implications for HE practitioners. Similarly, we are focusing on teacher assessments and feedback and not discussions of measurement related to high stakes, standardised summative assessments such as medical licensing exams. There may be lessons from those specialised subfields that would inform teacher-based exam design, for instance, but that inquiry was also outside the scope of our review.

Impacts of Higher Education Assessment and Feedback Policy and Practice on Students: A Review of the Literature 2016-2021
Edd Pitt and Kathleen M Quinlan

3 Assessment policies and practices

The Framework (Advance HE, 2015) and its associated guide (Elkington, 2020), which framed this review, began by defining assessment as the "means by which we assure and express academic standards" (Elkington, 2020, 5) but then called for a transformation of assessment to focus more strongly on learning and learners. Elkington (2020) argued for diversification of assessment methods with greater validity, authenticity and inclusivity; fewer larger summative assessments that address programme-level objectives; and more opportunities for formative assessment. He acknowledged that assessment has been traditionally difficult to change, and that modularisation has led to a proliferation of smaller, atomised and, potentially, decontextualised assessments that put students and academics into what Harland et al (2015) called an "assessment arms race". Elkington (2020) argued that improvements need to be made in terms of students' involvement in the assessment/feedback process, staff understanding of assessment, and infrastructure, such as harnessing the power of technology.

The Framework (Advance HE, 2015) promotes what has come to be called 'assessment for learning' (AfL) (Stobart, 2008). Several principles are central to AfL, including a) clarifying and sharing intended learning outcomes and criteria for success; b) designing effective assessment tasks that generate evidence of students' achievement against those learning outcomes; c) providing feedback that moves learners forward; d) engaging students as instructional resources for each other; e) and activating students as the owners of their own learning through processes such as self-assessment (Black and Wiliam, 2018). The Advance HE Framework mirrors these AfL principles, although it did not centre intended learning outcomes as strongly as Black and William (2018) did, or indeed, the theory of constructive alignment (Biggs, 2002), which implicitly underpins much of the work on assessment and feedback.

To support the call for transformation of assessment and be consistent with the broader AfL movement, we defined assessment more broadly. Essentially, an educational assessment is a procedure for making inferences about student learning. Assessment is when "learners engage in tasks, which generate data. These data become evidence when they are used in support of particular claims" (Black and Wiliam, 2018, 553). This definition puts learners' activities at the centre, rather than the academics who are expressing and assuring academic standards. Black and Wiliam's (2018) definition leaves open who might be involved in interpreting or making inferences about student learning, allowing the possibility for teachers, peers, students themselves, employers or other stakeholders to make those inferences and claims for a variety of purposes. This broad definition of assessment makes it challenging to draw the line between instruction and assessment/feedback, particularly when assessment and feedback is integrally embedded in broader instructional practices. We discuss this point as it arises below, particularly in relation to high-impact practices, authentic assessment and continuous assessment.

The Framework (Advance HE, 2015) emphasises 'innovative' assessment. Most of the literature reviewed here also investigated innovative practice, though 'innovative' depends largely on context. What is innovative in one country or field may be standard or typical practice in another. In general, where practices are well-embedded and established, there seems to be little research attention.

For example, there is a strong culture of testing in many contexts, with summative tests, exams and quizzes dominating. Some studies, such as those in the gamified learning section (section 3.4) represent variations or innovations around a traditional approach to quizzing students that may be driven, in part, by new technologies. Tests also can be improved by refocusing them, for example, on thresholds (Bajada et al, 2021); on higher order thinking skills (Carmichael et al, 2016); or toward greater authenticity (Villarroel et al, 2019. In sum, this review should be read as an indication of emerging, innovative trends in practice, rather than a reflection of what is typically occurring in universities. We start with the most transformative, highly effective practices in sections 3.1 and 3.2.

3.1 High-impact practices (HIPs)

High-impact practices (HIPs) are a set of specific, complex, highly effective learning practices. HIPs are defined by eight key elements: high expectations of learners, sustained student effort, meaningful academic and peer interaction, working with diverse others, timely and constructive feedback, real world application, public demonstration of competence, and reflection on and integration of learning (Kuh and O'Donnell, 2013). Thus, a public demonstration of competence (assessment) and feedback are key features of HIPs. HIPs include complex and extended learning activities such as community service learning, internships, capstone courses, collaborative assignments, undergraduate research, and richly designed e-portfolios (Kuh, O'Donnell and Schneider, 2017). HIPs were identified through students' high self-reported improvements during HE on a set of essential learning outcomes measured by the National Survey of Student Engagement (NSSE) (Kuh, O'Donnell and Schneider, 2017). The NSSE has spawned similar engagement surveys around the world, including the United Kingdom Engagement Survey (UKES). In the US, HE institutions are seeking to embed one to three HIPs in each undergraduate's experience (Andrade, 2018).

Each of these high-impact practices offers opportunities to enhance feedback and students' engagement with various sources of feedback (Quinlan and Pitt, 2021) throughout the learning experience, as well as presenting challenges to assigning marks for summative purposes. As we discuss three HIPs below for which we found relevant literature, we will focus on how assessment and feedback is optimised within them. We will not focus on their overall impact on students' learning outcomes, as each of them has a considerable body of evidence of impacts on students' learning beyond the scope of this review. For example, community service learning is associated with personal, social learning and career development benefits (Eyler, Giles, Stenson and Gray, 2001). Note that we address e-portfolios in 3.2.4 below, rather than here because the practices described in literature in this period did not meet the eight HIP criteria.

3.1.1 Capstone courses

Thompson and Houston (2017) provided a rich example of how to blend formative and summative assessment within an Australian capstone course for paramedics. Their capstone prompted students to look back on, review and integrate their learning through early formative assessments such as a pre-course diagnostic multiple choice test, while subsequently engaging in a series of activities such as problem-based learning (PBL) and practical scenarios. The course culminated in a final summative

oral interview exam that simulates those used by industry professionals in the hiring process. Following Knight (2002), they emphasised that the distinction between formative and summative assessment is not one of different tools or techniques of assessment, but of use. Each assessment generates information; when that information is primarily used for sense-making to further learning processes by teachers and students, it is formative. When the information is primarily used in making claims about competency for communication to external audiences, it is summative. According to students' responses, Houston and Thompson's (2017) capstone built a programme of assessment that bridged the formative and summative divide by clarifying purposes and creating motivation for students to engage with assessment-cum-learning activities.

3.1.2 Work-integrated learning

Internships

Work-integrated learning, which has gone by various terms, including work-related and work-based learning, has focused primarily on placements, internships or practicums in workplaces in which students are learning while working. Typically, these have been extended opportunities such a 'year-in-industry', 'sandwich years' or cooperative education in which an entire semester or year in workplaces alternates with time in the classroom. More recently, in the UK, apprenticeship learning has become an alternative route to a bachelor's degree in which learners spend 80% of their time 'on-the-job' with workplace mentors while spending 20% of their time in formal learning, opening routes for learners who might not otherwise thrive in traditional classroom-based HE (Saville et al, 2020). These kinds of workplace immersions are considered a high-impact educational practice as described above.

Students may be assessed by workplace supervisors based on workplace performance and/or through reflective logs, portfolios or other tasks set by academic supervisors. In Iran, Akashe and colleagues (2020) interviewed law internship supervisors, experts and academic staff and surveyed law students about desirable methods of assessing internships. Workplace supervisors emphasised exams with practical performance tasks, conducted by academics: "they should be able to write a bill, issue an order or a verdict regarding a case" (Akashe et al, 2020, 17). Students also preferred these kinds of simulated, practical assessments over various other common assessment approaches such as portfolios, observations, diaries and checklists (for more, see section 3.2.1 below on simulations). Caution may also be warranted in using diaries, as it is challenging for students to engage in high-quality reflection on their experiences and growing competencies without focused instruction and structured prompts (Vidiella and Garcia, 2016).

Aligning assessments with intended learning outcomes is vital (Biggs, 2002). Ramsgaard and Østergaard (2018) explicitly designed aligned assessments for health and nutrition students' six-month internships in Denmark. They assessed students' ability to identify opportunities in a practice setting, to create actions from ideas, and to transform ideas into valuable activities in a practice setting. They used portfolios and a final oral exam in which students presented (20 minutes) about these learning outcomes and then participated in a 20-minute dialogue. Students reported that

this exam was an important learning experience and motivated them, appreciating the opportunity to set the agenda themselves and be 'in charge'.

Hallenberg and Haddow (2016) incorporated service learning into a newly designed Criminology for a Just Society course in the UK. They assessed students with reflective blogs completed across the course of the academic year, a case study in which students critically reflected on the policies and practices of their chosen volunteering organisation in relation to sustainability and justice, and an end-of-module conference in which students gave short presentations on their experience to the volunteering organisations.

In an alternative to traditional internships that bridged the university/workplace divide, Teichgräber et al (2021) trialled a structured approach to a half-day clinical radiology immersion experience in Germany. Rather than passively observing physicians, students spent one hour each in two workplaces in which they were assessed on the completion of constructively aligned mini clinical evaluation exercises (for more on these kinds of exams, see 3.2.1 below). In the first workplace, students read chest X-ray images and created a medical report with ratings from resident physicians. In the second workplace, they were assessed on how they informed a patient on preparation for an MRI or CT scan. Students and clinicians were satisfied, and students performed well, though there was no group with which to compare students' performance or satisfaction. Teichgräber et al (2021) described how they scaffolded instruction before the immersion experience with short videos and during the immersion experience, culminating in students' performance of these key clinical tasks.

Some articles offered rich case examples of innovative work-integrated learning where it hadn't previously existed. For example, Heitkamp et al (2018) introduced a five-day work-shadowing component into dental education in Germany. Based on self-assessed competencies, students reported increases from pre- to post-intervention in the areas of communication and dental practice management as well as some clinical skills. Both students and participating dentists were satisfied with the experiences. Likewise, Pruett et al (2018) evaluated the impact of interprofessional out-of-court family divorce services on 18 student interns from law, psychology and social work, with particular attention to law students. The interdisciplinarity of this clinic was a key innovation.

Live case studies

Because few students have opportunities for high-quality internships, academics are developing other ways to engage students in authentic, work-integrated learning opportunities (see also section 3.2 on authentic assessment below). Kaider et al (2017) reviewed 1500 assessments in an Australian university on two dimensions: authenticity of assessment tasks to work-related tasks, and proximity of the assessment to the workplace. They developed and tested a three-by-three table (low, medium, high in each dimension) with examples of assessments that fit into each of the resulting nine cells. This authenticity/proximity framework is useful for classifying different assessments and prompting discussion about how to make existing assessments more authentic and/or more proximal to workplaces. Kaider et al (2017) concluded that it was easier to offer authentic tasks than it was to engage with employers, though they suggested ways that authentic assessments could be brought

into greater proximity to workplaces. Some ideas included doing projects for or with industry in 'live business projects', also sometimes called client-initiated or client-sponsored projects, live case studies or collaborative industry-led projects. These live cases also involve public performances, either through reports or presentations back to collaborating employers. Well-designed live case studies meet the eight criteria for high-impact practices.

In Chile, Valenzuela et al (2018) provided the most robust evaluation of live case studies, offering a model for similar research on the impact of assessment practices. They compared students' performance in a newly developed 'live cases' condition to the previous year's cohort in which the students did the same assessment activities but used fictional cases and data. Students' grades and scores on various aspects of the report rubric were significantly higher for the live cases than the fictionalised cases, with large effect sizes. The live approach also seemed to benefit lower-performing students the most, which suggests its potential usefulness in closing achievement gaps between less and more educationally, socioeconomically or culturally advantaged students. Students also rated their teachers much higher and had overall more favourable perceptions of their learning and academic practice (Valenzuela et al, 2018).

Lu et al (2018) studied students' reactions to an international business plan competition carried out by students in an international marketing course in Australia. Four or five teams of four to five students were assigned to one client and competed in the development of a plan for international market expansion. The pedagogical design incorporated both cooperation within the team and competition between teams, though there was no discussion of the mechanisms for individual accountability within the team (as is emphasised under group assessment in section 3.2.3 below). In focus groups, students reported more future-oriented thinking, gains in self-confidence, career preparation, ability to apply knowledge to real world contexts, increased motivation and interest via interaction with a real client, and the need to adapt to external constraints.

Offering a richer description of individual accountability within groupwork, Schonell and Macklin (2019) documented spending several years honing their live case studies to create a less resource intensive, but still pedagogically sound work-integrated learning for large classes in Australia. To reduce time burden on markers and create greater authenticity to the workplace, they shifted from individual written reports to group reports, but encountered student dissatisfaction with peers who did not work very hard. In response, they added explicit instruction on group work and introduced various process measures for individual student accountability, including meeting minutes, and reports on task allocations and completions, which contributed to individual marks (see group assessment section 3.2.3 below). Most students agreed that the learning activities helped them achieve the learning outcomes and that the assessments measured those learning outcomes. Business collaborators who received students' consulting reports were impressed with the quality of their work.

Based on the literature from the past five years, 'live case studies' seem to be highly effective educational practices, but the research on them is largely confined to business courses. Similar pedagogical designs are seen in service learning in a variety of disciplinary contexts (eg history, arts and design, sociology, environmental science) in which students work on a project or 'live briefs' in a

community/service context, rather than for a business client. However, our search did not uncover studies of their effectiveness. The challenges of connecting with and maintaining relationships with other stakeholders may hinder implementation.

3.1.3 Undergraduate research

There is more than two decades of evidence of the effectiveness of mentored undergraduate research experiences on students' retention (eg Nagda et al, 1998) and interest in pursuing a PhD (Seymour, 2004), particularly among minoritised students. As with other workplace internships, students perform and receive feedback from mentors and peers throughout the research experience, thus assessment and feedback are an integral part of the overall experience. Normally, students engage in a public capstone presentation of their research. Our search uncovered three studies that added incrementally to the body of literature on undergraduate research experiences, though none included control groups. Two of the studies extended research into course-based undergraduate research experiences (called CUREs) in the US, a novel extension of traditional independent undergraduate research projects (eg undergraduate dissertations). Students in a US-based CURE reported gains in professional skills, understanding of the research process, and how researchers work and think (Sangster et al, 2016). Students from a variety of disciplines who took a one-semester American introductory science research course reported personal gains related to research, to thinking and working like scientists, to attitudes and behaviours of a scientist, and to gains in skills as reported on the Undergraduate Research Student Self-Assessment (URSSA) (Sandquist, 2019).

While most of the earlier research focused on science, technology, engineering and mathematics (STEM) subjects, two of the recent studies included both STEM and non-STEM research experiences (Sangster et al, 2016; Stanford et al, 2017), and found no significant differences between the two groups in gains in knowledge, skills and attitudes as measured by the URSSA (Stanford et al, 2017). Nguyen and Walker (2016) also found that students' undergraduate final year dissertations were praised by both lecturers and students in two sociology departments (one in the UK and one in Vietnam) for promoting students' independent learning skills.

This topic is ripe for a systematic review, as many North American institutions have implemented variations on undergraduate research and CUREs in different disciplines and at different levels (eg first year versus final year). Meanwhile, institutions around the world may include final year undergraduate theses/research experiences. Analysing different kinds of support structures and instructional designs (eg size of peer groups, research coaches versus administrators) may help identify the most cost-effective ways of implementing this high-impact practice. Healey and Jenkins (2009) prepared helpful case examples and an overall framework for integrating research into undergraduate education more than 10 years ago for the then Higher Education Academy. While many UK universities engage students in final year research projects, North American research suggests that early involvement through core courses may yield other timely benefits, such as supporting greater progression of racially minoritised students into research careers (Bangera and Brownell, 2014).

3.2 Authentic assessment

Because authentic assessment is a broad and important concept in assessment, we will first discuss definitions before moving on to studies on its effectiveness. As there is considerable literature on different types of authentic assessments, we will review research on the demonstrable impacts of each of four common types of authentic assessments separately: simulations, problem-based learning, assessment of groupwork, and (e)portfolios.

3.2.1 What is authentic assessment?

Villarroel and colleagues (2018) reviewed literature on design dimensions required to bring authenticity to the assessment of classroom learning, as opposed to work-based or community-based learning, as is common in the HIPs described above. They analysed 112 articles on authentic assessment between 1988 and 2015 to identify key characteristics of authentic assessment. Villarroel et al (2018) identified 13 characteristics, which they grouped into three main categories.

First, Villarroel et al (2018) argue that authentic assessment involves realism, which emphasises the relevance of problems to everyday or professional life and its practical value. Realism may be a central feature of an assessment such as a simulation, or it can be introduced in traditional individual assessments by contextualising knowledge through, for example, asking students to write essays in relation to relevant case scenarios. For example, Zavhorodnia et al (2021) found that law and historical studies students reported higher confidence in practice-based skills after preparing for and engaging in practice-based exams in which they either analysed legal cases or characterised historical source materials. Realism may also be achieved through authentic performances/products, in which students demonstrate their knowledge and skills in tasks that simulate employment situations. For example, in a course on diversity in the US, Kang (2019) introduced an oral exam in place of a written exam to ensure an authentic assessment of students' ability to explain concepts.

Second, authentic assessments offer cognitive challenge, focusing on higher order thinking, application, problem solving and decision-making. Third, authentic assessments must develop students' evaluative judgment through, for example, transparency of criteria and opportunities for feedback. Evaluative judgment means that students understand what constitutes good performance so they can self-regulate toward improvement (Villarroel et al, 2018). See section 4.9 on exemplars and section 5 on peer feedback and peer assessment discussion, for more ways to enhance students' evaluative judgment.

Villaroel et al's (2018) definition can be broadened in several ways, though. First, authenticity is usually gauged in relation to professional workplaces (Villarroel et al, 2018), but that reference point can be challenged. McArthur (2020), for example, proposed a broader reference to the real world, not just the world of work, paying particular attention to social justice and public good. Relatedly, authentic assessments are also typically defined in terms of cognitive skills, while attending less to students' emotional and ethical development, a criticism that has been raised about higher education instruction and assessment more generally (Quinlan, 2011; Quinlan, 2016). More work could be done to create

and robustly evaluate authentic assessments through the lens of students' holistic development (Quinlan, 2011; Quinlan, 2016) and public good concerns.

Rather than define authenticity in relation to workplace practices, Quinlan and Pitt (2021), in developing the related concept of "signature assessment", defined authenticity in relation to the intellectual, social and material practices of the discipline. In an example of attention to the discipline itself, Bajada et al (2021) re-designed a first-year business statistics assessment to focus on mastery of a set of six key threshold concepts for the field, based on a whole-of-degree approach that considered the expectations of graduates in professional roles. Rather than awarding part-credit for partial understanding of key concepts, threshold assessments were designed to test understanding of the whole threshold. Students had three opportunities to pass a holistic assessment of each threshold, with earlier attempts followed by feedback. Threshold assessments contributed 60% of their module mark. The probability of attaining a credit or above in advanced subjects requiring statistics increased with students' increase in understanding of the thresholds (for more on mastery learning, see competency-based assessments below).

Cognitive challenge is also subject to discussion. Most authors define the knowledge, skills and attitudes to be addressed by authentic assessment in terms of a revision of Bloom's taxonomy of the cognitive domain (Anderson and Krathwohl, 2001). But even cognitive desired outcomes are open to debate and more precise specification. For example, Alexander (2019) argued for the importance of relational reasoning, defining it as the ability to discern meaningful patterns within any informational stream, including perceived similarities (analogical), discrepancies (anomalous), contradictions (antinomous), and contrasts (antithetical). Various professionals engage in these forms of reasoning: doctors diagnosing cases, science and mathematics teachers teaching, and engineering students designing new products. Understanding the skills used by professionals from studies of professional practice itself can help shape what needs to be assessed in authentic assessment.

Finally, greater clarity could be brought to the field between the authenticity of the products of students' work and the process they used to produce it. For example, Nilolaeva and Korol (2021), in an English as a Foreign Language (EFL) context in Ukraine, sought to assess both the product of students' translation and their translation process. When students provided both their translation of a 250-word text and reflections on their process, they reported high learning value of the process assessment measures and further use of those reflective materials for self-assessment and translation process enhancement (Nikolaeva and Korol, 2021). Understanding the process students used also allowed teachers to give feedback directly on learning and thinking processes (Nikolaeva and Korol, 2021).

3.2.2 How effective is authentic assessment overall?

Sokhanvar et al (2021) systematically reviewed 26 papers published between 2010 and 2019 to identify advantages to students' learning experiences and employability skills through authentic assessment. Their review lists the methods used in the studies but did not apply any evidence standards for selection or reporting. Most of their studies involved student self-reports via surveys or

focus groups with no comparator groups, making it difficult to make causal claims. In those studies, students tended to report positively on their engagement (eight studies), satisfaction (seven studies) as well as key employability skills including communication (seven studies), collaboration (six studies), critical thinking and problem solving (seven studies), applying knowledge to practice (nine studies), reflectivity and self-awareness (10 studies), and self-confidence (six studies) (Sokhanvar et al, 2021). Their studies were mainly in education, business, nursing, and law in the Global North. In the current review, we emphasise studies that meet evidence standards for causality (Office for Students, 2021). In particular, we specifically highlight quasi-experimental or experimental studies and studies with performance measures beyond students' self-rated competencies.

In an Australian laboratory experimental study with random assignment to conditions, James and Casidy (2018) found that undergraduate business students gave higher ratings on probable satisfaction with a hypothetical unit that contained an authentic assessment than those with non-authentic assessment briefs. Satisfaction was higher among students who rated their desire to pursue an interesting career higher, suggesting that career-oriented students will be most satisfied with authentic assessments.

In a quasi-experimental study in a Masters of Business Administration (MBA) course in accounting decision-making in the UK, Simpson (2016) compared students' pass rates and satisfaction before (three semesters) and after introducing case-based scenarios into seminars and an authentic summative assessment. The percentage of students who passed ranged from 55%-65% in the traditional style and rose to 73%-83% after the innovation. Contrary to expectation, higher performing students seemed to benefit the most. Students' enjoyment rose from 58%-68% to 89%-100% after the changes. While it is not clear whether the changes were attributable to the changes in teaching mode or changes in the summative assessment per se, both involved opportunities for students to practice and perform in authentic scenarios and formats.

Similar results were found in an Australian electronics engineering course in which group-based, authentic mini-projects were developed to replace traditional, passive problem sheet activities in tutorials (Shahnia and Yengejeh, 2019). Shahnia and Yengejeh (2019) developed five types of activities: design, datasheet/application note, catalogue/brochure, article-reading, and troubleshooting. Each had its own intended learning outcomes and were situated in authentic scenarios that built in real-world limitations. Students in the authentic activity tutorials scored substantially higher on the final exam than students in the traditional tutorials. Most students had not previously experienced similar activities and wanted more of them, particularly the design and troubleshooting activities. Satisfaction and self-reported learning gains were also high (Shahnia and Yengejeh, 2019). It is not clear whether these benefits mainly came from the authentic nature of the task or through increased interactions with peers and teachers, which the students particularly appreciated.

Way et al (2021) elaborated on the criterion of realism in authentic assessment, emphasising the importance of emotional content, temporal dimensions and unpredictability. These features were specifically built into an online mimetic occupational health and safety management simulation that unfolded as a story over nine weeks, with characters brought to life through emotional video phone

calls to the students at time-sensitive decision points. Gaps and inconsistencies in information, along with unpredictable twists and turns in the narrative also contributed to the realism. Students who experienced the mimetic simulation were compared to the previous cohort who experienced the same course but analysed the material as a historical case. Students who experienced the mimetic simulation watched more lecture videos, made more discussion forum posts, and viewed discussion forums more than the previous cohort who had the static, information-focused simulation. In short, the additional elements of real-world fidelity were associated with greater student engagement in the online module (Way et al, 2021).

3.2.3 Impacts of simulations

Simulations allow students to practice skills and apply knowledge in safe settings, typically on campus rather than in the workplace (Kaider et al, 2017). Thus, simulations are a form of authentic assessment, rather than a work-integrated learning experience as reviewed in section 3.1.2 above. Often, but not always, they are facilitated by technology. We first provide an overview of different kinds of simulations used in different fields before focusing specifically on new research on simulations in health professions and then in business.

Simulations are common in health professions education where they typically involve students role-playing professionals. Standardised patients (SPs) are trained actors who are enlisted to role-play patients in clinical scenarios in which students must demonstrate specific communication and/or practical skills. Technological advances now enable students to interact via various sensors with virtual patients (VPs) embedded in computer-generated 3D environments on virtual reality platforms. VP characteristics and behaviours can be programmed and adjusted along various parameters, enabling students to practice under different conditions (Consorti et al, 2012).

Simulations are also used in business education, though the literature on simulations in healthcare education and in business education have typically developed separately from each other, often published in discipline-specific pedagogical journals. Simulations in healthcare are typically short, stand-alone events of 90 minutes to a half day. In contrast, business simulations can be term-long projects (eg Valenzuela et al, 2018; Way et al, 2021), overlapping in our categorisation scheme with group assessments and authentic assessment more broadly. Objective Structured Clinical Exams (OSCEs) used in health professions education are short simulations that focus on individual assessment, whereas business simulations often emphasise teamwork and group assessment. Computerised, interactive simulation games are also increasingly used over the past five years in business education and are often structured as games with team and individual scores presented as feedback after each round (Brazhkin and Zimmerman, 2019).

The field would benefit from greater cross-fertilisation between these siloed applications of simulations. For example, OSCEs are spreading from medical education to other health professional education fields and spreading geographically (Al-Azzawi et al, 2016). They might have wider applicability in other fields, where they could be adapted to interactions with clients. Likewise, the design of extended ‘live cases’ (like those reviewed in section 3.1.2 above) could be brought into

health professional education to enable students to follow a patient over an extended period of time, allowing them to understand patients, typical sequela, and coordination of ongoing care in a broader context.

Simulations also offer unique feedback opportunities. Feedback may be provided by an observer, but consequential feedback (Quinlan and Pitt, 2021) is also built into the simulations. For example, dental students training with both a traditional Phantom Head simulation and a haptic virtual dental station experienced gains in relevant psychomotor skills after three months of weekly practice (Shahriari-Rad et al, 2017). Haptic devices offer feedback related to the sense of touch that guides students’ psychomotor behaviour. In business simulations, individual role players and the whole teams may get feedback on key performance indicators such as return on investment (Brazhkin and Zimmerman, 2019). The role and use of consequential feedback in simulation research warrants further investigation.

3.2.3.1 Simulations in health professions education

We focus on studies where the assessment and feedback aspects of the simulation are discussed explicitly. We do not report on studies in which simulations were only discussed or evaluated as teaching tools. Nicola-Richmond and colleagues (2016) designed a three-hour simulation session for first year occupational therapy students in Australia in which students practiced manual handling, interviewing, and health and safety documentation under supervision in small groups and with feedback. Students reported significantly higher post-simulation skills in the key, targeted areas from pre-simulation to post-simulation. They also reported reduced anxiety and higher preparation for upcoming work-integrated learning and practical exams (Nicola-Richmond et al, 2016). Phillips et al (2017) reported comparable results for a similar three-hour role-play session with physiotherapy students, also in Australia. Robinson and colleagues (2020) found that Australian university speech pathology students’ performance on a second communication scenario with a virtual patient (VP), as measured by a clinical educator, was significantly higher than their first VP encounter. No comparator condition was tested. Yet, particularly given the costs of VPs, it is important to test their effects relative to other kinds of simulated learning situations such as using role-playing peers or SPs.

Two new studies follow earlier research on VPs that show mixed results on performance (Consorti et al, 2012). In a randomised control trial in Germany, Seifert et al (2019) compared dentistry students’ use of VPs in four scenarios with traditional, interactive small group teaching (SGT) of the same four cases. Both groups of students reported significant gains in relevant self-reported competency from pre- to post-test, though there were no differences between the two groups. VP students made much higher gains on a theoretical test compared to the SGT students, though SGT students caught up with VP students six weeks later, after both had been in clinical training (Seifert et al, 2019). In contrast, Raupach et al (2021) found that another German cohort who had studied emergency medicine using a set of separate VP simulations for 90 minutes weekly over six weeks demonstrated significantly and meaningfully better diagnostic and therapeutic reasoning in the short term than a cohort who had experienced traditional teaching. The treatment group who used VP simulations also performed better than the control group in a follow-up 1.5 years later.

The design of simulation debriefings matter. Paloncy (2020) challenged standard best practice in basic life support simulation training in an American randomised control trial comparing CPR skills of students under two conditions. The control group participated in a standard short, high-fidelity real-time cardiovascular emergency simulation with a standard debrief. The experimental group did the same simulation plus debrief, followed by a 15-minute supervised practice in which a trainer provided corrective feedback on key component skills based on readings from the simulation software. The average performance in the first simulation was 19% for both groups (compared to 75%-100% for advanced performers and 50%-75% for intermediate performers). Two weeks later, students in the control condition performed no better in the simulation than they had the first time (with average scores of 17%). Students in the experimental condition who had also participated in the supervised skills-training practice session scored an impressive 72%. While the International Nursing Association for Clinical Simulation and Learning guidance indicates that either a debriefing or supervised practice should follow a simulation, Paloncy (2020) demonstrated that a debriefing alone was insufficient for students to make adjustments to psychomotor skills such as hand placement, compression depth, timing, or forcefulness of ventilation.

A key question for educators is how realistic a role play or simulation needs to be or how 'high-tech' it needs to be. In the post-simulation training sessions, Paloncy (2020) took advantage of the computerised feedback afforded by a high-fidelity emergency cardiovascular simulation that is not available with the more commonly used low-tech Resusci-Anne, which does not measure actual volume of the mannequin's lungs. This feedback was likely a key element in achieving the profound skill increases she reported with her experimental group. In contrast, Phillips et al (2017) compared treatment conditions in which physiotherapy students interacted with trained standardised patients using real hospital beds to a control condition in which students played the role of the patient using plinths as beds. They found no differences on student performance on a subsequent OSCE assessing the same skills, though students in the standardised patient (SP) condition praised the realism of the SP condition. In South Korea, Park et al (2019) randomly assigned students to either role-play a smoking cessation counselling session with peers or with standardised patients. Students' learning gain from pre- to post-test and short- and long -term retention, were no different between the two conditions, although students in the SP condition reported higher self-confidence after their learning experience (Park et al, 2019). Role-playing peers are a cheaper and simpler option than paying SPs or designing complex VPs. As we discuss under the peer feedback section (section 5), much of the benefit of peer feedback accrues to the peer giving the feedback. It is notable that in the peer role-playing condition (Park et al, 2019), students rotated among three roles over three rounds: doctor, patient and peer observer giving feedback. In the SP condition, students did not participate in giving peer feedback or in role-playing the patient. They practised as doctor over three rounds, receiving feedback from the SP only. Thus, some of the learning benefit from the peers' role-playing conditions may have accrued from students giving peer feedback.

As described in the section on work-integrated learning (section 3.1.2), Valenzuela et al (2018) found that interacting with a real client in a live business case (a form of simulation) led to higher student performance than mocked-up client scenarios. A term-long project involving a real client and a range

of complex skills is quite different from a three-hour session focused on technical, practical skills. Likewise, Way et al (2021) found increased student engagement in a nine-week online narrative simulation with emotional interactions with fictionalised characters, time-pressured decisions, and plot unpredictability compared to analysis of the information of the case presented historically. However, most of Way's students were mature students with professional experience, thus they may have been better able to handle the increased pressure added into their mimetic simulation.

Considering that simulations often contain emotional aspects that may be stressful for students, particularly for younger students who are less likely to have professional experiences (Aamlid and Tveit, 2021; Brazhkin and Zimmerman, 2019), educators need to consider how they build in social support. Social support seems to be a key factor in self-perceived learning in group-based simulations (Thomas et al, 2019). In sum, the importance of realism and work proximity for different kinds of simulations remains an open question and may depend, in part, on the level of experience of the students and how much social support is provided for managing realistic scenarios.

3.2.3.2 Game-based simulations in business

Van Esch et al (2020) reviewed a variety of papers that showed mixed effects of game-based simulations in marketing, in part, it seems, because the experience of 'losing' in game-structured simulations may put students off. They compared a group of students who were randomly assigned to play Marketplace Live, a business simulation widely used in US HE, across six weeks in the second half of term to a control group who experienced the same course but did not participate in the simulation. Students who participated in the simulation experienced greater increases in enthused participation pre- to post-test, but no differences on two other aspects of engagement. There were no significant effects on grade point average (GPA) between the experimental and control conditions, though the number of students was small (n=24 in simulation group). The nature of the assessments used to determine GPA and their alignment with the intended learning outcomes of the game were not discussed.

It is vital that simulations, particularly game-based simulations, be aligned with the desired learning outcomes and that assessment of students' performance on them are also constructively aligned. In a longitudinal qualitative study, Brazhkin and Zimmerman (2019) traced students' learning through a series of five rounds of play with The Fresh Connection team-based simulation. While individual role scores based on role-specific key performance indicators or team scores on team KPIs are often used as part or all of students' grades, their study sought to determine what and how students were learning through the play. They emphasised that assessment needs to reference the desired learning outcomes, including strategic decision-making, appreciation of cause-effect relationships, planning and data analysis, and effective communication within the team, meaningful team collaboration, and reasoned decision-making (Brazhkin and Zimmerman, 2019). Scores generated by the simulation itself do not do so on their own.

Some studies, perhaps where simulation was being introduced in settings or on topics where it was new, simply reported students' satisfaction and improved self-confidence in key skill areas after

simulation learning experiences (eg Omer, 2016; Torda, 2020). More studies comparing simulations to other kinds of activities, with appropriately aligned performance measures, are needed to ensure that these game-based business simulations have positive impacts on students' learning. Such research can also focus on establishing the conditions required to make them effective, such as preparation, scaffolds, debriefs and grading schemes.

3.2.4 Impacts of problem-based learning

Problem-based learning (PBL) is a method that engages students actively in working through problems, often presented as cases contextualised in relevant professional contexts. It was initially developed in medicine, where there are decades of research on its effectiveness. Research during our review period has now spread to other health professions programmes and beyond. The method involves small group interactions with facilitators that enable teacher and peer feedback, as well as study sessions drawing on relevant information that must be applied to the case. Liu and colleagues (2019) conducted a meta-analysis of 34 experimental studies (n=4406 students) comparing PBL to lecture methods in pharmacology education. The results revealed that PBL had a positive effect on scores on theory tests and that PBL methods are superior to lecture-based teaching methods in improving students' self-study, interest in learning, team spirit, problem solving, analysis, scope of knowledge, communication and expression. Given that PBL involves continual performance and feedback through an active learning process, at least some of these gains are attributable to authentic, formative assessment.

It has long been recognised that problem-based learning (PBL) needs to present complex, challenging problems. PBL case formats and processes have been established, refined and tested in health care fields, with demonstrable results (eg Liu et al, 2019). But less evidence exists outside of those fields, where the design of problems and the conditions under which problem-based learning is effective are not yet resolved.

Research is scant during the period of our review, but it has focused on how to scaffold students' engagement with problems and the effects of doing so. For example, Budner (2020) developed integrated problems in chemistry which challenged students to use multiple concepts in a single problem. However, the teacher provided hints about which concepts to apply at various points in the problems and prompted students to systematically review previous concepts. The author also tested this integrated-problems approach in three different contexts: completed during a class (ungraded), providing access outside of class (ungraded), or offering the integrated problems as graded problem sets outside of class. There were no significant differences in performance across conditions, but students working through the problems in class with teacher feedback tended to engage with it as a learning activity. Students doing the problems outside of class found it difficult without teacher assistance and, in the graded condition, reported the mark as the main incentive for completing the problems (Budner, 2020). We discuss the role of grades later in this review (section 4.3).

In Wilkin's (2016) redesign of an assessment task in which accounting students had to apply key professional principles to a business case study, students were prompted to pay attention to concepts

and ways of thinking that had been absent in earlier iterations of the assessment design. The author introduced scaffolding through structured application of a set of principles, followed by class discussion, and found that students reported better engagement with critical thinking and demonstrated better subsequent critical thinking following this scaffolded approach (Wilkin, 2016).

During the period of our review, more evidence has emerged for the effectiveness of problem-based learning in health professions education. Researcher-teachers in other fields are still working on how to design and scaffold problems and the educational conditions in which students will learn best with them.

3.2.5 Impacts of groupwork assessment

Because team working and collaboration are often cited as key skills students will need in their future employment, group assessment can be a form of authentic assessment. While there is a copious literature on peer assessment (see section 5.1), there has been less research over the past five years on the design of group assessments or their impact. Studying the impact of group assessment on student achievement is complicated by the variety of other variables that impact on student achievement, some of which may interact with group assessment. Care needs to be exercised in the design of group work assignments and students' preparation for those tasks, including scaffolding of new skills or attitudes required for success (McGahan et al, 2016; Paschalis, 2017; Taylor et al, 2021). Student perceptions of group assessments, like other assessment tasks, is improved by clarifying the task and its purpose and value (Sedghi and Rushworth, 2017).

Two key Australian studies have demonstrated that student performance tends to be better in group assessments than individual assessments, though student satisfaction is not always enhanced. First, controlling for various confounders, Gamlath (2021) studied 98 modules in an Australian business school, of which 35 had group-based assessments. Student progression was higher in modules with group assessments and the effect was stronger with higher weighting of group assessment. However, they also found that student satisfaction was lower in units with group assessments. Second, Headrick et al (2021) found that, on average, sport and exercise students improved their individual multiple-choice quiz scores after discussing the items with peers. This collaborative testing approach was associated with significantly higher scores on motivation-related variables among first year students (enjoyment, perceived confidence, effort and usefulness) compared to third year students. Third year students found the collaborative discussions more stressful. Qualitative feedback suggested that first year students benefited from learning about different ways to approach such questions, while third year students were already entrenched in their own methods and found it anxiety-provoking to question those methods (Headrick et al, 2021).

Group assessments may also help promote social integration, which is associated with greater retention if teachers assign students to mixed groups. Morley and Ablett (2017) designed and evaluated the impact of assessed group work on a diverse cohort of first year students in an Australian social work programme. Integrating group goals and individual accountability, 70% of the groupwork mark was allocated for self- and peer-assessed individual contribution and 30% for the quality of the

group's final product. This group work experience led to students feeling more connected to their classmates and a greater sense of belonging (Morley and Ablett, 2017). International science students studying in the UK also reported more benefits from multicultural group work than UK domestic science students and were more likely to seek out culturally mixed groups than UK home students were (Sedghi and Rushworth, 2017). Nonetheless, both UK home students and international students valued multicultural groups and agreed that teachers should allocate students to mixed groups. Teacher involvement in group assignment was also associated with higher self-reported performance in American social work classes (Postlethwait, 2016). Group work may not always proceed smoothly across cultural boundaries; students may need scaffolded reflection opportunities to support productive, inclusive, cross-cultural collaboration (Taylor et al, 2021).

Some studies suggest that group work may be an inclusive form of assessment that does not contribute to attainment gaps between majority and minority ethnic students. In the UK, Hill et al (2016) found that assessed group work does not adversely affect UK Black, Asian and minority ethnic (BAME) students. They also did not find an effect of the demographic composition of groups on achievement, though surveyed students tended to see random assignment into groups as fairer than students' own self-selection. Surveyed students in their study valued group work for helping them learn time management and skills in working with others that would contribute to their employability (Hill et al, 2016).

While teachers often balk at group grades out of concern that students may find it unfair or be less satisfied, such fears may be unfounded when assessment is well-designed. For example, Leung et al (2017) redesigned Canadian physical science lab exercises so that only one randomly selected student from a group of four was assessed on the lab skill and all students in the group were given the score obtained by the randomly selected student. There were no differences in student scores in the experimental versus control groups, though there were more perfect scores in the experimental group (Leung et al, 2017). Students in the experimental condition rated the assessment significantly higher on fairness (Leung et al, 2017). Peer instruction and peer learning also appeared to be higher when one student was randomly selected to represent the group on the skills assessment, though this trend was not significant (Leung et al, 2017). Clearly communicated conditions, enhanced peer instruction, random assignment of groups and random selection of students may have enhanced the perception of fairness. Marking duties were reduced by 75% in the experimental condition, suggesting the utility of such an approach on a wider scale (Leung et al, 2017).

It is vital to design group assessments that ensure individual students are held accountable for their contributions. A national survey of architecture students in Australia found that students were generally satisfied with their group work and appreciated that it helped prepare them for professional practice and develop social and interpersonal skills. They were most dissatisfied with unequal workload contributions and assessments that did not hold individual students accountable for their contributions (Tucker and Abbasi, 2016). Likewise, unequal division of labour was related to group conflict and lower student self-reported performance and learning among social work students in a smaller study (Postlethwait, 2016). These findings are consistent with research over decades on the

effective design of group work in education, which show that students need to both be working toward group goals and be individually accountable (Slavin, 2010). That is, providing an evaluation tool to identify individual contributions to group work is also important (Sedghi and Rushworth, 2017).

Students also benefit from instruction on and scaffolding of collaboration processes. Paschalis (2017) designed scaffolded processes of collaboration that were supported by a learning activity management system (LAMS) embedded in Moodle in a Greek computer science course. With their six-week project divided into phases, including specific instructions for subtasks and tools for collaboration, students performed better than a previous cohort who didn't have those supports (Paschalis, 2017). The research studies cited above generally support long-established key principles of effective group assessments.

Other studies provided detailed case studies of how to integrate novel group work projects and assessments. For example, Neal et al (2018) described the design of a 'Medieval Expo' for history students in Australia to learn how to work in teams to communicate in creative ways to non-specialist audiences. Task authenticity, collaboration, creativity and novelty of the tasks contributed to their effectiveness, as did teacher group formation, and a combination of group and individual accountability built into the marking structure (Neal et al, 2018). Taheri et al (2019) provided a useful summary of the history of Makerspace, particularly in engineering education, offering examples of group project designs that make use of such spaces and students' responses to those learning opportunities.

In sum, group projects can enhance students' collaboration skills and performance, and may help reduce attainment gaps. To achieve these benefits, though, students need random or teacher-assigned groups, scaffolding of collaboration and intercultural competence, group goals and individual accountability.

3.2.6 Impacts of portfolios and e-portfolios

E-portfolios are digitally mediated, learner-centred, intentional collections of work intended to capture sophisticated, multi-dimensional achievement (Cummings and Maddux, 2010). Traditionally, students autonomously select from work done over an extended period, on which they have had teacher feedback and, possibly, peer feedback. When fully implemented, portfolios can qualify as meeting the eight elements of high impact practice described in 3.1 above. Because of limitations in the way in which they are currently rolled out as e-portfolios, we have, instead, categorised them under authentic assessments, focusing on their key affordances of realism (authenticity), higher order thinking, and evaluative judgement (Villarroel et al, 2018).

E-portfolios are a possible solution to problems associated with standard grading practices which reduce various levels of accomplishment of multiple, complex learning outcomes into a single grade (eg Schwab et al, 2018; Jorre de St Jorre, Boud and Johnson, 2021). Instead of just a grade, an e-portfolio offers curated work samples illustrating different competencies that can be shared with wider audiences such as prospective employers. E-portfolios are gaining traction in Hong Kong

(Deneen et al, 2018) and Malaysia (Abd-Wahab, 2016) and in specific disciplines with professional associations, such as engineering (see discussion under competency-based learning in 3.3 below).

Research and practice on e-portfolios in this five-year period still appears to be at an initial stage, focusing mainly on student perceptions of what an e-portfolio should contain (Abd-Wahab, 2016) and students' acceptance of and attitudes toward e-portfolios both as an assessment vehicle and an educational technology (Deneen et al, 2018). It appears that in some cases a technological platform has been introduced, but academics are still learning how to make best use of its pedagogical affordances. Practice with and research on e-portfolios also seems to be complicated by differing interpretation of their role and purpose. Are they seen as final products that showcase student accomplishments, or are they assessed as a work-in-progress, emphasising the quality of student's reflection? Relatedly, implementing them at module level is quite different than at a programme level, thus there are 'grain size' issues that affect the degree of sustained effort normally expected of a high impact practice. Finally, there were discussions around the potential affordances versus burdens of learning to use multi-media representations in students' work.

Unsurprisingly, then, there are different implementations of e-portfolios, which makes it difficult to evaluate their impact on students. After their university's rollout of the Mahara e-portfolio platform in Hong Kong, Yang, Tai and Lim (2016) analysed students' qualitative reports of the use of e-portfolios in three different first year courses and found that key elements of high-impact practices and even of authentic assessments were largely lacking. One or more of these affordances were also missing in most of the small studies found in this literature review, though regular teacher feedback – one of eight key elements of high- impact practices – was most commonly reported. Only one study, in teacher education, focused on the reflective aspects of portfolios (ie building evaluative judgment) (Alzouebi, 2020). They emphasised the process of selecting artifacts and assignments and accompanying reflective narratives (Alzouebi, 2020). Students scored well on a rubric with ratings of appropriateness of artifact selection, adequacy of reflective writing, appropriate standards, and use of appropriate format, software, and mechanics. Students were also satisfied with the assessment (Alzouebi, 2020).

There were also instances in the literature in which some key features of portfolios were present, but the assessments were not referred to as portfolios. For example, Gozzard and Zadnik (2021) introduced a diary worth 5% of students' final marks into an Australian introductory astronomy course which students made regular astronomical observations. These diaries offered an authentic task, observations plus reflections, as well as autonomy in choosing what and when to observe. The format of presentation and the quality of work evidenced considerable effort and engagement. Students in the course with the observations had a larger gain in astronomical understanding and attitudes from pre- to post-course than students in another section without them (Gozzard and Zadnik, 2021).

In a UK example, Christie and Morris (2021) conducted focus groups with 16 students in four different types of disciplines about the effects of blogging on student engagement. They found that the structure of blogging contributed to students' engagement in the courses. Although students assembled their blogs into "portfolios", there was no discussion of any reflection or self-assessment

required in selecting elements for inclusion in the portfolio. Wakeling, Aldred and Hains-Wesson (2018) used a similar activity in which 21 first year food science students in Australia wrote 350 words each week in an e-portfolio platform and received feedback. Students' responses were mixed, though most appreciated the opportunity to encounter new knowledge about food science, the variety of topics they could address, the opportunity to consolidate their ideas about the subject, and regular feedback from instructors. Again, no mention was made of reflection on the corpus of work, suggesting that many initiatives under the label of 'e-portfolios' or 'portfolios' mainly refer to either the technological capabilities of linking to other sources or including multimedia files or simply as a repository for student work in a situation of continuous assessment. Blogs and these sorts of portfolio entries were compared by some students to discussion boards, which we discuss in section 3.5.1 below as examples of continuous assessment.

Mohamadi (2018) investigated the impact of three writing assessment designs on the writing skills of 130 Iranian EFL students: a scaffolded individual writing task; individual portfolio writing with substantial self-assessment of their work; and collaborative writing exercises with a peer. He conducted pre- and post-tests after each segment of the course that used these different conditions. All designs led to significant increases in writing skills, but the portfolio-writing and collaborative writing tasks had successively larger effect sizes, though the sequence (scaffolded up to individual self-assessed work to collaborative work) may have led to their greater effectiveness. These different assessment designs incorporated – to varying degrees in each – some of the evidence-based key educational principles discussed thus far, including: realism (authenticity), higher order thinking and evaluative judgement. Collaboration with peers was also introduced in the final condition. Self- and peer-assessment were particularly powerful. Those will be discussed in greater detail in section 5 below. This literature on portfolios and e-portfolios highlights the need to clarify the nature of their implementation in terms of its key educational principles and elements. The effectiveness of this assessment practice, like many others, likely rests on the extent to which evidence-based principles and elements (of high-impact practice or, minimally, authentic assessment) are incorporated in any given implementation, rather than use of the technology.

3.3 Timing: competency-based learning

Competency-based learning (CBL) focuses on clearly defined and communicated technical and professional competencies. This educational approach presents an alternative to traditional assessments insofar as students are not permitted to progress from one unit to the next unit until a certain level of mastery is achieved. Thus, timing of the assessment is a critical defining feature. Henri, Johnson and Nepal (2017) conducted a systematic review of 60 articles focused on competency-based learning (CBL) in engineering published between 2005 and 2015, in part to determine the assessment methods used in CBL.

Interest has risen in CBL in engineering in many countries in the past 35 years because it suits adult learners who may enter and exit at any time of the year and are balancing education and work. It also suits employers who want to be assured of achievement of key competencies, and it is facilitated by

newer technologies. Henri et al (2017) found a variety of assessment methods being used, often in combination, and noted that little work had been done to identify best practices in assessment.

Traditional pencil and paper quizzes, exams and lab exams were the most commonly used assessments in CBL (Henri, Johnson and Nepal, 2017). Group work, such as projects and presentations, were also common. Portfolios, 360-degree assessments in which students receive feedback from multiple sources, and surveys and questionnaires were also referenced across the papers. Their review suggested that while CBL assessment methods are well suited to technical skills and knowledge, the field has not yet addressed the challenges of assessing professional competencies such as group working and communication. Surveys and questionnaires were mentioned as tools for assessing professional competencies, but it appears that more work needs to be done in this area.

Because CBL tends to compartmentalise learning into discrete units, educators need to give further consideration to assessments that ensure students can synthesise and integrate multiple smaller units into a 'big picture'. Only five of the papers Henri et al (2017) reviewed discussed portfolios, which were used to enhance students' motivation, engagement, ability to self-reflect, iterative enhancement of work, and to assess their communication skills in representing their own work. Electronic portfolios enabled students to showcase work in a variety of media (see section 3.2.4 above for more on portfolios and e-portfolios).

Six of the papers Henri et al (2017) reviewed referred to 360-degree assessments in which feedback from multiple sources (self, peers, experts, teachers) was integrated. This method appeared to enhance students' learning outcomes and promote professional skills and attitudes. It was used in combination with other assessment methods and addressed both technical skills/knowledge and professional skills/attitudes.

Online assessments are particularly valuable in the self-paced CBL environment, and time-efficient and sustainable. They were used for pre- and post-testing and included a variety of formats, including simulations and games that tested specified competencies. Some of the individual studies in our review offered case examples of how assessments were revised in competency-based learning (eg El-Maaddawy and Deneen, 2017; Vidiela and Garcia, 2016).

More work is needed to identify best practices in the special case of competency-based assessment. It will be particularly challenging to bridge demand for assessments of atomised, 'bite-sized' competencies in strict progressions with the need for students to engage in complex, authentic practices that integrate multiple competencies over sustained periods of time.

3.4 Timing: continuous assessment

The Bologna Process emphasised the importance of continuous assessment (CA), prompting many changes in curricula across Europe over the past 15 years and associated teaching and instructional development changes (Quinlan and Berndtson, 2013). Attention to continuous assessment is also seen in other parts of the world during the period of the review, including Ethiopia (Sarka et al, 2017).

Studies classified as 'continuous assessment' focus mainly on the timing of assessments – as occurring early and/or continuously through the course. Typically, these assessments contributed marks toward a final grade (ie serving a summative purpose), but because of their timing, they also communicated information back to students about their performance (ie a formative purpose). Thus, there is some overlap with the sections below on formative assessment (section 3.6) and summative assessment (section 3.7).

Sarka et al (2017) demonstrated the importance of preparing students for new forms of assessment. Forty Ethiopian students were surveyed prior to a training intervention and more than half believed they would perform poorly on a variety of continuous performances, but after the training intervention and experience with a range of assessments (individual assignments, group assignment, presentations, exams), most students believed they would perform well on those assessments. Sarka and colleagues' study, though small, reminds us of the importance of helping students adjust to new assessment demands and build their confidence and awareness. Helping students prepare for a different assessment culture and approach is a general rule of thumb across any new form of assessment and is reflected in research on various kinds of transitions (eg Carter, 2020; Rovagnati, Pitt and Winstone, 2021). Early diagnostic assessments can also be used strategically to help students orient to new standards, particularly when accompanied by clear criteria about expectations and requirements and personalised, constructive feedback (eg Palmer et al, 2018).

Unfortunately, available studies were generally not designed in a way that allows us to draw conclusions about the impact of continuous assessment on students' performance. Most of the studies (eg Pearce, 2018) suggested that students valued CA opportunities to motivate study throughout the course and that, predictably, 'practice makes perfect', though that was not always the case (eg Bruwer and Ontong, 2020; Akhtar et al, 2020). However, when students self-selected into experimental versus control groups; when students in experimental groups had CA prompts, practice and feedback compared with students in a control group with no intervention whatsoever; or when comparison groups had different outcome assessments, it is not possible to discern the underlying mechanism or process that led to observed outcomes. Achieving study designs that allow causal inferences is imperative, possibly by comparing a pre-revision cohort to a reformed course cohort.

One study that shed doubt on the value of CA was Reboredo (2017), who studied the same economics course taught over six years (n= 589) to Spanish economics students and found that CA grades were a poor indicator of final exam grades. They pointed to the importance of how students interpret their CA grades, students' perceived likelihood of achieving a better mark, and the number of points already earned in the course. They depicted students as rational, strategic actors in allocating scarce study time resources in the context of competing study demands. Although they did not reference the significant research on the movetoward students' use of feedback, their interpretation of their finding echoes that trend. See the discussion of feedback in section 4.

Similarly, in a qualitative study of eight students who had failed and repeated an introductory calculus module, Dibbs (2018) found that students' engagement with formative feedback results to improve their understanding, their study skills, self-assessment, and self-efficacy in the subject was critical to

their success. Importantly, teachers' perceived caring seemed to initially prompt students' engagement with the formative assessments before they began to embrace assessments independently (Dibbs, 2018). As with the substantive literature on feedback addressed in section 4.5 below, these findings highlight the importance of how and why students engage affectively, cognitively and behaviourally with continuous assessments that provide formative feedback, not just its presence, absence or design features.

Research prior to the review period suggested that effects of any assessment system on students' learning processes and learning outcomes are mediated by their perceptions of this assessment system (Struyven, Dochy and Janssens, 2005; Lizzio and Wilson, 2013). Students' perceptions of assessment systems were outside the scope of our review. Nonetheless, we note that only a few of the studies that aimed to consider impacts on satisfaction, engagement or performance deliberately focused on students' perceptions as a mediating variable. For example, Hirst (2016) explored students' perceptions of a Pecha Kucha formative assessment activity, concluding that students may have seen it more as surveillance and extra work than as a learning opportunity. How students perceive an assessment should be considered in future evaluative research on the outcomes of specific assessment designs.

Vaessen et al (2017) centred issues of student perceptions and that students' perceptions need to be considered more systematically in the assessment literature. They demonstrated that any assessment activity can be perceived in various ways by students, and those perceptions lead to differential effects on students' motivation and learning strategies. They developed a questionnaire that focused on four dimensions of students' perceptions of continuous formative assessment: 1) value (eg useful in motivating regular study); 2) negative effects such as stress and lower self-confidence in the face of low marks; 3) positive effects such as lower stress for the final exam and higher self-confidence when receiving good marks; 4) formative function such as identifying areas in which they did less well to focus their study. The questionnaire would benefit from further development grounded more clearly in key theoretical constructs and tested on a larger group, but it provides a useful early step toward more robust, multi-item measures in a field that is dominated by single item questionnaires. Students' negative perceptions of formative grades explained (mediated between) grades and students' motivation. In turn, students' intrinsic motivation predicted students' final grades even when controlling for formative assessment grades (Vaessen et al, 2017). The other aspects of students' perceptions of regular graded individual formative assessments did not have an effect on either motivation or final grades. While this study did not refer to the substantial body of psychological literature showing that stress and lower self-efficacy and self-concept all can have negative effects on academic achievement, its key findings are consistent with that research. That is, assessment designs and feedback associated with formative assessments can trigger either positive or negative emotional responses which affect learning processes and outcomes. These emotional responses have effects on final grades that are above and beyond the grades on the formative assessments. These effects on emotions and self-confidence, though, may vary depending on the level of intrinsic motivation or interest in the subject that students come in with (Vaessen et al, 2017). See also section 4.6 on emotions and feedback.

In sum, this section has focused on the timing of assessments and the potential impacts on students of receiving marks and feedback earlier – and, potentially, continuously – in a course. Rather than providing a clear picture of the effects of continuous assessment on students, research over the past five years raises questions about how students perceive the assessment system, interpret their grades to 'game the system', and emotional responses to early feedback, which affect subsequent performance. These questions go beyond just whether students are satisfied or not with the assessment system, but to how they understand the value and purposes of assessments in the context of their own goals and busy lives.

3.4.1 Discussion boards, blogs and micro-blogging

Online, asynchronous group discussions are typically used to promote student engagement, satisfaction and the development of student learning communities, particularly in online and hybrid courses. Effective discussion boards engage students regularly in posting and responding to fellow students' posts with teachers actively involved in querying and stretching students. By nature, then, they are a form of continuous assessment.

Teachers often wonder, though, 'how much is too much?' In a controlled study with 524 mostly adult (>25 years old) students in accelerated online courses, Selhorst et al (2017) found that reducing the requirements from two discussion board postings to one each week appeared to reduce course withdrawal without harming student achievement if discussions were not assigned a large weighting on grades. Similar research with younger students in longer online courses is needed.

Blogs offer a richer Web 2.0 technology than most discussion boards, enabling students to use various audio-visual materials, interact with each and showcase their work more broadly than only to a teacher. Blogging has become popular, particularly in language learning, as it enables students to practice writing in an authentic context. The findings of several small scale or qualitative studies without controls suggested that task authenticity, wider audiences, continuous assessment with opportunities to look back on their work, student autonomy/creativity/personalisation, clear assessment criteria with exemplars, and the creation of a learning community through reading and commenting on peers' work may be key to students' satisfaction when implementing blogging (Hung and Huang, 2016; O'Brien and Freund, 2018; Christie and Morris, 2021; Morris et al, 2019) and microblogging (Santoveña Casal, 2019).

3.5 Purposes: formative assessment

Popham (2011, 270) defined formative assessment as "a planned process in which assessment-elicited evidence of student's' status is used by teachers to adjust their ongoing instructional procedure or by students to adjust their current learning tactics." Formative assessment can provide students with specific, just-in-time information that they can use to improve their performance.

Over the past five years, as seen under several other topics, there was considerable interest in how technologies can be used in the instructional setting. For example, several studies considered the use, and effects, of personal response systems, audience response systems, or classroom response

systems (CRS) (using a variety of platforms including Socrative, Plickers and Mentimeter). These technologies may involve dedicated devices ('clickers') or rely on students using their own mobile devices. Using these technologies, teachers typically pose students a question, which they answer individually and anonymously. Data is collected and displayed in real time, providing formative assessment data that helps teachers to provide just-in-time feedback, such as clarifying a concept where there is considerable misunderstanding. Students also see their knowledge gaps. Typically, questions are multiple choice questions, but some systems, such as Socrative or Google Slides, allow students to generate their own content either individually or via collaboration (eg Paul et al, 2020).

Elmahdi et al (2018) found that 166 education students of different levels in Bahrain, in different classes, who experienced Plickers (a paper-based CRS system) three times during a term reported that they were highly satisfied with the tool; it positively affected their engagement; allowed checking for understanding, equal opportunities to participate and a savings in learning time; felt exciting and fun and broke the routine; and was easy to use. The students regarded the various uses and purposes of formative assessment to be very important. Mshayisa (2020) found similar responses on a survey of 85 food science students in South Africa, though he also found that students particularly valued the element of anonymity. Some studies, especially those in the Global South where internet access can be unreliable, focused more on how to create formative assessment opportunities using e-learning approaches that draw on more commonly available mobile phone technologies (Mir et al, 2019).

Not all studies of satisfaction with formative assessment focused on new technologies. For example, some focused on describing and documenting students' responses to well- designed formative assessments that incorporated key design features, such as clarifying purposes and criteria and providing peer, self and teacher feedback or providing early induction into key skills (eg Naghdipour, 2017; Nkealah, 2019; Cooke, 2016). Others emphasised student satisfaction with activities that promoted conceptual understanding, such as creative exercises in chemistry (Giwelski et al, 2019) or concept maps in mathematics (Hammad, 2020) or that prompted students to engage with readings (Heinicke et al, 2017).

A challenge with interpreting research on effects of formative assessment on performance is that, even when randomising, samples may be small and many studies do not involve any controls (eg Naghdipour, 2017; Nkealah, 2019; Palmer et al, 2018) or, if they do, it may not be a proper control. Without a true control, enhanced performance for those in treatment conditions may simply be due to already well-established practice and testing effects (eg Amasha et al, 2018; Alahmadi et al, 2019; Hammad, 2020; Jeong et al, 2020; Pearce, 2018; Umar, 2018; Davis et al, 2020). Nonetheless, practice and testing effects alone are important educationally, particularly if students value the approach. See Roediger and Karpicke (2006) for a discussion of how to leverage the potential educational benefits of the testing effect, in which practice at retrieving information strengthens memory for it.

Studies on the effects of testing and frequency of testing on simple recognition and recall tasks are usually done in laboratories. However, Foss and Pirozzolo (2017) conducted a series of studies in HE classrooms with controlled differences between conditions. These studies demonstrated strong

Impacts of Higher Education Assessment and Feedback Policy and Practice on Students: A Review of the Literature 2016-2021
Edd Pitt and Kathleen M Quinlan

testing and frequency effects on lower order thinking tasks, even when items on the same topic were worded differently on the first and subsequent tests. Other classroom studies, such as those reviewed below, provide comparable controls in which the questions focus on the design of more complex (higher order) assessment tasks, with different kinds of designs or foci as controls. These quasi-experimental designs allow us to tease out the key assessment design elements that contributed to success.

Adams (2020) reported on the effects of redesigning UK chemistry laboratory assessments on students' engagement and satisfaction as compared to the prior year's cohort. Rather than assessing students' experimental design skills, Adams's (2020) redesign clearly focused on practical skills through stated learning outcomes and aligned criteria. He provided the cohort of first year students with 15 weeks of formatively assessed labs, followed by a choice of submitting marks from four out of six summatively assessed labs, also called 'selective assessment' (Sobhanzadeh and Zizler, 2021). The percentage of students viewing the online videos demonstrating lab procedures nearly doubled (from 45% to 97%). Student satisfaction with key elements of the lab was also higher and clarity about the purpose of laboratories also rose substantially from the previous year.

In another chemistry study, also focused on creating constructively aligned laboratories, Young et al (2016) designed a course for non-chemistry majors in the US (eg health profession students) that alternated six to seven wet labs (using liquids, biological matter or chemicals) with six computer-based self-guided inquiry activities focused on critical reasoning, data interpretation and report writing. Students viewed technique videos and completed quizzes each week. Students were satisfied with the alternating activities. Their performance on practical skills, report writing and the overall exam were no different from other cohorts, which had the full semester's worth of traditional weekly labs (Young et al, 2016). As labs require considerable staff time and materials, using computer-based alternatives can be cost-effective.

Taylor et al's (2017) study offered controlled studies in two different settings that showed significant effects on students' conceptual understanding in biochemistry after engaging in 26 pre-class concept review sheets, in combination with in-class reinforcement activities (one of which involved clicker questions and discussions). Pre-class review activities alone did not show significant effects over either control conditions or null conditions, while the combination of pre-class plus either of two different types of in-class reinforcements did have an effect. They reported a trend (though insignificant due to small sample sizes) toward long-term learning gains on an exam eight months later for one of their experimental conditions.

In a randomised control trial in a large introductory chemistry course in the US, Scalise et al (2018) found that the nature of the formative assessments made a difference to students' performance. Treatment groups receiving formative assessments during the first half of the course that offered tests and feedback related to (i) conceptual support or (ii) conceptual support with differentiated instruction showed a statistically significant gain in learning outcomes over the comparison group that took part in take home exercises provided by a standard textbook publisher. The treatment and comparison groups were exchanged half-way through the semester. On average, the original treatment group continued to outperform the original comparison group in the second half of the semester.

These findings suggest early conceptual support may have sustained impact on understanding over time. Likewise, in introductory mathematics, conceptual support in the form of three spaced individual concept mapping activities coupled with collaborative discussions and teacher feedback was associated with increases in students' conceptual understanding and lower failure rates for the course overall compared to students in other sections, though the characteristics of the comparison group were not clear (Hammad, 2020).

Racially minoritised, at-risk, first-generation-to-attend university students may have specific needs that can be addressed through well-designed formative assessment. Carmichael and colleagues (2016) offered one of the few studies that considered how assessment and feedback designs might close attainment gaps between advantaged and disadvantaged students. They made locally evidence-based instructional and assessment changes to an introductory biology course over nine years. They added regular clicker-based activities in classes, increased classroom interaction, designed review sessions based on clicker data-mining, engaged students in self-correcting their work and structured self-assessment, and graded students' regular take-home assignments. Over that time, significant performance improvements of at-risk students were noted on a core set of the same 60 multiple choice items.

In the introduction to section 3, we referred to Black and Wiliam (2018)'s five key elements of effective formative assessment. Looking at the evidence from the past five years, we see these general principles reinforced. Simply providing opportunities for formative assessment and feedback does not appear to be sufficient; other elements must be in place. Because much of the value of formative assessment lies in the feedback provided to students, the design of formative assessments and the design of feedback should be considered in tandem. Section 4 addresses effective feedback designs.

3.5.1 Automated assessments

Interest in technology-enhanced continuous or formative assessment, discussed at various points above, has also included automated assessments. Some studies focused primarily on using multiple choice question (MCQ) functions of learning management systems, such as Moodle, as a way of incorporating more frequent, continuous assessments without increasing demands on staff time or sacrificing student satisfaction (eg Babo et al 2020; Lopez et al, 2016). Others used low weight, graded (less than 20% of overall marks), regular weekly quizzes to successfully prompt students to do preparatory work to enhance student engagement and understanding (Cook and Babon, 2017; Herbert et al, 2019; Pelkola et al, 2018; Young et al, 2016; Heinicke et al, 2017). Such tests are also called 'Readiness Assessment Tests' (Heinicke et al, 2017). Some teacher-researchers aimed specifically to create MCQs that engaged higher order student thinking, such as by using more complex stimulus (stems) like graphs, case examples or academic articles (Davis et al, 2020; Hancock et al, 2018; Herbert et al, 2019). Creating MCQs that prompt higher order or deep approaches to learning is important to promote better outcomes (Svirko and Mellanby, 2017). Some studies, conducted in low resource settings, highlighted concerns with connectivity or access to computers or computing spaces as a potential drawback with submitting online assignments (Padayachee et al, 2018; Reju and Jita, 2020).

Impacts of Higher Education Assessment and Feedback Policy and Practice on Students: A Review of the Literature 2016-2021
Edd Pitt and Kathleen M Quinlan

Immediate feedback available through automated assessments may help promote students' engagement. Herbert et al (2019) devised specialised systems for automatic marking in mathematics that prompted students to engage in problem-solving (not just responding to MCQs). To prevent plagiarism, the authors also designed multiple iterations of the same assessment and randomised question sequences. Like other studies reviewed here, their design sought to encourage students to engage with the content prior to attempting a relatively high-weighted assessment. They found that 94% of the students completed the new formative assessment with automated, immediate feedback compared to just 15% who had completed unmarked MCQs in the same course the year before. Students agreed that the assessment prompted them to study throughout the course. They also engaged more with questions and comments in online forums.

In addition to immediate feedback, automated assessments can offer students multiple attempts, which may support their longer-term performance. Buitrago-Flórez et al (2020) designed an automated system that enabled engineering students in Colombia to receive feedback during each step of a problem in an individual assignment and multiple attempts at the assignment. Those students did better on that assignment and the overall final assessment than a previous cohort that only did traditional pencil and paper assignments with delayed feedback. Students in the treatment group also reported high levels of satisfaction with the system, an average of three attempts, and high engagement with personalised assistance during assignment completion. Students cited multiple attempts and immediate feedback as particularly valuable (Buitrago-Flórez et al, 2020). Likewise, Davis et al (2020) showed in a quasi-experimental study with business students in the US that students in an unlimited-attempt quiz situation with immediate feedback performed better on subsequent exams than those with limited attempts.

However, immediate feedback, rather than multiple attempts, may be the key mechanism generating performance improvement. See section 4.7.1 on feedback timing for more on immediate feedback. Tila and Levy (2020) sought to separate the effects of multiple submission from immediate feedback. They found that economics students in a US community college improved their scores over successive attempts, even when they did not receive immediate feedback on accuracy. However, students did not score higher on multiple submission assignments than single submission assignments, nor were there treatment effects on subsequent assessments. Most students made use of the multiple submission opportunity, submitting an average of 2.89 times, and were very satisfied with the opportunity, believing that they did better on the multiple submission assignments, and that it reduced stress (Tila and Levy, 2020).

It appears that some students use multiple submission opportunities as a substitute for other study strategies, weakening the intervention's impact on subsequent performance. Tila and Levy's (2020) participants reported reviewing and studying further to identify their own errors before resubmitting. That is, they seemed to use the resubmission instead of other study strategies they might otherwise have used before doing a single-submission assignment.

Some students also appear to access resources that are not available during final exams when completing formative quizzes, which reduces the correlation between formative quizzes and final

exam performance. Pelkola et al (2018) used STACK in mathematics to create weekly online mastery quizzes designed to encourage students to retake versions of the same exercises until they reached at least 80% mastery. Students managed to raise their scores from initial attempts to eventual mastery on the weekly quizzes from an average of 12% to an average of 88% correct. However, students' final mastery on those exercises was only weakly correlated with their performance in subsequent formal tests. This discrepancy may be because some students reported using calculators, books or other tools to complete the quizzes, although this practice was discouraged. Although Lopez et al (2016) demonstrated clear positive correlations between Spanish physics students' academic performance in online formative tests and their final marks, their study of outliers is revealing. Some students performed well on the Moodle tests, but failed the final exam, which suggests the need for further inquiry into students' use of Moodle tests compared to traditional, closed-book, timed exams and the longer-term impacts of different student behaviours related to formative assessment.

Some papers described in detail how specific technologies can be adapted to accommodate the challenges and needs of specific disciplines. For example, to assess students' plant identification in a botany course, Conejo et al (2016) developed an extension of the Siette assessment framework that uses geographical positioning and/or QR codes to trigger questions attached to a specific location. Mainly, the system simulated individualised, immediate teacher feedback in the field for students working independently asynchronously. Students who used the system demonstrated better performance than students using traditional forms of formative assessment, and they self-reported higher learning gains. Allagui (2021) used augmented reality to create dynamic visual writing prompts that generated better EFL descriptive writing than a traditional writing prompt. In mathematics, both Herbert et al (2019) and Pelkola et al (2018) used technology in innovative ways to address subject-specific challenges.

In sum, students appear to engage well with formative automated assessments and are satisfied with it, particularly if there is a small amount of credit provided for completion and it offers multiple submission opportunities. Readiness assessment tests appear to help students stay on top of material during the term and come to class better prepared for discussions and activities. However, effects of formative quizzes on subsequent exam performance are mixed. Immediate feedback may be a mechanism that enhances performance. Multiple submissions that encourage mastery learning may lead to higher subsequent exam performance. However, student's overall performance may not be improved if students take quizzes in conditions that are unlike subsequent exams or if they simply use practice with multiple quizzes in place of other effective study strategies.

3.5.2 Gamified learning: adjusting traditional assessments

Gamified learning adapts traditional learning or assessment processes to include game-like elements such as collaborative-competitive structures with points accruing on leader boards to teams that compete against other teams. Gamification is theorised to promote learning behaviours that are associated with better learning outcomes (Landers, 2014). For example, students who are having fun may work harder, which promotes their learning. Gamified learning shares some similarities with but differs from a full-fledged, serious game (Sailer and Homner, 2020). Serious games and simulation

games tend to be more authentic, thus we have dealt with serious game-based learning in section 3.2.3 on simulations.

In gamified learning, there are various quiz programmes such as Kahoot! and Quizalize that have some elements in common with personal response systems (see section 3.5.1 on automated assessment) but contain game-like elements. This power of this kind of gamified learning may lie in immediate task-based feedback, as we discuss in section 4.7.1 below.

A meta-analysis of 19 different gamified learning experiments across various educational levels with a total of 1686 students (Sailer and Homner, 2020) found small, significant effects of gamification on cognitive, motivational and behavioural learning outcomes. However, when considering only those studies that met the highest standards of methodological rigor, only cognitive learning outcomes were demonstrated. Gamification may be effective for motivational and behavioural learning outcomes when it is applied in competitive-collaborative settings, in contrast to situations with no interaction among students (Sailer and Homner, 2020). That is, the social element may be a key mechanism promoting motivation and learning.

Evidence is still emerging, as many early studies suffered methodological flaws that have resulted in conflicting results over a series of systematic reviews prior to Sailer and Homner's 2020 meta-analysis. There are also many moderating variables associated with the instructional design (eg whether games include a narrative component) and settings (in class versus out of class; as a flipped design or not; time period over which gamification was tested) that need to be taken into account in future studies, as well as studies that include differences in learner characteristics and the extent to which students actually engaged with the affordances of the designs.

The meta-analytic findings were supported by several robust empirical studies uncovered in our search. For example, in a control group study in the United Arab Emirates, Eltahir et al (2020) found that using the game Kahoot! as an in-class formative assessment with students studying Arabic grammar yielded higher student motivation, engagement and achievement than a similar paper-based test. In Spain, Guardia et al (2018) reported higher student performance and self-reported student skills with collaborative-competitive structured (eg points and team leader boards) Kahoot! activities than traditional MCQ tests. In Germany, Sailer and Sailer (2021) found positive learning outcomes and positive effects on motivation using Quizalize, though the novelty of the activities may have been a contributing factor. Wilkinson et al (2020) found improvements in grades among UK students who used a 15-minute gamified quiz just before an anatomy assessment compared to students who used their usual revision methods. Students cited immediate feedback as one of the advantages.

Technological advances make it easier to add gamification into traditional assessment activities such as quizzes. Doing so is likely to yield small benefits to students, particularly in terms of knowledge gains. However, the type of gamified quizzes described above may not address higher order thinking skills. They also do not necessarily transform assessment practice to the extent that high-impact practices or authentic assessments do, where gains in student experience, engagement and performance may be larger and more meaningful.

3.6 Purposes: summative assessment

Much of the emphasis in the field over the past five years has been on formative assessment and the ways in which both assessment and feedback design can promote student learning, including large-scale studies in new country contexts (eg China) that replicate previous findings elsewhere about the value of feedback (Guo and Shi, 2016). Nonetheless, summative assessment remains an important role for institutions in certifying competence to external stakeholders such as employers or admissions panels for advanced study.

To fulfil that communicative role, summative assessment needs to be designed to accurately predict future performance, particularly in high stakes professions. In a systematic review of 18 studies in health professions education, Terry et al (2017) found that most studies (15) used objective structured clinical assessment, 12 of which showed significant positive relationship with clinical performance. However, the amount of variance explained varied considerably (1.9% to 39.7%), as did the structure of the OSCEs themselves, ranging from three stations to 35 stations. Larger numbers of stations were typically associated with better reliability (Brannick et al, 2011) and were also associated with better predictive validity in the Terry et al (2017) review. Terry et al (2017) also reviewed four studies on written examinations, finding little support for their use as predictors of clinical performance. Other professions may benefit from similar studies about how well different assessment types predict students' professional performance.

3.6.1 Constructive alignment and equivalencies across formats

Emphasising constructive alignment (Biggs, 2002) between learning outcomes and assessment has led to greater attention in the literature to specifying higher order learning outcomes and creating more authentic assessments, though this ideal has not always been translated into practice. For example, Stowe and Cooper (2017) found that 93% of a sample of organic chemistry exam questions from four elite US universities were not aligned with authentic scientific or engineering practices, despite rhetoric used to promote the courses. They offered guidance about how exam questions could be improved to gather evidence of students' engagement with these key practices.

A greater focus on learning outcomes allows programmes to map assessments against learning outcomes to consider assessment across an entire programme. Mapping assessments across a programme ensures that students are not under- or over-assessed. It also enables the tracking of students' progress across a programme on learning outcomes (see also section 3.3. on competency-based assessments).

Assessing against explicit learning outcomes also supports flexible and inclusive assessments, enabling students to choose different formats for demonstrating their achievement (eg written essay versus presentation). Providing choices that allow students to play to their strengths and preferred learning strategies can result in higher marks (McGahan et al, 2016). That said, in New Zealand, Hammond and Brown (2021) analysed correlations between students' scores (n=1185) on different assessments and on different learning outcomes and found only weak correlations on the same

learning outcomes across three different assessment formats: essay, oral presentation and reflective activity. The essay yielded the most differentiation in learner performance, suggesting it may have been more difficult than the other assignments. Their study raises questions about equivalence of different assessment methods when retrofitting learning outcomes assessment onto existing course assignments, something they argue is a common occurrence in their New Zealand context (Hammond and Brown, 2021).

In contrast, in Spain, Rivero et al (2017) found strong correlations between students' coursework marks (in a variety of formats) and final exam marks, although 60% of students did better on the coursework than the final exam. Doing better on coursework was associated with students' preference for coursework, lower university entrance exam scores, concurrently working, repeating the course, and Spanish versus English language of instruction. These results have implications for closing attainment gaps and warrant further investigation in other contexts. In general, students preferred coursework assessments over a final exam (Rivero et al, 2017).

In the US, Kahl (2016) studied the impacts of a system in which students' coursework assignments, or their exams, were automatically more heavily weighted by 15%, according to students' performance at the end of the module. This system led to more than a 1% increase in students' grades (an improvement of a fraction of a grade eg a B- to a B for 48% of the students), with more substantial differences for medium and low-performing students than high-performing students. Because they found larger impacts on students' grades when the differential weighing was increased, they urged caution over assuming equivalence of coursework and exam marks (Kahl, 2016).

In sum, more attention needs to be paid by practitioners to creating authentic assessments that are aligned with important learning outcomes and predict future performance. Greater attention to alignment with higher order learning outcomes may facilitate flexibility and inclusivity, but more research is needed to understand how to best actualise those principles. Coursework rather than exams may help level the playing field and address those principles. That said, the field needs to consider coursework in a more nuanced way rather than simply in contrast to exams. There are a variety of different types of assignments available, that offer different affordances, as seen in sections 3.1 to 3.5 above.

3.6.2 Online exams

Butler-Henderson and Crawford (2020) conducted a systematic review of 36 articles on online examinations published between 2009 and 2018. They concluded, based on seven studies, that most students prefer online exams over traditional paper and pencil exams. Students see online exams as saving time and offering flexibility. Although most students trust the results of online exams, more than half think it is easier to cheat (Butler-Henderson and Crawford, 2020). One of the empirical papers we reviewed also supported that conclusion. Ilgaz and Adanır (2020) studied Turkish students' perceptions of e-exams. Students were generally positive and some students, such as those who commuted long distances or had a disability, preferred the convenience of taking tests at home. Some students expressed concerns about time management, usability, and anxiety about potential

problems, although a forum was available for help during their timed test (Ilgaz and Adanir, 2020). Butler-Henderson and Crawford's (2020) review yielded inconsistent results on whether online exams provoked greater anxiety than traditional paper-and-pencil exams.

Anxiety is likely to be higher for students with less experience of online exams. Thus, some students will benefit from specific preparation for computer-based assessments, as Walker and Handley (2017) found based on input from UK masters' students, many of whom were international students and/or mature students. Therefore, those authors developed a framework with recommendations for teachers to help prepare students: thorough orientation to the testing procedures, including the rationale and suitability for the programme; guidance to students on e-assessment, including ensuring IT proficiency for assessment such as working with unfamiliar technology under time pressure; test-taking strategies for online exams; and attention to assessment design and exam interfaces (Walker and Handley, 2017).

In Australia, Ambikairajah and Tisdell (2019) helped students familiarise themselves with the online system with a mock exam in a campus computer lab that used short answer questions. Students were more satisfied with the appropriateness of the assessment and clarity of information about the exam requirements than previous years with paper-based exams, though there were no differences in their overall course satisfaction.

Some research was concerned with whether online assessments yielded the same results as paper-and-pencil assessments. Butler-Henderson and Crawford (2020) summarised four studies that found no significant differences in final exam scores between online and traditional exam modalities. Nonetheless, they reported that cheating was the most common theme encountered in their review, with studies about students' perceptions of the ability to cheat, staff perceptions of the risk of cheating, studies of technical approaches to cheating and technical approaches to cheating prevention. They found less attention to social and ethical interventions to reduce cheating, suggesting the need for more attention to those potential solutions (Butler-Henderson and Crawford, 2020).

The Covid-19 pandemic forced many institutions around the world to shift their exams to remote online delivery from early 2020. Jaap et al (2021) studied the impact of that shift in Edinburgh's Medical School, finding that an open-book, synchronous, timed multiple-choice exam with randomised question sequence yielded exam scores among fourth- and fifth-year students that were consistent with those students' previous performance. The exam was not proctored, nor were students' identities checked, other than that they needed to use a link sent immediately before the event to their university email address to access the exam. They were required to sign an honour code declaration before beginning. Few technical difficulties were reported, but more students preferred the invigilated computer lab setting to remote exams because they were concerned about fairness and a level playing field. For UK students in general, fairness in marking matters (Sutherland et al, 2018).

While Jaap et al (2021) did not emphasise authentication of students' identities, this issue has underpinned traditional proctored exams where photo identification is checked, and student behaviour is monitored during the exam. A wide range of technical solutions to authentication were identified in

Butler-Henderson and Crawford's (2020) review, though no clear conclusions were reached about the pros and cons of different approaches. With the shift to online exams in many parts of the world due to Covid-19, this subfield has likely moved on quickly, though the timing and specific focus of our review may not yet reflect these new developments.

3.6.3 Beyond grades: other ways of representing student achievement

Jorre de St Jorre, Boud and Johnson (2021) argued for reform of university systems that communicate students' competencies through GPAs or degree classifications and transcripts with marks aggregated across tasks at a module level. These systems, which are typical around the world, are mismatched with the needs of labour market and students seeking to demonstrate employability skills. Instead, they proposed a system that would enable richer documentation of valued skills and graduate attributes sought by employers, such as digital portfolios linked to digital credentials ('badges').

Flournoy and Bauman (2021) presented a worked example from an HEI in Canada demonstrating how such a process might be implemented. Students self-assessed each programme-level intended learning outcome (ILO) in three steps. First, the students said whether they achieved an ILO (yes or no). If yes, they provided evidence supporting that claim. Finally, they provided a reflection that explained why their evidence demonstrated that ILO. These self-assessments were then validated by academic staff. Although this process was designed to evaluate curricular effectiveness as part of larger institutional accountability processes, it could be adapted to fulfil the 'assessment for distinctiveness' approach advanced by Jorre de St Jorre, Boud and Johnson (2021). In effect, Flournoy and Bauman (2021) created a high-quality e-portfolio (see section 3.2.6 for more on portfolios and e-portfolios).

3.7 Design process: co-assessment

Overall interest in the sector in co-creation of education with students has also extended to co-assessment. First, we discuss studies of students' participation in writing exam questions before moving to a broader definition of co-assessment.

An online, Web 2.0 platform called PeerWise has now become common across many countries and universities to facilitate students' engagement in assessment co-creation, as well as peer assessment activities. Doyle et al (2019) demonstrated high student engagement, satisfaction and self-reported learning gains with an assignment in which business students wrote an MCQ on an assigned unit, along with distractors and a model answer that provided 5% of their total mark. In a follow-up study, Doyle and Buckley (2020) randomly assigned students to create exam questions based on one of five topic areas, then assessed whether students performed better in that topic area than the other four topics for which they did not create questions. They found significantly better performance on short answer questions on the topics for which students engaged in MCQ question creation than other topic areas, although students had access to all the MCQs generated for all topics for revision purposes (Doyle and Buckley, 2020). A useful follow-up study would be to compare the co-assessment activity with a similar teacher-directed revision activity, allowing a better test of the student-generation element of the activity.

In a robust study with large student numbers (n=1,200), internal and external controls, and researcher-rated levels of student engagement, Hancock et al (2018) concluded that student engagement with PeerWise results in an average mark uplift of 4%. Students were assigned to both generate and peer review MCQs. The better the engagement, the greater the gains students made on those topics. PeerWise engagement benefited students across the ability spectrum, and particularly those at the lower end. Having been trained on generating good quality MCQs that test higher order thinking, student questions were often high-quality and engaged students in dialogue about the substantive content. These studies suggest that short, low-weighted formative assessments involving co-assessment through a platform such as PeerWise have demonstrable impact on students' learning and performance, particularly when the peer-review affordances are incentivised. The fact that this tool harnesses crowdsourcing across the student body and demands relatively little teacher time makes it a promising approach for large classes. See also the section on peer assessment in section 5.1 below.

Quesada et al (2019) defined co-assessment more broadly, highlighting different stages and ways in which students might jointly negotiate with teachers toward consensus about assessment: during planning, development or deciding a mark. They privileged the process of dialogue and negotiation with teachers in co-assessment, a process that was not highlighted in the PeerWise studies above. They found that 470 education students in four different modules were generally satisfied with the dialogue with their teacher that led to the design of a group task and a negotiated final grade, noting that it helped them learn from their mistakes and strengthen their communication and relationships with peers and teachers. Kaur et al (2017) described a rich case example of how education students collaboratively designed inclusive group assessments that involved negotiating criteria for group assignment, choice of topic and modes of presentation, triangulation of both individual and group assessment data to arrive at final scores, and rubric co-design. In a study in China, Zhao, Zhou and Dawson (2021) investigated a peer assessment loop with international business students through co-constructed rubrics, rubric-guided learning, rubric-referenced peer review, and rubric-referenced reflection on students' perceptions of their assessment experiences. This integrated approach re-created real world business settings, with the co-constructed rubrics acting as a compass which steered students to developing professional competencies, values and the ability to assess their peers' work and their own. However, we identified no other literature on the impacts of other processes of co-assessment on student satisfaction, engagement, or performance. Further research on this innovative, emerging practice is warranted, particularly in disciplinary contexts beyond education.

In section 4 we have addressed a wide range of assessment practices. In doing so, we have emphasised key principles of high-impact practices and authenticity as underpinning effective practices. In the next section, we turn our attention to feedback design and practices. We will integrate recommendations across the three focal areas of the Framework (Advance HE, 2015) in the final recommendations section (section 6).

4 Feedback policies and practices

4.1 Introduction

The Framework (Advance HE, 2015) that guided this review privileges developing student engagement with feedback. However, the term engagement is broad and imprecise in measuring what happens during feedback processes and practices. In this review, we have concentrated on finding evidence relating to student satisfaction with feedback, the effect on learning behaviour and, where possible, performance improvements as a result of feedback-related interventions. The Framework (Advance HE, 2015) was released when feedback research was undergoing a substantial change in orientation. During the period of our review, researchers have begun to question our understanding of feedback within HE. There appears to be two cultures or paradigms of feedback. In the traditional approach, the focus is mainly on educators delivering feedback information to students. The second emerging paradigm, which has gained prominence in the last five years, instead emphasises student roles in generating, processing and using feedback (Carless, 2015; Winstone and Boud, 2019; Nicol, 2021). A recent, more learner-centred redefinition of feedback has emerged in which "students make sense of information about their performance and use it to enhance the quality of their work or learning strategies" (Henderson et al, 2019b, 1402).

Students struggle with feedback as a one-way transmission because information is often received too late in a cycle for students to take action (Carless, 2006) or they experience difficulties in decoding educators' statements (Sadler, 2010) and need dialogue with feedback providers to make sense of it (Winstone et al, 2017). New-paradigm research has begun to focus on how feedback is delivered to students (Ryan, Henderson and Phillips 2019; Mahoney, Macfarlane and Ajjawi, 2019), what students do with feedback, how it effects their future learning (Henderson, Ajjawi, Boud and Molloy, 2019) and the sociocultural dynamics of feedback interactions (Esterhazy and Damşa, 2019).

In 2018, Carless and Boud introduced the concept of feedback literacy, which represents the capabilities students need to benefit from feedback processes. Student feedback literacy involves seeking, generating and using feedback; making sound evaluative judgments; and working with emotions productively (Carless and Boud, 2018; Molloy, Boud and Henderson, 2020). Feedback literacy has attracted significant recent interest, including theoretical explorations in terms of sociomaterialism (Gravett, 2020) and sociocultural approaches (Chong, 2020); conceptual work on strategies for embedding feedback literacy within the curriculum (Malecka, Boud and Carless, 2020); discipline-specific explorations and interventions (Han and Xu, 2019; Noble et al, 2020) and a new learner-centred feedback literacy framework (Molloy et al, 2020). This work is advancing our understanding of how to help students develop, monitor and progress in the areas of feedback seeking, making judgements and using feedback to improve future work. This work puts students – rather than feedback providers – at the centre of feedback effectiveness.

This review reflects some of the older paradigm thinking about 'feedback as telling' as well as the newer paradigm interpretation of feedback focusing on student roles in generating, processing and using feedback. We begin the review by highlighting areas of feedback practices and processes that come after grading of student work. We begin by setting out the feedback expectations literature, then move on to the role of grades in feedback, the building of relationships through feedback and students'

emotional reactions to feedback and technology enhanced feedback. The second part of the review explores recent advancements in formative feedback that, in part, respond to some of the challenges and limitations of too much emphasis on feedback after grading. In particular, we explore formative feedback for learning, technology-enhanced formative feedback, types of feedback that can be provided, and exemplars. The review concludes with a discussion of the role peers can take in assessment and feedback processes. At the heart of the review, we focus on a wider discussion of the student's role in feedback practices and processes and how enactment of feedback can be achieved.

4.2 Feedback expectations

What constitutes effective feedback continues to be debated. We found that educators' and students' conceptions of 'effective' feedback are often misaligned. We saw research attesting to the virtues of educators and students engaging in ongoing dialogue surrounding enactment of feedback and refinement of work in progress (Carless and Boud, 2018; Dawson et al, 2019; Pitt and Carless, 2021). Yet, studies continued to show that students expect educators to give them high-quality information and guidance without considering their own role in the process (van der Kleij, Adie and Cumming, 2019; Winstone and Carless, 2019).

Countless studies in the last 20 years suggest students expect feedback and are not always satisfied with it. In the UK, Winstone et al (2016) indicated that psychology students felt the provision of good feedback information was the biggest necessity for lecturers to offer. Sparrow, Smith, Petronzi, Wilson and Roeschlaub (2020) revealed child and youth studies students wanted personalised feedback, not just standardised in-text comments and to be able to access feedback from different sources. For example, students preferred in-text comments and audio feedback and some form of follow-up meeting, which gave them the impression of feedback choice, ownership and responsibility. However, from an educator's perspective, replicating feedback in multiple locations is extra work.

Improving student satisfaction seems to be the agenda driving efforts to improve feedback. This agenda holds educators responsible for ensuring student satisfaction by providing more detailed, higher quality and speedier feedback. The concept of student satisfaction with feedback is not just a UK or Western centric issue. In India, Gopal, Singh and Anggarwal (2021) found that prompt feedback from an instructor had a positive effect on student satisfaction. In an Indonesian context, Basith, Rosmaiyadi, Triani and Fitri (2020) reported that students were satisfied with their interactions when the lecturer provided constructive feedback and when they had more opportunities to discuss it with course instructors. In the UK, Mulliner and Tucker (2017) reported that most staff and students thought that individual verbal (face-to-face) and individual typed feedback were effective forms of feedback. Interestingly, a significantly higher proportion of staff (86%) felt that that individual verbal (face-to-face) was very effective compared to students (63%).

Satisfaction with feedback is not the same as actual changes in learning behaviour, though. A number of studies reported what students thought would improve their performance without demonstrating actual changes in performance. In terms of learning behaviour, Mulliner and Tucker (2017) found most students (93%) said they always acted on feedback, and 72% agreed they understood how to make

best use of their feedback. In contrast, only 4% of educators thought students always acted on feedback (Mulliner and Tucker, 2017). Ngo and Ngadiman (2019) reported that providing students in Indonesia with immediate teacher feedback allowed them to assess their understanding of course content and improve their performance. Winstone et al (2016) explored students' thoughts about the qualities of good feedback. Winstone et al (2016) found that when they coupled students' perceptions of feedback with their approaches to learning, students classified as deep learners reported being less inclined to want improvement-related feedback. It may be that modularised curricula, a common feature of HE globally, plays a part in the feedback students want. When moving from one unit to another, students may find it easier to action feedback relating to skill development in subsequent work than feedback focused on the current task.

National-level student satisfaction surveys provide useful information about what students understand about and expect of feedback. MacKay et al (2019) studied the UK National Student Survey (NSS) responses for their institution and found that students described feedback as something educators do to students, suggesting a more passive role for students in the process. Students also reported they felt frustrated by feedback processes when educators did not attend to elements such as relationship-building, caring and understanding. For feedback to be useful, students first needed to understand what was expected of them and wanted examples of, or opportunities to, understand what constituted quality in their discipline. Subsequently they wanted opportunities to improve their grades, but feedback was sometimes ineffective because it was either received too late or came only after work had been submitted for summative marking. Another study from the UK echoed MacKay et al's (2019) findings. Francis, Millington and Cederlof (2019) carried out six focus groups with 28 undergraduate geography students. Students were asked about their conception of feedback in terms of its function, their experiences of good and bad feedback, how much they engaged with it and how it could be improved to help them in their learning. Students understood the purpose of feedback, but their desire to receive it was motivated primarily by wanting to improve their grade. Dissatisfaction occurred when they perceived the feedback as vague, incongruent with the grade awarded or without opportunities for further dialogue. Francis et al (2019) argued that to overcome this dissatisfaction, feedback opportunities need to be more agentic for students. They suggested increasing formative opportunities prior to summative work, where students can enact the feedback they receive and reducing feedback following summative work when there is no opportunity to enact it. In Norway, Vattøy, Gamlem and Rogne (2021) found that teacher education students wanted specific, understandable and actionable feedback that was timely prior to summative submission. Students were frustrated and less likely to use feedback when it came after the work had been submitted for summative marking. Students explained that educators' feedback was often monologic and transmission focused.

Providing students with feedback that aligns to the grade descriptors has been suggested. Despite offering this alongside more future-orientated feedback, Shafi, Hatley, Middleton and Templeton (2018) highlighted that education studies students in the UK only read half of this feedback. Nonetheless, the majority of students said feedback was more important than the grade received as it allowed them to understand how they can improve in the next assessment. About a third said they

returned to the feedback while they were writing their next assignment and a quarter indicated they sought further dialogue with their lecturer. Few students indicated more proactive engagement with the feedback in areas such as planning for the future, rewriting their feedback into language they more readily understood or consulting other sources. For most students, then, feedback was a static object that did not initiate much change in student learning behaviour.

Old feedback paradigms may be driving student dissatisfaction with feedback. Ali, Ahmed and Rose (2018) studied 447 psychology students from all three undergraduate years in a UK university. Generally, the students perceived their feedback experience negatively, worsening as they progressed in their studies. They reported a lack of dialogue within feedback exchanges, lack of improvement-focused feedback and no opportunities to revise work to enact feedback. Students did not readily make use of, enact or build into their learning process the feedback they have been provided, but they attributed these failings to the nature of feedback they receive and the feedback systems in which they were embedded. It is well-established that dialogic feedback affords students more agency within feedback exchanges (Steen-Utheim and Wittek, 2017; Vattøy and Smith, 2019). However, the power of dialogue may be lost when feedback only occurs following summative marking. Responding to this issue in Norway, Vattøy et al (2021) argued that even though students expressed a desire for formative assessment-related feedback, the culture and dominant practice of teacher educators was one of summative assessment with feedback after submission. This policy likely does not promote more proactive recipience during the feedback process, which Winstone et al (2017) argue is necessary to facilitate students' engagement with feedback.

Students' expectations also vary by year of study. Wei, Sun and Xu's (2021) surveyed 427 English for academic purposes (EAP) students across undergraduate years in a Chinese university. In the open-ended parts of their questionnaire, clear differences in feedback expectations between first year and final year students emerged. First year students expected personalised dialogic feedback, task-specific and constructive feedback. Final year students expected more future-orientated solutions with built-in chances to enact this feedback. Wei et al's (2021) findings suggest that students in their final year of study had transitioned from wanting more transmissive and evaluative forms of feedback in their first year towards wanting opportunities to enact, self-evaluate and improve future work as advanced students. This study highlights the need for educators to adjust their feedback practices to students' level of study. Final year students appeared to accept reduced opportunities for feedback, on the understanding they were expected to do more with this feedback and exercise a level of autonomy and self-direction in its enactment. Due to their relative inexperience in HE, the first-year students' expectations of what constitutes feedback may have reflected their experiences prior to university. Educators' approaches to transitioning students towards greater independence and agency is essential for their feedback literacy development over the longer term (Molloy et al, 2020).

Students still often fail to understand the role that formative feedback plays in improving learning. In Australia, Douglas Salter, Iglesias, Dowlman and Eri (2016) concluded that their students perceived feedback as only being related to summatively assessed work. Their students did not identify or consider other areas of their learning experience to be rich in feedback opportunities. In South Africa,

Impacts of Higher Education Assessment and Feedback Policy and Practice on Students: A Review of the Literature 2016-2021
Edd Pitt and Kathleen M Quinlan

Cohen and Singh (2020) reported that 55% of their fashion students said they did not really engage with formative feedback as it didn't always provide them with the scaffolding that resulted in self-development and performance improvement.

4.3 The role of grades in feedback

In this section, we review articles that looked specifically at the role grades played and their effect on students' perceptions of feedback, enactment feedback and wider learning behaviours. Marking, grading and feedback are fundamental to all HE provision. Both educators and students have an expectation marking will happen following all summative assessment. Indeed Chalmers, Mowat and Chapman (2018) argued that all marking processes are closely linked with feedback. On the one hand we have grading or evaluation of student work against predetermined standards that satisfy internal or external stakeholders that students have achieved the desired learning outcomes (Boud and Soler, 2016). On the other hand, we have feedback which is defined as "students making sense of information about their performance and using it to enhance the quality of their work or learning strategies" (Henderson et al, 2019b, 1402). The presence of grades often means that students don't read the feedback or fail to enact the feedback (Pitt and Norton, 2017).

Educators have various options for the kind of evaluative information returned to students with their grades. Azmat, Bagues, Cabrales and Iriberri (2016) carried out a randomised control trial over a four-year period in Spain with students studying business, economics, finance and law. In the control group, 354 students received grade outcome information. In the experimental group, 623 students received grade outcome information and information about their relative performance to the rest of the cohort. Students in the experimental group were sent an email every six months indicating their current grade and how it compared to the rest of the cohort. Seventy-two percent of the students in the treatment group checked to see where they were positioned relative to the cohort. On average they accessed this four times during the four-year period. Students who occupied the higher-ranking places in the cohort checked their ranking the most frequently. Students with lower grades checked the least. The treatment group students were significantly more satisfied than the control group, suggesting that being aware of (or, at least having the option to find out) one's relative position increased feelings of satisfaction.

Expectations for grade outcome seem to affect the way students process feedback associated with it. Pitt and Norton (2017) asked 14 UK sports studies students to discuss their experiences of 'good' and 'bad' work. Students brought one piece of work they felt represented bad work for them and another good work to the interview. In choosing the work to bring along, good work represented higher than personal grade expectation and bad work fell below their personal grade expectation. Students expected to receive feedback comments and that it should be the lecturers who generate and provide this feedback. Dialogic follow-up interactions with lecturers seemed to be reserved for querying grades that mismatched with their achievement expectation rather than discussing the substance of the feedback. Some students who achieved higher than their expected level of achievement ignored feedback. Some students achieving lower than expected grade outcomes demonstrated more

adaptive behaviours than previous literature had suggested. A dichotomy in interviewees' responses, with respect to the importance of grades, indicated that it is not simply a matter of the grade influencing how, or even whether, the student will pay attention to the feedback, but there is a more nuanced process whereby the students' predetermined grade expectations also seemed to affect their subsequent processing of feedback

4.4 Building relationships through feedback

The individualised nature of feedback and the creation of positive working relationships between educators and students is often emphasised within the literature. If relationships are central to creating personalised feedback, then it raises questions about the practice of anonymous marking. Pitt and Winstone (2018) sought to understand students' perceptions of anonymous marking and feedback and its impacts. Four hundred and forty-two (442) first year business, politics, pharmacy and French students in the UK experienced both anonymous and non-anonymous marking of coursework across different modules. Student performance data was collected for both conditions and a survey was administered following the completion of their modules to determine students' perceptions of each of the conditions. Performance outcomes did not differ according to whether their work was marked anonymously or non-anonymously. Feedback on non-anonymously marked work was perceived by students to have greater potential for learning than feedback on anonymously marked work. Students reported that anonymous feedback did not allow them to clarify things they had not fully understood, to initiate further feedback dialogues or improve their learning and study practices. Anonymous marking also depersonalised the educational experience. Anonymous markers could not reference students' previous performance, their use of feedback from previous assessments or be emotionally sensitive to that particular student. It is therefore not surprising that students thought non-anonymously marked work fostered a stronger relationship with the marker, enhancing their internal wellbeing and positive mindset. Anonymous marking misses a key opportunity to foster positive relationships between lecturers and students and limits the potential for feedback dialogue.

Previous research has demonstrated the benefits of dialogic interactions between students and markers through, for example, assignment cover sheets (Bloxham and Campbell, 2010; Carless et al, 2011) in which students ask for specific feedback and markers address these requests. Students gain a sense of agency when they have directly requested specific feedback and may be more likely to enact this feedback in future work (O'Donovan et al, 2008). In our study period, Keshavarz and Köseoğlu (2021) did a variation of the cover sheet exercise with 10 students from two 'English toward proficiency' (ETP) classes in a private university in North Cyprus. During the feedback-revision process, students responded to prompts about their written feedback preferences and expectations in "a letter to my teacher" activity. The activity gave students an opportunity to instigate a dialogue about the feedback they wanted on certain problematic aspects of their writing. Keshavarz and Köseoğlu's (2021) students felt the feedback was more individualised to them, and it made them feel a closer connection to their lecturers.

Using technology to administer feedback can also foster relationships between students and lecturers. In Australia, Lewis, Heath, Lim and Roberts (2021) exposed first year psychology and communication media students to weekly computer-generated personalised feedback messages about their progress. The software within the virtual learning environment (VLE) allowed instructors to produce 'if-then' rules which created predetermined messages relating to aspects of students' study outcomes. Students received feedback aligned to their performance outcomes, generalised guidance on how to improve for the future, and nudges to prompt engagement with wider learning materials. Focus groups with students revealed that, despite the fact the feedback messages originated from a software program, they still perceived them as being personalised to them. These messages appeared to increase the student's motivation to engage with the module and acted as a stimulus to complete various associated learning tasks. It fostered students' perceptions of connectedness with their lecturer, increasing desire to engage in more dialogic interactions seeking clarification. The software facilitated personalised feedback, but the success of the Lewis et al (2021) intervention was truly realised when lecturers were responsive to students' increased desire to engage in further dialogue. It seems that without this secondary engagement, the personalised feedback would not have necessarily initiated longevity of student engagement in feedback processes.

Although these studies show the importance of dialogue after students receive marks, more needs to be done to incorporate these opportunities throughout their course experience – during production of drafts, at the point of submission (eg cover sheets) as well as after summative marking. In their study about dialogic feedback, Ajjawi and Boud (2018) stressed how issues of power and discourse can influence how students interpret messages and subsequently act on their feedback. This is especially problematic when feedback is received after the students receive their mark. They argue that dialogic interactions, designed to break down barriers and foster working relationships work better as students are producing their work. We will discuss dialogue in formative feedback further in this review below.

4.5 Early and often: formative feedback for learning

Many of the problems raised in sections 4.2, 4.3 and 4.4 arise because of overemphasis on feedback that comes only after summatively assessed work has been submitted. Yet, formative feedback has long been recognised as a powerful contributor to student learning. Meta-analyses from the 1990s found an overall medium effect size of 0.41 for formative feedback across 131 studies (Kluger and DeNisi, 1996). A recent meta-analysis of 553 studies (Wisniewski, Zierer and Hattie, 2020) revealed an even larger effect size of 0.53 on cognitive outcome measures. In this section, we highlight how educational practices can be changed to privilege ongoing formative feedback prior to students submitting work for summative grading, and the effects of doing so.

Simply providing better written feedback – even in the formative stages – is not sufficient to produce performance gains. For example, Milne et al (2020) revamped support offered to a large final year Australian online nutrition module where students had traditionally received feedback on a draft assignment. They collected student satisfaction data for the same final year module in two consecutive years. They began from a position of 55% of students in 2015 agreeing that feedback had

helped them in their learning. In 2016 they resolved to provide more satisfactory feedback, not just a higher volume of feedback. The module traditionally used casual markers and so the authors instigated a rigorous process of support to develop their marking practices. Markers were provided with examples of previous markers' feedback, video guides on how to give useful feedback, and were paid to engage in a group marking exercise. More established markers from previous years also moderated and discussed the marking process with new markers. Students were given specific additional online seminars to build their assessment and feedback literacy and incentivised to explicitly reflect on how they had incorporated and enacted the draft feedback. The 2016 module received a feedback satisfaction score of 59% (from 81 respondents) and 61% in 2017 (from 86 respondents). Student satisfaction with feedback improved by 6% between 2015 and 2017. The additional effort, workload and staffing costs yielded only modest satisfaction increases and no observed improvement in student learning outcomes.

Students need multiple opportunities for feedback and repeated interactions about feedback over a module and course. Esterhazy and Damsa (2019) studied a 20-week undergraduate biology course at a large Norwegian university in which students completed three group assignments and participated in several iterative written feedback episodes with the educator. Students' interactions with each other, their teacher and the available resources mediated how students advanced their understanding of their discipline. Through the curriculum and feedback exchanges, they had opportunities to enact feedback, self-evaluate and improve future work. Educators created a culture of students finding solutions for themselves rather than telling them the answers, prompting students' responsibility for meaning making and acting on feedback. Similarly, Alfalagg (2020) demonstrated the effectiveness of English as foreign language (EFL) students in Yemen proactively participating in the feedback process. Students engaged in dialogue with their lecturers to receive feedback that identified gaps in their knowledge. Students thought the feedback was more meaningful and actionable as it was tied to a subsequent opportunity to redraft the work. Students improved their grade by 10% from the draft to the final submission. Thus, when students engage actively with various sources of feedback, the effects are positive.

It is important that educators provide feedback-rich learning environments which foster an increased appreciation of feedback. Thai, DeWever, and Vickle (2020) carried out experimental research with 104 second year education students in Vietnam. The students were assigned to one of two experimental conditions: with extra feedback (WEF, n = 52) and no extra feedback (NEF, n = 52). Students in the WEF condition received extra weekly feedback in the form of whole group teacher feedback relating to their quiz answers. The feedback was broad in scope and mainly focused on whether the answer was correct or not and hints for further explanation of the students' answers. Of this sample, 10-15 randomly chosen students also received individual feedback. In the NEF condition, students responded to the questions, posted their answers, but did not receive any teacher feedback. In both conditions students' mean learning performance increased from pre-test to post-test, but the WEF students scored comparatively higher. Interestingly, the WEF students and NEF students had similar self-efficacy scores at the beginning of the study but at the post-test WEF had increased by 6% whereas NEF had dropped very slightly. Students who received extra feedback on these formative

Impacts of Higher Education Assessment and Feedback Policy and Practice on Students: A Review of the Literature 2016-2021
Edd Pitt and Kathleen M Quinlan

tasks not only increased their performance over the course of the study, but also their self-efficacy and appreciation of feedback. Beccaria, Kek and Huijser (2019) found that students with a deep approach to learning were more likely to access and use feedback opportunities that enhanced their performance. Building a learning culture in which feedback becomes integral to the learning process may help shift students toward deeper approaches to learning.

Educators can make smaller changes from year to year that lead to more feedback-rich and effective learning environments. In Canada, Frank, Simper and Kaupp (2018) demonstrated how engineering educators incrementally changed the design of their feedback provision over three years. In 2013, the course was taught exclusively by didactic large lectures and feedback was only offered after the summative submission. By 2015 there were multiple opportunities to practice problem solving tasks, receive formative feedback from lecturers and peers, and self-assess performance as part of the summative assessment. Students completed pre-reading prior to class and then completed tasks in groups through an online platform. At the end of each class, the lecturer provided formative feedback designed to challenge and move student learning forward. Following the 2015 changes, the students reported being much more proficient in self-assessment and using feedback to improve than in the previous two years. (See also section 3.6 above on formative assessment).

To increase the dialogic and relational interactions between educators and students, O'Malley, Scanlon, Alpine and McMahon (2021) developed a five-minute feedback form (5MFF) for use in work-based Irish medical clinical placements. Students used the form weekly to self-reflect and self-assess their performance, identify areas for future development, and progress towards overall learning goals. Students reported that the 5MFF was easy to use and provided them with detailed and personalised feedback that promoted active engagement in the feedback process. The success was mediated by the degree to which the educator engaged with the 5MFF. Similarly, educators reported that students needed to actively seek feedback opportunities more, reflecting a need for them to understand and take responsibility for their role in the process (see Winstone et al 2021 for a wider discussion of student responsibility in feedback processes).

Two further examples – one from the USA and one from the UK – show that performance improvements occur when formative feedback on draft work is implemented. Uribe and Vaughan's (2017) formative assessment intervention was used with 89 students in a distance education course in the USA. In previous years, students could hand in one draft for formative feedback. The researchers changed this policy to allow two drafts, giving students personalised feedback on a first draft, time to enact the feedback into a second draft, and then a second round of feedback and enactment before submission. These opportunities were not in isolation though. Early in the module, the educators provided students with specific lessons on the purpose of feedback and how to enact it. This process afforded students multiple opportunities for feedback related to their work and then, more importantly, time to make sense of this feedback and to enact it. Following this intervention, over 50% of the cohort achieved exemplary marks compared to previous years when only 20% achieved such a mark. In the UK, working with second year ecology students, Hill and West (2020) discussed the assessment (an essay) during class time and asked students to write a draft and then discuss it individually face-to-face with their educator to generate feedback. Subsequently, near the end of the course, students

attended a seminar where they participated in a group marking exercise with two essays from the previous cohort and compared their marking with that of the educator. They marked their own draft essay, reflected upon previous educator feedback and the feedback generated in the group marking exercise. All of these formative feedback steps were designed to help students judge the quality of their own work in reference to others and the specific assessment criteria. The students who participated in the intervention received a mean mark of 63%, which was 7% higher than the cohort in previous years who were not exposed to the intervention. Students experienced an intervention which engendered positive feedback learning behaviours such as seeking, processing, reflecting on and enacting feedback in subsequent iterations of their work. This approach appeared to have a positive effect in the long term as students reported continuing their behaviour within their final year.

In the studies reviewed, many educators were designing formative tasks and associated feedback opportunities to incrementally help students develop prior to submitting their final summative piece. We need students to understand the potential impact that feedback can have on their learning over the longer term. Isolated formative feedback initiatives within a module may, on their own, not produce either increases in satisfaction or learning outcomes, as we saw with Milne et al (2020). Promoting such feedback practices earlier within a student's experience, repeatedly over their educational experience, and across multiple modules, may help them to access and learn how to use feedback for the longer term.

4.6 Students' emotions about feedback

In the past five years, researchers have focused on investigating students' emotional reactions to grades and associated feedback. Carless and Boud's (2018) feedback literacy model emphasised that practitioners and students need to understand the "affective" domain if students are to develop effective learning behaviours. Negative emotions can interfere with students' use of feedback information.

Negative emotions can arise from both negatively and positively worded feedback, though negatively worded feedback may be more problematic for students. Based on research in Australia, Rowe (2017) concluded that critical comments from educators can reduce students' self-esteem and perceived self-efficacy and lead to negative emotional reactions. But Bulut, Cutumisu, Aquilina and Singh (2019) found that students responded poorly even to well-crafted feedback tailored to students' strengths and weaknesses to elaborate deficient areas. Administering this feedback via a computer did not minimise possible harmful effects on students' self-esteem, and did not translate to immediate adoption, processing, or feedback-seeking. While Pitt and Norton (2017) found that negatively worded feedback sometimes spurred students on to improve next time, such students were in the minority. Most students found negatively worded feedback hard to process and did not see how it could help their future performance. Talking with 23 undergraduate psychology students in focus groups in the UK, Lynam and Cachia's (2018) found while the students appreciated and valued the supportive feedback their lecturers provided, stress and anxiety following feedback impeded their engagement with assessments. Students seemed to experience more positive emotions when the assessment afforded opportunities to be expressive, and associated feedback recognised their efforts.

Impacts of Higher Education Assessment and Feedback Policy and Practice on Students: A Review of the Literature 2016-2021
Edd Pitt and Kathleen M Quinlan

Students' expectations also affect their emotional responses. In Nepal, Neupane Bastola and Hu (2021) explored master's students' experiences of supervisory feedback during their compulsory English education and engineering thesis. Four hundred and thirty-four (434) students completed a questionnaire. Consistent with previous research in this area (Han and Hyland, 2019; Molloy, Noble and Ajjawi, 2019), the students reported positive affective engagement with supervisory feedback when they experienced support and genuine interest from their supervisors. Students also wanted feedback to be oral, written and positively worded, with detailed points for improvement. They responded negatively if these expectations were not met. Negative reactions were also associated with student unwillingness to accept feedback that required them to do extra work or consider conceptual changes (Neupane Bastola and Hu, 2021). Likewise, in the UK, Pitt and Norton (2017) argued that "emotional backwash" played a major role in students processing and using feedback following a disappointing grade outcome. Thus, students' pre-conceived concept of what constituted a good grade, or a bad grade, affected their reactions.

Students' expectations of feedback are also culturally conditioned, affecting their emotional responses to feedback. Ryan and Henderson's (2018) survey of 4,514 students across different disciplines compared international and home (Australian) students' perceptions of feedback. The majority of domestic students did not find feedback comments too critical, whereas the majority of international students said they did. Nearly 20% of the international students found their feedback comments upsetting, compared to domestic students (12.5%). Ryan and Henderson (2018) advocated practitioners consider potential cultural differences between international and domestic students and avoid standardised approaches and expectations about students' emotional responses to feedback. Similar points have been raised within a UK context where Rovagnati, Pitt and Winstone (2021) researched continental European students' experiences of feedback within the UK. Rovagnati et al (2021) advocated moving away from considering international students' reactions as deficits, suggesting that an increased awareness and understanding of intercultural competencies may help all students' emotional reactions in feedback situations.

Hansen and Mendzheritskaya's (2017) also found differences in emotional reactions to negative feedback between Russian, German and US students in a lab-based study with 383 students from these countries. Students read vignettes of a university lecturer giving negative feedback to a student on their performance. The researchers used a 2 × 3 × 3 experimental design to determine whether the negative feedback situation (private versus public), the lecturer's display of emotion (no emotion versus anger versus pity), and the cultural-educational contexts (German, Russian, US) affected the participants' predictions of the student's affective and behavioural reactions to the negative feedback. German and US participants reported stronger emotions in response to the negative feedback than the Russian participants. Russian students attributed failure more often to external, unstable and controllable factors and adopted more approach tendencies after receiving the failure feedback. Cultural-educational and situational contexts affected the way students responded to the emotions a university lecturer displayed, and those emotions can shape students' learning behaviour.

With repeated exposure and support, students can adjust to new feedback cultures. Li and Curdt-Christiansen's (2020) studied five Chinese international students' experiences of UK higher education following previous study in China. They discussed how the students' understanding of teacher feedback developed through successive experiences. Initially students found the feedback wording harsher than they were used to, which provoked negative emotional reactions that had to be overcome before they could enact feedback. The authors encouraged lecturers to consider the language they use in feedback and think about cultural implications. Repeated learning loops and associated feedback over the long-term helped students to engage cognitively with lecturer feedback.

Feedback also affects students' self-assessment accuracy and self-efficacy, which are both important to students' subsequent enactment of feedback. In a lab setting in Korea, Kim and Lee (2019) studied the effects of a nursing examiner's positive and negative verbal feedback on students' self-assessment accuracy, emotional responses and self-efficacy. They randomly assigned 110 second year South Korean nursing students to a negative feedback (NF) or positive feedback (PF) group. The NF group demonstrated greater self-assessment accuracy, but more negative emotions than the PF group. In the PF group, positive emotions and self-efficacy scores were higher. This lab-based study has implications for classroom-based settings, suggesting that practitioners should be aware of the effect negative versus positive feedback might have on students.

Students' self-efficacy also seems to play a vital role in the enactment of feedback for improvement. Adams et al (2020) surveyed 232 second year students from 10 different subject areas in the UK. The students' academic self-efficacy mediated their perceptions of feedback and grade outcome. Highly efficacious students were more likely to feel positive and competent and, therefore, were more disposed to accept challenging feedback encounters, interpreting them as opportunities to improve. This study corroborated Winstone et al's (2017) finding that self-efficacy mediates time spent by students reflecting on feedback information. Students with higher self-efficacy were more likely to interpret the feedback positively and enact it to achieve their goals. Conversely, students lower in self-efficacy were more likely interpret negative feedback in less adaptive ways (Adams et al, 2020). Crucially, Adams et al (2020) suggested that feedback itself should not be the main emphasis; rather, practitioners should focus upon raising students' self-efficacy, so they attend productively to the feedback and enact it in subsequent opportunities.

In sum, students may have negative emotional reactions to feedback which, in turn, impedes their ability to process, interpret and act on that feedback. Negative reactions may be more likely if the feedback is, itself, worded harshly or negatively or if students are unaccustomed to receiving feedback that challenges them. Students who are more emotionally mature, more experienced with the culture of feedback, or have higher self-efficacy tend to be more able to respond constructively to feedback. Practitioners may be able to help students by offering early formative tasks which are feedback rich and build in discussion surrounding how to interpret, react and enact this feedback. A key function of feedback literacy development is helping students to appreciate feedback and manage their emotional reactions. Integrating this development within the curriculum seems a fruitful way to contextualise it

Impacts of Higher Education Assessment and Feedback Policy and Practice on Students: A Review of the Literature 2016-2021
Edd Pitt and Kathleen M Quinlan

within the disciplines or profession's ways of doing things. Developing students' capacity to learn from feedback is a vital graduate attribute that will benefit them long after they leave HE. As the field of feedback literacy development grows, more research is needed specifically in understanding how to help students manage the potential negative effects of emotional backwash in feedback situations.

4.7 Technology-enhanced feedback

We have seen an increase in new digital tools aligned with assessment innovation (Bearman, Boud and Ajjawi, 2020; Bearman, Dawson, Ajjawi, Tai and Boud, 2020). Technologies can simply replicate existing processes or they can make existing processes more efficient, which can bring incremental benefits to students or teachers. Some technologies, in the right hands, have the potential to transform what we do. Our challenge is to consider how we can use technologies to radically improve educational practices. In the following sections we will analyse how different types of technology-enhanced feedback have been operationalised and their effectiveness.

4.7.1 Immediate, automated feedback

During the period of our review, a number of applications have been studied that afford opportunities for immediate, automated feedback. These applications allow educators and students to gauge understanding of the theory and concepts being discussed in classroom settings or to create bespoke automated formative tests that feed into aligned summative assessments. In Spain, Kahoot has been used at the end of practical engineering lessons to gauge students' understanding (Cantero-Chinchilla et al, 2020). In Norway, Egelandsdal and Krumsvik (2019) reported that 85% of students from law, psychology and social science disciplines (n=400) perceived a formative feedback clicker intervention in their lectures as useful for self-monitoring, enhancing their content understanding and revealing gaps in their knowledge. The feedback initiated a change in their learning behaviour to address these gaps. For more evidence on the value of immediate automated feedback, see also section 3.5.2 on gamified learning and 3.6.1 above on automated assessments.

The automation capabilities of the technology can improve the quantity, quality, and timeliness of feedback on straightforward knowledge tasks. Students can be set formative computer-based tests (CBT) they complete on their personal devices. Most students (86%) in Nardi and Ranieri's (2019) Italian study reported that the immediate feedback this method afforded was the best part of their experience. However, the method of assessing students formatively also offered them the chance to modify their answers prior to submitting them. Just over half of them (55%) were satisfied with this opportunity, suggesting that they would rather receive the automated feedback following their initial answers than changing answers as they went along. In a Taiwanese sight translation context, Chang (2016) found that students greatly valued the instant teacher feedback provided in class which helped them to understand the problems in their renditions. One criticism of technology enhanced automated feedback is that it generally follows on from closed quiz-based questions which can promote surface learning. (See also section 3.6.1 above.)

Virtual reality is also an emerging technology that can offer more nuanced feedback on complex student performances. In Holland, Belboukhaddaoui and van Ginkel (2019) studied the impact of immediate versus delayed feedback within a virtual reality-based activity. Thirty students were tasked with presenting their work to a virtual audience. One group received immediate computer-generated feedback, while the other received delayed feedback at the end of the presentation. The immediate feedback was displayed on the virtual reality screen while the students were presenting, giving them cues relating to their eye contact with the audience or the speed of their voice. The students in the delayed feedback condition did not receive any cues on the screen. Rather, they received feedback on these areas at the end of the presentation. Subsequently, both groups presented to a physical audience of peers within the classroom and were assessed in a similar manner to the virtual reality practice session. The results revealed no differences in performance outcome between the two conditions. Thus, in-the-moment performance-based feedback through virtual reality did not enhance students' performance compared to delayed feedback at the end in a subsequent task.

Technologies can also facilitate more discursive feedback on complex skill development. We will discuss those affordances in the next sub-section.

4.7.2 Written versus audio/video feedback

Prior to 2016, researchers considered the merits of written feedback (Jolly and Boud, 2013; Vardi, 2013) and audio feedback (Gould and Day, 2013; Lunt and Curran, 2010) separately. During the review period, researchers began to compare these different modalities. Knauf (2016) directly compared students' perceptions of audio versus written feedback with 52 online social work students in Germany. Students submitted an essay for summative grading for which they received audio feedback and a grade by their lecturer. The students then waited a further 14 days before they received written feedback. Students' perceptions of this change in feedback practice were dichotomous. Some preferred the audio feedback and others the written, but the audio feedback felt more personal than the written feedback. Woolstencroft and de Main (2021) gave final year UK business students group-level feedback one week after their summative essays were marked. Subsequently they provided three minutes of individual audio feedback before revealing their grade. Forty-three of the 182 students provided input on their experience of the intervention. The process of having to listen to the feedback prior to receiving the grade was positively received. Those achieving higher grades reported using the audio feedback to go back into their essay to see where there were issues. They also felt the feedback was more personalised to them. Students said it was a welcome departure from their usual experiences of feedback when they looked at the grade and, if it met their expectations, they would ignore the feedback. Thus, this intervention interrupted students' counterproductive habits.

Evidence is mounting for the advantages of video feedback (Mahoney, McFarlane and Ajjawi, 2019; Grigoryan, 2017) and screencast feedback (Mathisen, 2012; Turner and West, 2013). Students have reported valuing video feedback because it is clear, detailed and provides visual and aural communication of feedback information (Anson et al, 2016; Elola and Oskoz, 2016; Mayhew, 2017).

Students also report that video feedback feels more personalised, caring and engaging (West and Turner, 2016; Stannard and Mann, 2018). Mayhew (2017) reported that students reviewed their video feedback several times and often made detailed notes about their feedback as they watched. As educators, we want to encourage students to adopt these kinds of positive feedback behaviours.

Video feedback seems to be different from written feedback in several ways that may explain its effectiveness. Cavaleri et al (2019) evaluated whether audio and video feedback helped students engage with the feedback more successfully than conventional written methods. They employed a longitudinal, mixed method design to quantify impacts and explore differences between written feedback and recorded audio-visual feedback. Twenty undergraduate students in Australia were assigned an academic skills advisor and submitted two draft assignments to the advisor for language and literacy feedback. One of their papers received written feedback via the comments feature in MS Word. The other paper received audio-visual feedback, which included screen capture video feedback, audio commentary and minimal written comments. The students then redrafted their work for summative submission. Written feedback tended to be more directive whereas video feedback included more suggestions, explanations and elements of relationship-building through praise. Video feedback comprised commentary as well as image, text and movement on screen which all contributed to the meaning-making process. Written feedback was typically direct, compact and concise, but spoken feedback tended to repeat and recycle information, which helped reinforce points and clarify the intended meaning. Cavaleri et al (2019) marked the work and then compared both pieces of work for each student, in each condition, to see how much feedback had been addressed from the draft to the summative submission. Seventy seven percent of the written-only feedback was used in a subsequent revision, compared to 88% of the video feedback. When students received video feedback they were 2.17 times more likely to make successful changes to their work than written feedback alone. This intervention seems to have had the most profound effect on students at the lower end of the grade spectrum. When receiving written feedback, lower achievers address only 53% of the feedback comments, compared to 78% when they received video feedback. Thus, video feedback may be an important element in closing attainment gaps.

Audio and video feedback appears to help support the positive relationships between student and teacher we highlighted in section 4.4 above and is important in reducing negative emotions (section 4.5 above). Rasi and Vuojarvi (2017) investigated replacing written summative feedback with formative audio feedback on draft case analysis work. Students listened to audio feedback, made changes to their case analysis and explained the changes they had made. Audio feedback was perceived as being personal whether or not the students met with their lecturer face-to-face. Students indicated they wanted both text and audio feedback as both had their benefits. The audio feedback was effective given that 70% of the comments were addressed by students in the summative work.

Video feedback may be as good as in-person verbal feedback for some students. Seifert et al (2020) carried out a randomised, controlled and blinded study with German dental students to compare direct expert feedback given face-to-face, individualised video feedback and general video feedback. Students were randomly assigned to one of the three feedback methods. All students were shown an

instructional video that detailed a dental procedure and were asked to film themselves carrying out the actual procedure. They received feedback aligned with one of the three experimental conditions. Six weeks later they were summatively assessed via an OSCE. In all groups, regardless of feedback modality, significant performance improvements on the dental procedure task were seen.
The comparison between each group in the post-test showed no significant differences between the three groups. These findings suggest that it is the verbal/visual interaction that matters, rather than whether it is captured on video. Given the logistical difficulties of individual meetings with each student, recorded video feedback is more feasible for most academics. Thus, the technology brings an ideal educational practice into reach. It is surprising, though, that general feedback was as effective as individualised feedback. Given that generalised feedback is much more efficient from a lecturer's perspective, these conditions need further study in other contexts.

Unlike the other studies reviewed, Hunukumbure, Smith and Das (2017) used videos to stimulate recall and dialogic feedback among a group of students, rather than just replicating "teacher telling". Hunukmbre et al (2017) filmed third year medical students while they completed their summative OSCE. Twenty-eight (28) students attended a feedback session where the video clips were used as a stimulus for feedback dialogue. Following each individual video clip, the featured student was encouraged by the lecturer to reflect on their performance, and peers were asked to provide constructive feedback. This dialogic interaction focused on actual evidence of authentic practice. Hunukumbure et al's (2017) videos provided a true and accurate reflection of the student's actual performance in the OSCE, rather than relying on fuzzy or conflicting memories of events. The video could be paused at certain moments, rewound and reviewed to emphasise issues and stimulate specific peer feedback. Indeed, some students were unaware of their mistakes, and this process helped them to better understand the mark they received and areas they could improve upon next time. Similarly, the videos of excellent performance acted as exemplars to other students (see section 4.9 on how to use exemplars effectively). As we will highlight and discuss later in this review, some students did struggle with the peer feedback element (see section 5.2 below). Some students were overly positive and avoided negative comments, as they feared upsetting their peers.

4.8 Types of feedback offered

The type of feedback students receive and its effect upon subsequent performance has also been discussed vis-a-vis online formative opportunities. As seen in section 4.7.1 and in section 3.6.1 above, technology provides opportunities for immediate feedback, which seems to be beneficial for student learning. Here we consider the effects of elaborative versus corrective feedback.

4.8.1 Elaborative versus corrective feedback

In most of the studies we reviewed, elaborative feedback was better than corrective feedback or no feedback. In their systematic consideration of the feedback literature Butler and Woodward (2018) critiqued the benefits of elaborative feedback over corrective feedback for enhancing student learning. In Canada, following multiple choice quizzes, second year physiology and psychology students preferred elaborative feedback over feedback containing only the correct response, or no feedback at all (Smith et al, 2019). In Croatia, across three online quizzes with electrical engineering students, Petrovic, Pale and Jeren (2017) manipulated the feedback students received. They were either given no feedback, corrective feedback indicating the correct answer, or elaborative feedback involving the correct answer plus a short explanation including a misconception that might have led to the incorrect response and a reference to the course material. Students had higher knowledge gains with elaborative feedback than in the corrective feedback and the no feedback conditions.

Mujtaba, Parkash and Nawaz (2020) conducted an experimental study with 90 business communication students. Two markers were trained in the use of a rubric to standardise marking of students' written drafts. Standard corrective feedback was provided to all, while the treatment group also received additional short explanatory comments. Redrafting was carried out a further two times and results showed statistically significant difference between those students in the treatment group versus the control group on the final summative submission. While in the control group improvements were seen based on minimal corrective feedback, the addition of further explanatory comments assisted the treatment group to improve their mark by 11% overall. Zeglen and Rosendale (2018) randomly assigned 65 US HE students to a control group (n=22) or one of two treatment groups (both n=21). All students participated in a series of instructional games. The control group received no specific indication of the correctness or incorrectness of their responses but did find out their final score. Treatment group one received elaborate feedback after each game relating to the correctness of their responses and why this was the case. Treatment group two had the same information as treatment group one, with the addition of visual hints on the screen prior to answering each question. The results showed no significant difference in comprehension between the three groups. In terms of remembering the material, the students in treatment group two did score significantly higher than the other two groups. Their experience of elaborate and visual hint-based feedback may have improved their performance by enabling transfer from working memory to long-term memory across the tasks.

In Holland, Thijssen et al (2019) conducted a randomised control trial with 463 first year biomedical students. Students completed seven formative tests during a four-week course, where half the students received elaborative feedback and the other half no feedback. No significant differences were found the two groups on the final exam. At first glance it may seem like the intervention had no demonstrable impact upon students. However, those who received elaborative feedback reported spending more time engaging in further reading of the material. Such an intervention may not have provided immediate improvements compared to the control group, but it may have had a positive impact on student learning behaviours that they would take into subsequent assessment opportunities.

4.9 Exemplars: helping students understand quality standards

The use of exemplars for helping students understand assessment criteria and standards has emerged as a promising practice in the last 15 years. Exemplars are samples from within a discipline of how others have responded to an assessment task or brief. Despite the name, exemplars are not necessarily the best examples, but typically represent a range of performance. Exemplars are

typically anonymous samples drawn from previous cohorts, rather than samples from peers in the same class, which are addressed under peer assessment below (section 5). Studying exemplars can help students understand assessment criteria and appreciate assessment standards. In essence, they help students understand what constitutes quality (or lack therefore of) within a given assessment. Discussing exemplars among peers and with teachers can help students to develop their evaluative judgments and apply this new understanding to their own work (Carless and Chan, 2017; Carless, 2020; Nicol, 2021). Often this process is scaffolded through the use of marking rubrics to help students see how criteria are achieved (Panadero and Jonsson, 2020). These rubrics can either be designed by the lecturer or co-constructed with students to enhance their engagement with and awareness of the specific criteria (Bacchus et al, 2020).

A recent systematic review carried out by To, Panadero and Carless (2021) identified the conditions under which exemplar use benefited student outcomes. Studies that offered evidence of demonstrable impact on student learning outcomes typically employed two or three opportunities for exemplar use and often within the first two years of the student's experience (To et al, 2021). They also found that most studies on exemplars have been carried out in western centric contexts and published in the last six years.

Discussions of exemplars help students understand what is required in a task and may support self-regulation of their performance. Students enrolled on Australian education courses across a range of study levels reported that exemplars were useful to them in learning about levels of quality and helped them interpret the criteria (Grainger, Heck and Carey, 2018). Most importantly, the exemplars helped students to analyse what they needed to do in the task ahead. Broadbent, Panadero and Boud (2018) carried out a study with 1500 psychology students. In tutorial groups of 30 students each, teachers used one exemplar of high-quality work with detailed associated explanations relating to marking criteria to help students understand the standards expected in a summative journal assessment. Students also received detailed audio feedback following submission of each section of their journal-writing assessment. Students performed significantly better than previous cohorts that did not experience the exemplar approach. Hawe, Lightfoot and Dixon (2017) explored first year business students' development of self-regulation following their experiences of exemplars. Their results attested to a range of benefits including students' self-perceived enhancements in self-efficacy and self-monitoring.

Students, particularly those earlier in their education, may benefit from teacher assistance to make the most of exemplars. In Hong Kong, Carless and Chan (2017) asked teacher education students to analyse two exemplars, share their judgements with a peer, and then discuss these within a lecturer-facilitated session. Students often generated feedback that was not critical enough or not aligned specifically to criteria. Carless and Chan (2017) argued that the lecturer's facilitation, through elaborating students' thoughts and judgements to unpick and develop their diverse opinions, was crucial to making the exemplar process successful. To and Liu (2017) used a similar process with first year business students in a Chinese context and also found that some students struggled to reconcile disputes with their peers when differences in judgements about the quality of the exemplars emerged.

On an animal science programme in Australia, Hendry, White and Herbert (2016) gave students advice about marking and rubrics before they graded two exemplars using the lecturer- designed marking rubric. Subsequent peer-to-peer discussion and a follow-up lecturer-facilitated explanation helped students to clarify their judgements. Thirty-nine responses to a follow up questionnaire revealed that 64% of students said they had spent time marking the exemplars on their own and 56% used the marking guide to make their decisions about its relative quality. The majority of students (72%) reported disagreeing with their peers' interpretation of the exemplars. As a result, 44% of students aligned their marking decision with that of the lecturer. Thus, the final step in the intervention, in which the lecturer clarified how they marked the exemplar, appeared to be important in offsetting students' potential misunderstandings. This final lecturer-led session was the most popular part of the process.

Repeated opportunities within the curriculum for discussion of exemplars are also very important. Repetition may help students learn how to negotiate disagreements productively with peers, which may be a useful and necessary part of the process of critical reflection. Hawe and Dixon (2017) worked with undergraduate teacher education students in New Zealand in a 12-week module. Students completed three 500-word pieces of writing and then, later in the module, a longer essay. Students engaged with annotated exemplars of 500-word written pieces in a three-hour teaching session during which they assessed the quality of these exemplars against their own drafts. They subsequently compared their own work to their peers through several formative peer review processes during the module. They appreciated their peers' feedback, finding it both supportive and challenging. These exchanges sometimes yielded disagreement between peers in their approach to writing for the assessment. Over multiple dialogues, though, students critically reflected on the quality of their own work and that of their peers. Thus, disagreements in peers' interpretations can be used constructively to foster critical reflection.

In sum, discussions of exemplars can help students learn about quality standards, if used early, repeatedly and with appropriate teacher support. Early discussion affords students more time to realise the impact of their use. We have found that lecturers are often hesitant about dedicating precious class time to discussing exemplars. However, if assignments are authentic reflections of the skills required of graduates in a field, then helping students to learn the quality standards for those performances are an essential part of students' education. Peer discussions of exemplars are also frequently used. We will elaborate on how to engage students productively in peer feedback in section 5.2 below. Peer assessment and peer feedback processes are distinguished from exemplars in that the work samples are produced by a students' peers in the same cohort, which introduces dynamics we will consider in section 5.1.

5 Peer assessment and peer feedback

Peer assessment and peer feedback are well-established in HE and have well-documented positive impacts on student learning. The terms are closely related, contested and sometimes used interchangeably. For the purposes of this review, we define peer assessment as students assessing their peers' work against set criteria and providing both qualitative feedback and a numerical grade. Within peer assessment activities, students can be both the assessor and assessee. Peer feedback focuses on the actual content, value and use of the qualitative feedback information generated and received by peers. The goal of peer assessment and peer feedback processes is to build students' capacity to self-assess, although this is implicit in much of the research. The Framework (Advance HE, 2015) refers to self and peer assessment together, which is supported by the literature. The following sections discuss peer assessment and peer feedback separately, with occasional crossover.

5.1 Peer assessment

In peer assessment, students make judgements about the quality of their peers' work, give indicative grades and provide feedback comments designed to help their peers improve (Min, 2016). Peer assessment has been used with many different types of assessment (Bijami et al, 2013). In the period of our review, there have been several systematic reviews (van Popta et al, 2017; Panadero and Alqassab, 2019) and meta-analyses (Huisman et al, 2019; Double et al, 2019; Li et al, 2020) that demonstrate the positive effects of peer assessment on student learning outcomes and key conditions for its success. Double, McGrane and Hopfenbeck (2019) conducted a wide-ranging meta-analysis of the effects of peer assessment on academic performance in primary, secondary and tertiary education settings. Across these stages of education, participating in peer assessment moderately improved students' performance outcomes compared to no assessment and teacher assessments. Peer assessment was less effective in higher education than in primary and secondary contexts, though. There were no differences between anonymous and non-anonymous peer assessment. However, Double et al (2019) reported that across the 141 studies included in their review, peer assessment was not superior to self-assessment. Nonetheless, Double et al (2019) concluded that peer assessment has educational performance benefits.

Recent research has shown that HE students do not always value peer assessment. Melser, Lettner, Bawer, Puttinger, Holzinger (2020) asked 423 second year Austrian medical students about their attitudes to peer and self-assessment. Most students felt that peer assessment was, at times, not objective enough, they didn't take it seriously and it had little impact on their learning. They suggested that self-assessment was more helpful in improving their knowledge over the longer term. Within a Chinese context, Zou et al (2018) investigated students' attitudes and participation in peer assessment. Students reported a generally positive attitude toward peer assessment. However, they had concerns about interactions with peers and the efficacy of their peer's assessment judgements. Students lacked prior experience, teacher feedback was preferred and a reluctance to take responsibility for peer assessment was seen. In a Confucian-influenced culture, teachers are considered as the knowledge authority and the proper assessor, which may have contributed to students' perceptions (Zou et al, 2018). Similar results were found by Park (2018) in a Korean tertiary setting, where the students felt reluctant to accept their peers' assessment judgments compared to those by the lecturer. Chinese engineering and English majors in Zou et al's (2018) study indicated that a further barrier for peer assessment was students' negative perceptions about how many times they might need to engage with it before seeing a performance improvement. However, for Korean students in Joh's (2019) study, competitiveness and a desire to keep up with their peers stimulated them to participate and perceive peer assessment as beneficial to their learning.

Although evidence suggests its effectiveness, teachers also often express reluctance. Australian academics interviewed by Adachi, Tai and Dawson (2018a) expressed concern about personal power relationships within the peer groups and students' distrust of their peers' ability to assess correctly. peer assessment also creates more peer interaction and active engagement with material, both of which are key elements of high-impact educational practices (see section 3.1). Lavrysh (2016) found that Ukrainian English for Specific Purposes (ESP) students' perceptions of peer assessment were favourable because classroom-based sessions were more interactive, evaluative, critical and collaborative and provided useful feedback to improve learning outcomes. Peer assessment affords students opportunities to examine their own and their peers' performance (Schunn et al, 2016). This can lead to a period of self-reflection (Liu et al, 2019), instigating a reciprocal consideration of strengths and weakness of the work produced (Panadero and Alqassab, 2019; Seifert and Feliks, 2019).

Self-assessment is likely a key mechanism in explaining the effectiveness of peer assessment activities. Gezer-Templeton et al (2017) applied an emerging approach called "exam wrappers" to prompt students' reflection in their exam preparation, improve their study skills and enhance future exam performance. Students completed the exam wrapper within one week of each of three exams in an introductory level course. The exam wrapper asked students: 1) How did you prepare for the exam? 2) What types of questions on the exam were most challenging for you? Why do you think they were challenging? 3) What changes to your study habits do you plan to make when preparing for the next exam? Students also reported the amount of exam preparation time, how far in advance they started studying for the exam and reflected on their experience and perceptions of the exam wrapper after exams 2 and 3. While the study design makes it difficult to attribute changes to the intervention itself since there were also other activities related to teaching study skills occurring in the course, students tended to increase their study time for later exams and to start studying earlier. Most students agreed that the exam wrappers helped them improve their study habits and half of them thought it helped them improve their exam scores. More than half also applied the same kind of post-exam reflection to their other exams (Gezer-Templeton et al, 2017). The authors suggested that the intervention may be particularly effective for students in the middle of the grade distribution.
Exam wrappers are a way of promoting self-assessment of study processes. Similar questions could be discussed among peers in processes like those we have described in this section. While the focus of Gezer-Templeton et al (2017) was on exams, "wrappers" could be applied to other forms of assessments students undertake.

5.1.1 Reliability of peer judgments

As we defined it, peer assessment involves not only feedback (on which students may disagree, as we noted under 4.9 above), but also scores or grades. Thus, the reliability of students' judgments is important if students use those scores to improve drafts or they become part of students' final grade, as sometimes occurs in massive open online courses (MOOCs). Here we consider conflicting evidence over reliability and suggest evidence-based ways of improving the quality of students' peer judgments.

Some research has found peer assessments were consistent with teacher assessments when using a lecturer-designed rubric or scaffold. Lecturers play a key role by creating the criteria on which peer judgements are based (Murillo-Zamorano, 2018; Ramon-Casas, Nuno, Pons and Cunillera, 2019). Li et al (2016) found that students' assessments of their peers' work was comparable to the lecturer's assessment. Ramon-Casa et al's (2019) research with 365 Spanish psychology undergraduate students found consistency of assessment judgements between students and lecturers on a lecturer-designed rubric. Consequently, students receiving and enacting the associated feedback can improve their performance outcomes (Double et al, 2019). Meek, Blakemore and Marks (2017) explored peer assessment within a biomedical science MOOC with 190 students who participated in a peer review process. Students used three questions to scaffold their review, with no other training provided. The lecturers judged the standard of peers' qualitative judgements on a 300-word writing task as high. However, students who achieved the higher end of grades appeared to be more adept in providing higher quality feedback.

Other researchers have questioned the reliability of peer judgments, specifically in a MOOC context. In China, Li et al's (2021) analysis of more than 5,000 peer assessments across 18 different assessments in three different MOOCs demonstrated a huge range in variability of marks given at both ends of the marking spectrum. Student-generated marks also did not align with those a lecturer would give. Similarly, within a Spanish MOOC, Garcia-Loro et al (2020) reported that across 63 peer assessment activities, the reliability of peer assessment was very low. In both these research studies, the authors recommended that peer assessment should not be used in MOOCs given the variability in participants' ability to make suitable judgements.

As shown under 4.9 above on exemplars, students need repeated exposure and teacher guidance to hone their evaluative judgments. Specifically, training in peer assessment may be needed. In Li et al's (2020) meta-analysis of 58 studies, they concluded that participating in peer assessment showed a .291 standard deviation unit increase in performance compared to not participating. But rater training was the most critical factor explaining the effects of peer assessment. In studies in which students were offered training, peer assessment effect sizes were greater than when it was not offered (Li et al, 2019). Investigating the effect of teachers' and senior students' feedback on the writings of a group of less experienced Thai EFL students, Sritrakarn (2018) reported beneficial results but emphasised the necessity of teacher intervention to ensure the positive impact of senior students' comments.

As we saw in section 3.2.5 above, students are often involved in assigning grades to peers during group assessment. While we argued that peer judgments of students' contributions are a primary way of operationalising individual accountability in groupwork, student bias in marking may be a concern. Research from Australia found that although students were routinely able to make accurate judgments about their peers' formative work, peer grading bias was apparent when such evaluations were summative and counted towards their own final grade in group work (Sridharan et al, 2019).

The process of making a judgement of peers' work is complex; there are multiple areas the marker must consider. Students need to have a working understanding of the rubric itself, reflect on and consider any inputs the lecturer may have made, read and re-read their peers' work and then make a final evaluation of its merits against the set criteria (Chen et al, 2020). This process is complex even for a seasoned practitioner, so it is unsurprising that students need training and multiple opportunities to improve their peer assessment judgments. Doing so, though, can enhance student performance on key tasks, support individual accountability assessment in group work and, potentially, reduce lecturers' burdens in high enrolment courses such as MOOCs.

5.1.2 Technology-enhanced peer assessment

Prior to our review, questions about the role that technology played in peer assessment were beginning to emerge. Research since 2016 has focused on understanding how online peer assessment may differ from traditional face-to-face models. Zheng, Zhang and Cui (2020) conducted a meta-analysis of 37 empirical studies published between 1999 and 2018. Technology-facilitated peer assessment had a significant and medium effect on learning achievements, with an overall mean effect size of 0.576. Students who had technology-enhanced peer assessment opportunities outperformed students who did not. (See also the studies reviewed using PeerWise in section 3.7 above, which reached similar conclusions).

Examples from within the peer assessment literature focus mainly on technology-enhanced asynchronous written exchanges that are beginning to replace face-to-face peer assessment (van Popta et al, 2017). In his systematic review, Ashenafi (2017) reported that very few studies in the sample explored synchronous dialogic interactions within peer assessment. In a rare example, Liu, Lan and Zhang (2018) used an online platform for peer assessment in a Chinese business course and found that synchronous group-based discussions surrounding differences in peer assessment judgements helped students in their learning.

Student preference for, and satisfaction with, online peer assessment has been reported. In a Turkish computing course that used blogs to facilitate peer review, Gurer (2020) reported that students welcomed the ongoing interaction with their peers and the chance to experience different perspectives that were not always available in other face-to-face classes. Blogs helped the students to interrogate and evaluate other students' work, reflect on how this related to their own work and provide peers with feedback. One useful affordance of technology-mediated software is that lecturers can add specific review questions, which ask students to indicate how useful the peer feedback was

and whether they enacted this in their subsequent work (Patchan et al, 2016; Misiejuk et al, 2020). Such questions may prompt students to think more about feedback enactment.

The online format seems to make some students feel more comfortable in offering more challenging and critically focused feedback and assessments of peers' work. McCarthy (2017) attributed students' preference for online peer assessment to the fact that face-to-face interactions can often be less critical in tone due to fear of upsetting people or losing friendships with certain peers (Kilickaya, 2017). In an Iranian online English for Academic Purposes course, students published their ongoing draft work through online wikis and received peer and lecturer feedback (Asoodar, Atai and Vaezi, 2016). After addressing the peer feedback, they published the work to a cohort-wide blog. Students were satisfied with the opportunity to compare their work to peers. The anonymous review process allowed for more critical feedback to be given without fear of seeing the recipient's emotional reaction. Although the source of the feedback was anonymous, everyone could see the feedback that was generated. Asoodar et al (2016) contend the online blog also enabled students to engage in an asynchronous dialogue regarding the meaning of the feedback, which was visible to all students. Meaning making in this regard was developed within peer interactions and spread to the wider group.

In sum, online peer assessment tools offer affordances that overcome some of the barriers associated with face-to-face peer assessment. It appears to be effective because it creates a more socially orientated interactive environment (Li et al, 2019) while anonymity facilitates greater student criticality and self-evaluation (Zheng et al, 2020)

5.2 Peer feedback

Peer feedback occurs during peer assessment, but it can also occur in interventions that don't involve a teacher-created rubric or assignment of grades or scores. Research on peer feedback was well established prior to this review period. Research since 2016 also confirms its effectiveness overall. Experimental research has demonstrated that when students have opportunities to generate and receive peer feedback compared to those who do not, they outperform them on writing gains (Simpson and Clifton, 2016; Zhang and McEneaney, 2019).

In the past five years – even in peer feedback rather than peer assessment situations as discussed in section 5.1.1 above – students questioned the expertise of peers as a source of feedback. In Jeon's (2018) study, 29 Korean university students enrolled in Practical English Writing course from various years and subject discipline followed a four-stage process of pre-writing, drafting, revising and editing. After each stage, students received feedback from either their peers or the lecturer. They then refined their work in light of this feedback prior to submitting for summative grading. Seeing their peers' work-in-progress helped them to identify their peers' mistakes and then go back to their own work and notice similar issues. Students' post-test anxiety scores were also significantly lower than their pre-test anxiety score. However, the benefits may have come mainly from seeing peers' work, as in the exemplar research in section 4.9 above, rather than from peers' feedback per se. Jeon (2018) stressed that students raised concerns about peers' expertise and were more likely to trust the feedback from the lecturers.

Impacts of Higher Education Assessment and Feedback Policy and Practice on Students: A Review of the Literature 2016-2021
Edd Pitt and Kathleen M Quinlan

Even a peer marker's self-rated expertise can influence students' perceptions of the feedback. In a Dutch context, Dijks, Brummer and Kostons (2018) investigated student perceptions of anonymous peer feedback when only the self-perceived expertise of the reviewer was provided. Forty teacher training students gave their peers anonymous feedback on a piece of draft work using a lecturer-developed rubric. Students made judgements relating to the quality of the work and indicated their own self-perceived expertise in providing feedback. When the work was sent back to the author, the recipient only saw the feedback and the reviewer's self-perception of expertise. Dijks et al (2018) found that students preferred feedback from students with higher self-rated expertise, but it made no difference to enactment of that peer feedback. Dijks et al (2018) point out that we must be careful interpreting results from one episode of peer review and associated feedback. Throughout our discussion of feedback, we have emphasised the importance of students' multiple engagements with feedback interventions.

The issue of students enacting peer feedback was apparent in Jurkowski's (2018) experimental study with German educational psychology students. Students received peer feedback on their draft work before refining for summative submission. Students in the control group received no guidance, but students in the experimental condition received question prompts designed to guide and assist them in processing the peer feedback. Those in the experimental group considered the received feedback more and integrated more of the correct peer feedback comments and less of the incorrect than the control group. However, this still did not translate into a significant increase in students' overall performance between drafts, as they appeared to ignore 50% of the peer feedback comments they received.

5.2.1 Training and peer feedback

Training in providing peer feedback has been shown to improve the quality of peer feedback and students' wider feedback behaviours (Carless and Boud, 2018; Huisman et al, 2018). Both Simpson and Clifton (2016) and Zhang and McEneaney (2019) stressed how important opportunities for structured training are prior to engaging with peer feedback opportunities. Zhang and McEneaney (2019) argued that their training helped students to develop the level of criticality needed to readily recognise issues within their peers' work and then apply this to their own work. In a Turkish EFL class where grammar error correction is the most commonly generated peer feedback, Uymaz (2019) reported that students were only given two hours of training. The training consisted of showing students exemplars of work that had been previously marked by the lecturers. In four separate assignments, the first written draft received peer feedback and students could redraft. Despite the limited prior training, peers revised 69% of the errors correctly in their second drafts (Uymaz, 2019).

Effective training consists of students actively engaging and learning from 'doing' rather than being told. As Huisman et al (2018) argued, students become very familiar with the meaning and application of assessment criteria if, when they are constructing the peer feedback, they have to consider these criteria. In Taiwanese dance education, Hsia, Huang and Hwang (2016) reported that those engaging in peer commenting via an online dance feedback system provided their opinions and interpreted others' comments without the need for additional guidance or assistance. Students performed better

following receipt of peer feedback and associated grades compared to just peer feedback or just grades alone. The feedback generated by students in this experimental condition was similar to that of the lecturer.

5.2.2 Anonymity in peer feedback

During the review period, researchers focused on whether peer feedback was more effective under anonymous or non-anonymous conditions. Throughout section 4 above, we have seen that teacher feedback is more effective when it is part of an ongoing relationship (section 4.4), offers opportunities for dialogue over time, is more expansive and personal as in video feedback (section 4.7.2) and is elaborative (section 4.8.1). Anonymity mitigates all of those features (section 4.4). Thus, it may be counter-intuitive that peer feedback would benefit from anonymity, though we have seen how it is an advantage in technology-enhanced peer assessment (section 5.1.2).

Evidence on peer feedback more generally also shows some support for anonymity, at least as students learn how to engage in peer review. Within second language writing in Holland, van den Bos and Tan (2019) reported that it was a combination of anonymity and the four weeks of training their students received which helped them to readily understand and be able to generate good feedback for their peers. Students in the anonymous peer feedback group received higher final grades than students in the non-anonymous group. As discussed in section 5.1.2 above, anonymity may have enabled students to feel comfortable giving more critical feedback, underpinning performance improvements. Within Belgian teacher education, Rotsaert, Panadero and Schellens (2018) assigned 46 Belgian third year education studies students to an anonymous peer assessment condition early in a module. Students engaged in several peer assessment activities for four weeks, beginning with anonymous feedback before transitioning to a non-anonymous peer feedback structure. After switching to the non-anonymous condition, Rotsaeart et al (2018) found the quality of student feedback did not decrease. They suggested initial anonymity can scaffold and improve the quality of feedback over time prior to non-anonymous interactions. Thus, student training, as discussed in section 5.2.1, may be essential in shifting from anonymised to non-anonymised conditions.

5.2.3 Dialogue in peer feedback on and offline

Despite the evidence on anonymity above, studies in the review period also showed that students appreciate meaning making through dialogue in peer feedback, just as they do with teacher feedback. Zhu and Carless (2018) investigated the role of oral dialogue within peer feedback with second year EFL students in China. Feedback providers liked hearing from the recipient about their feedback, and receivers welcomed the negotiation of meaning through interactions. In Schillings et al's (2021) small study of eight Dutch biomedical students, face-to-face dialogue increased students' sense of responsibility in the feedback process. Providers of peer feedback thought their role in the dialogue was central to the receiver's understanding their written feedback. Receivers were positive about using this feedback within subsequent redrafts of their work prior to submission. But, in both Schillings et al (2021) and Zhu and Carless (2018), students reported concerns about having sufficient expertise

and felt they needed additional input from their lecturer, echoing findings we showcased in sections 5.1 and 5.2.1 about the importance of training.

While technology can facilitate asynchronous, written peer feedback opportunities (see sections 5.1.2 and 5.2.2 above) students may still benefit from synchronous, verbal dialogues. In China, Zheng, Cui, Li and Huang (2018) conducted a randomised control trial with 64 undergraduates to test the impact of peer dialogue after written feedback. Twice students wrote short essays and posted them to an online platform that randomly assigned three peers to review their work using a teacher-created rubric. Students in the experimental group engaged in follow-up synchronous dialogue with their assessors. Those in the control group participated in no dialogue. Those in the experimental group significantly improved their writing performance compared to students in the control group. The inclusion of the synchronous discussion helped students clarify their understanding of the feedback and how they might enact it.

Technologies can also provide opportunities for greater student agency than prescribed in-person peer feedback activities. For example, in Kay, Hardy and Galloway's (2020) UK study, 3,000 students across six large undergraduate science courses used the popular free online question generator PeerWise to provide peers with feedback on the quality of their questions. Through the software, students could discuss questions and feedback with their peers, self-pace their learning and decide how they wanted to engage with peers. In a study with EFL students in Malaysia, Saeed et al (2018) quantified students' actual engagement with and enactment of reciprocal asynchronous peer feedback. They found that students mostly attended to revision orientated comments, consistent with findings in 4.8.1 suggesting students prefer and benefit more from elaborative feedback.

Finally, in a novel approach, Zamberlan and Wilson (2017) enlisted final year students as peer tutors within a first-year undergraduate design studio module in Australia. The peer tutors helped their less experienced peers develop their design process skills and visual communication techniques by discussing their own experiences. Less experienced students were exposed to different sources of feedback throughout the module, which gave them agency in enactment decisions. Students commented that these more experienced peer tutors were more approachable, had longer to engage with them and entered into meaningful dialogic interactions that supported their learning.

5.3 Benefits to providers versus receivers of peer assessment/feedback

Previous peer feedback research has generally found that students who generate feedback benefit more than students who receive it (Nicol, Thomson and Breslin, 2014). During the timeframe of our review, researchers have continued to demonstrate benefits for the provider (Cao et al, 2019; Huisman et al, 2018; van Popta et al, 2017), as well as a range of learning benefits for the receiver (Cao, Yu and Huang, 2019; Huisman et al, 2018; Jurkowski, 2018). Huisman et al (2018) found that Dutch first year education and child development students who provided peer feedback considered this beneficial to their writing performance. The process of providing peer feedback prompted them to reflect on the task-specific processes and criteria, initiating changes to their own work in light of it. That is, giving peer feedback made them better at self-assessment.

Other studies since 2016 directly compared receiving feedback with providing feedback (eg Ion et al, 2019; Schunn, Godley and DeMartino, 2016), concluding that while both are useful for students, improvements in learning and performance are more frequently associated with providing than receiving (Wu and Schunn, 2021). Likewise, providers of online peer feedback in van Popta et al's (2017) study improved in skill acquisition, critical reflection and knowledge construction. Chen et al (2020) developed an online peer feedback system In Taiwan called the e-musical theatre system, which provided example guidance videos, teacher-created rubrics, and a place to carry out the peer assessment. They created three groups: 1) peer scoring group, 2) peer commenting group, and 3) peer commenting and scoring group. Using lecturer-created rubrics, group one evaluated peers' videos by scoring, group two evaluated peers' videos by giving comments and group three evaluated peers' videos by providing both comments and scores. Students in group three (peer commenting and scoring) performed better than those in the other two groups. This was attributed to and corroborated by their VLE records that showed group three had engaged more fully than the other groups in supplementary guidance surrounding the use of the assessment rubrics. The students had also applied these more carefully when giving their feedback, perhaps because they needed to offer both scores and comments. Thus, giving peer assessment with comments prompted greater learning from and about the assessment standards, which benefitted their own performance.

Peer feedback receivers also benefit as long as they engage with the feedback, re-emphasising a key principle from the teacher feedback literature (section 4) about students' role in feedback interpretation and use. Zong, Schunn and Wang (2021) reported that in a Chinese undergraduate computer programming course, students' task-related improvements were mostly related to the depth of comments rather than the volume of comments they received from peers. Crucially, regardless of the depth, students needed to cognitively engage and enact these peer feedback comments to improve. Nonetheless, Zong et al (2021) highlighted that while both parties benefitted from the process, the greatest gains were for the provider.

Both providers and receivers of peer feedback reported increases in molecular life sciences related knowledge and high levels of satisfaction in Noroozi and Mulder's (2017) study with Dutch undergraduate students. Huisman, Saab, van den Broek et al (2019) reported that receivers of peer feedback improved their writing performance when they were willing to enact their peers' feedback. Alanson and Robles (2017) designed learning activities around students' ongoing e-portfolios in an American careers' education programme. Thirty-nine students completed a pre and post intervention confidence survey, shared their e-portfolios with peers, received and generated peer feedback and redrafted their e-portfolios. Sixty-seven percent of the students indicated they had readily engaged with the peer feedback process, 80% reported using the peer feedback to improve their final e-portfolio and 92% of students also reported using the formative feedback from lecturers to improve their work. Although Alanson and Robles (2017) did not analyse student work for enactment, the results suggest that when meaningful formative feedback and enactment opportunities exist, students engage and use feedback from both their lecturers and peers.

However, Huang (2018) has questioned the longevity of benefits that peer feedback provides beyond, for example, the immediate task. This finding seems to chime with Noroozi and Mulder's (2017) argument that generating peer feedback relating to a specific task may only be applicable for that task or closely related tasks. These findings are consistent with evidence presented earlier on teacher feedback, in which we argued for the importance of authentic tasks that reflect key disciplinary or professional skills and feedback that is focused on those transferable skills (eg writing, argumentation, problem-solving, decision-making).

6 Conclusions and recommendations

6.1 Conclusions

Discussion surrounding authentic assessment was a prominent feature in our review suggesting that it has the potential to reflect real-world demands, challenge students, help them to evaluate their own performance and promote engagement, satisfaction and improvement. In this review we have also discussed numerous well-established areas within the assessment and feedback literature. It is clear that these established processes and practices have been enhanced by the use of new technologies. We have discussed both the positive and negative effects technology can have on student satisfaction and learning outcomes. The feedback paradigm shift discussed in section four of the review emphasised students taking more active roles in generating and enacting feedback from multiple sources.

The Advance HE Transforming Assessment in Higher Education Framework (2015) was used to guide our review. At its centre is transforming assessment. We concluded from this integrated review that feedback should also be at the centre of the Framework. That is, we argue that assessment and feedback should be considered together to promote student success. The three focal areas of innovative assessment, feedback practices and self and peer assessment usefully structured this review. A strength of this review was bringing together literature on these three areas, which are often treated separately. Therefore, we suggest revisiting and updating the six key tenets for transforming assessment to integrate these three focal areas and reflect recent evidence. Finally, we would suggest that the process of transformation explicitly seeks to integrate assessment and feedback design.

6.2 Recommendations

In this section, we summarise the key lessons from this report in the form of recommendations to practitioners and policymakers that will support the implementation of evidence-based practice drawn from the literature from 2016-2021. We also offer recommendations to researchers to ensure the field generates high-quality evidence that will continue to advance practice in assessment and feedback in HE.

6.2.1 Recommendations for practitioners

1 **Aim to incorporate at least one high-impact practice into each student's experience.** Integrative capstones, work-integrated learning, undergraduate research experiences, and programme-level reflective portfolios are all examples of practices that embed public demonstrations of skills and knowledge (ie rich assessment) and feedback along with six other key educational principles. Combining powerful educational principles and practices creates high impact.

2 **Use, evaluate and refine authentic assessments.** Authentic assessments are 1) realistic, 2) challenging and 3) support students' understanding of quality standards and their ability to self-assess. They promote student satisfaction, engagement and achievement.

   a. Ensure assessments and related feedback are aligned with complex, higher order intended learning outcomes.

   b. Prepare students for new assessment demands. Students are often nervous when facing new types of assessments; they benefit from early, low-stakes exposure and preparation.

   c. Scaffold, scaffold, scaffold. When engaging with complex problems and new processes (such as collaboration or cross-cultural teamwork), students need prompts, structures, templates, guides, practice, informational supports, resources and emotional supports.

3 **Attend to the specific evidence-based requirements of each type of authentic assessment.** For example:

   a. Be cautious when considering high-tech or high-fidelity simulations, since cheaper, lower-tech options might be equally effective.

   b. Attend to the design of related learning activities in and around high-tech simulations, as these are essential to ensure that learning is harnessed from the unique feedback opportunities afforded.

4 **Incorporate carefully designed groupwork across each student's programme.** Group assessments can enable the practice of valuable collaboration skills, improve student performance, support the integration of diverse students, address EDI issues and reduce marking burdens on staff. To ensure groupwork fulfils its potential, prepare students for collaboration through scaffolding cross-cultural teamwork skills and processes, assign groups (don't rely on student self-selection), create group goals, and ensure individual accountability in final products. Given the centrality of the teamwork and communication skills involved, scaffolding might occur over the course of a programme, with multiple opportunities for collaborative projects.

5 **Shift the culture of feedback to emphasise and support students' use of it.** Feedback is more than 'telling'; students need to understand it and use it. Increasing students' sense of responsibility in the feedback process and promoting more proactive reading, interpretation, and enactment (use) of feedback has been shown to improve students' learning. Lecturers should foster a culture of shared responsibility around feedback. Feedback interventions can be used across a students' programme to build such a culture.

6 **Treat feedback as part of an ongoing, positive relationship, not a one-off event.**

   a. Opportunities for dialogues about feedback with lecturers and peers improves students' learning outcomes. Feedback interventions are particularly effective when they occur within formative opportunities (ie drafts), so students can act on feedback to improve their work before submitting it for a mark.

   b. Design multiple opportunities across a programme for teacher and peer feedback. It is better for students to generate than receive peer feedback; generating feedback can improve students' enactment of feedback and subsequent outcomes.

   c. Anonymous marking misses opportunities to build productive feedback dialogues and ongoing educational relationships.

7 **Help students understand assessment criteria and standards.** Assessment criteria and standards aligned to authentic learning outcomes define quality work in the field. Thus, they are central to learning in the discipline or profession, not just something that is done to earn a mark. Learning and quality standards can be achieved by spending time within curricula designing and reviewing assessment criteria or rubrics and working with and applying criteria against exemplars or students' actual work-in-progress. Practice with assessing and giving feedback to peers helps students internalise quality standards.

8 **Be aware that students come with different experiences of, expectations of, and preparedness to learn from feedback.** Feedback experiences and expectations may be influenced by different cultural or educational backgrounds. Students also have different levels of emotional maturity needed to cope with constructive feedback. Feedback that enhances students' self-confidence can help them cope with the higher expectations teachers may communicate in feedback.

9 **Increase opportunities for peer assessment and peer feedback.** These activities can help students develop the necessary skills to engage in self-assessment that will help them in their work. They work best when there are repeated opportunities.

10 **Scaffold support prior to and during episodes of peer feedback and peer assessment.** Practitioners should design peer assessment and feedback opportunities that promote dialogue about standards, quality and enactment of feedback so they improve students' learning. Practitioners should look beyond one instance and use multiple opportunities to build upon each other, across a programme of study.

11 **Evaluate educational technologies in relation to the educational goals they support and principles they afford. Make use of the affordances of immediate feedback, personalised feedback, video and screencast feedback and anonymous, asynchronous peer assessment.** Technologies can effectively support gamification, peer interactions, community-building, wider sharing of student work, greater creativity in assessment products, greater authenticity in processes or products, and immediate, personalised feedback. Benefits accrue from using these affordances of technology. Attend to whether a technology is simply replicating a low-tech option or whether it presents opportunities for incremental improvements through these educational principles or more radical improvements to educational processes. We recommend that educators attend to and design for the underlying educational principles that are likely responsible for associated gains in satisfaction, engagement or performance.

12 **Explore social, ethical and behavioural interventions in relation to academic integrity, not just technical or legal solutions.** Concerns about academic integrity have come to the fore particularly in relation to online exams used during the pandemic, though other formats are also susceptible to plagiarism and academic misconduct.

6.2.2 Recommendations for policymakers

To ensure that policies support, encourage and reward educators for implementing the recommendations above, we recommend the following.

1 **Higher education institutions should review their assessment and feedback policies to ensure that they are consistent with the recommendations to practitioners above.**

2 **Review the assessment and feedback questions that are asked in student evaluations of teaching at all levels to ensure they are consistent with the practitioner recommendations above.** Assessment questions might focus, for example, on students' satisfaction that assessments reflected real-world demands, were challenging, promoted their understanding of what constitutes a good performance in their field, and fostered teamwork and communication skills. Feedback questions might emphasise student satisfaction with assistance in interpreting their feedback, opportunities to generate feedback for themselves or peers, and opportunities to enact feedback received. Such changes would reflect current understanding of effective assessment and feedback practices and encourage educators to implement the practitioner recommendations above.

3 **Create quality assurance codes that will encourage higher education institutions and practitioners to implement the recommendations above.**

4 **Professional bodies should engage with higher education providers to re-consider assessment formats.** Professional bodies should consider how authentic assessments and or alternative assessment formats could be used within subjects that currently rely on more traditional written knowledge tests.

6.2.3 Recommendations for assessment and feedback researchers

To advance our understanding of demonstrable effects of assessment and feedback practices on student outcomes, researchers should:

1 **Move beyond student satisfaction-based studies.** While satisfaction is a metric commonly used in student evaluations of teaching or experience surveys, it does not always relate to or confirm student learning.

2 **Develop and consistently use more robust measures of engagement-related outcomes.** Engagement is an important but very broad concept. The field would benefit from greater conceptual clarity that could guide more robust and comparable measurement.

3 **Use external measures of performance.** Many of the studies relied on students' self-reported competencies or learning, which is expedient but often does not provide a good reflection of actual learning. Researchers should consider quasi-experimental studies. Control-based designs will help tease out policies and practices that are likely to have the greatest impact on students, particularly their performance. Comparing one cohort to another, while imperfect, may be a more feasible approach to these kinds of studies than randomising students into groups.

4 **Use multiple data sources.** Research which uses data from multiple sources such as observations, learning analytics and student performance outcomes, may afford more sophisticated mapping of student learning behaviours. Drawing together research across different disciplinary domains will help to highlight key lessons that can be learned.

5 **Consider equity and inclusivity.** Very few studies sought to determine differential impacts on historically disadvantaged or marginalised students. To close attainment gaps and ensure educational equity, researchers need to consider who benefits most from specific assessment and feedback policies. There was surprisingly little on the demonstrable impacts of inclusive assessments. More work needs to be done to ensure equivalencies of different formats for assessing the same intended learning outcomes.

Acknowledgements

We thank the Advance HE Advisory Group members for their feedback on an earlier draft of this report and Veronica Rovagnati and Lauren McGuiness for their research assistance.

7. References

Aamlid, H and Tveit, B (2021) 'Simulation as a joint learning activity in clinical placement-interaction between first-year nursing students and qualified nurses', *Journal of Clinical Nursing,* 31 (1-2): 250-261. Available at: doi.org/10.1111/jocn.15903

Adachi, C, Tai, J and Dawson, P (2018) 'A framework for designing, implementing, communicating and researching peer assessment', *Higher Education Research and Development,* 37 (3): 453-467. Available at: doi.org/10.1080/07294360.2017.1405913

Adams, A, Wilson, H, Money, J, Palmer-Conn, S and Fearn, J (2020) 'Student engagement with feedback and attainment: the role of academic self-efficacy', *Assessment and Evaluation in Higher Education,* 45 (2): 317-329. Available at: doi.org/10.1080/02602938.2019.1640184

Adams, C J (2020) 'A constructively aligned first-year laboratory course', *Journal of Chemical Education,* 97 (7): 1863. Available at: doi.org/10.1021/acs.jchemed.0c00166

Ajjawi, R and Boud, D (2018) 'Examining the nature and effects of feedback dialogue', *Assessment and Evaluation in Higher Education,* 43 (7): 1106-1119. Available at: doi.org/10.1080/02602938.2018.1434128

Akashe, Z B, Esfahani, A R N, Nili, M R and Tabatabaei, S M S (2020) 'Investigating the desirable assessment methods of the performance of the BA law students during internships', *International Journal of Work-Integrated Learning,* 21 (1): 13-24.

Akhtar, M and Saeed, M (2020) 'Assessing the effect of agree/disagree circles, exit ticket and think-pair-share on students' academic achievement at undergraduate level', *Bulletin of Education and Research,* 42 (2): 81-96.

Alahmadi, N, Alrahaili, M and Alshraideh, D (2019) 'The impact of the formative assessment in speaking test on Saudi students' performance', *Arab World English Journal,* 10 (1): 259-270. Available≈at: doi.org/10.24093/awej/vol10no1.22

Alanson, E R and Robles, R A (2017) 'Using electronic portfolios to explore essential student learning outcomes in a professional development course', *Asia-Pacific Journal of Cooperative Education,* 17 (4): 387-397.

Al-Azzawi, A, Nagavi, B G, Hachim, M Y and Mossa, O H (2016) 'The implementation and development of an objective structured clinical examination in the community pharmacy course of a select Gulf-region academic institution' (Ras Al Khaimah College of Pharmaceutical Sciences): A pilot study', *Innovations in Education and Teaching International,* 53 (1): 60-72. Available at: doi.org/10.1080/14703297.2013.866328

Alexander, P A (2019) 'Individual differences in college-age learners: The importance of relational reasoning for learning and assessment in higher education', *British Journal of Educational Psychology,* 89 (3): 416-428. Available at: doi.org/10.1111/bjep.12264

Alfalagg, A R (2020) 'Impact of teacher-student writing conferences on frequency and accuracy of using cohesive devices in EFL students' writing', *Asian-Pacific Journal of Second and Foreign Language Education,* 5. Available at: doi.org/10.1186/s40862-020-00104-z

Ali, N, Ahmed L and Rose, S (2017) 'Identifying predictors of students' perception of and engagement with assessment feedback', *Active Learning in Higher Education.* Available at: doi.org/10.1177/1469787417735609

Allagui, B (2021) 'Writing a descriptive paragraph using an augmented reality application: An evaluation of students' performance and attitudes', *Technology Knowledge and Learning,* 26 (3): 687-710. Available at: doi.org/10.1007/s10758-019-09429-2

Alzouebi, K (2020) 'Electronic portfolio development and narrative reflections in higher education: Part and parcel of the culture?', *Education Information Technology,* 25 (2): 997-1011. Available at: doi.org/10.1007/s10639-019-09992-2

Amasha, M A, Abougalala, R A, Reeves, A J and Alkhalaf, S (2018) 'Combining online learning and assessment In synchronization form', *Education and Information Technologies,* 23 (6): 2517-2529. Available at: doi.org/10.1007/s10639-018-9728-0

Ambikairajah, A and Tisdell, C C (2019) 'E-examinations and the student experience regarding appropriateness of assessment and course quality in science and medical science', *Journal of Educational Technology Systems,* 47 (4): 460-478. Available at: doi.org/10.1177/0047239518822016

Anderson, L W and Krathwohl, D R (2001) *A taxonomy for learning, teaching and assessing: A revision of Bloom's Taxonomy of Educational Objectives.* Complete Edition. New York: Longman.

Andrade, M S (2018) 'Learning gain – A US perspective', *Higher Education Pedagogies,* 3 (1): 46-48. Available at: doi.org/10.1080/23752696.2018.1508361

Anson, C M, Dannels, D P, Laboy J I and Carneiro, L (2016) 'Students' perceptions of oral screencast responses to their writing: Exploring digitally mediated identities', *Journal of Business and Technical Communication,* 30 (3): 378-411. doi.org/10.1177/1050651916636424

Ashenafi, M M (2017) 'Peer-assessment in higher education – twenty-first century practices, challenges and the way forward', *Assessment and Evaluation in Higher Education,* 42 (2): 226-251. Available at: doi.org/10.1080/02602938.2015.1100711

Asoodar, M, Atai, M R and Vaezi, S (2016) 'Blog-integrated writing with blog-buddies: EAP learners' writing performance', *Journal of Educational Computing Research,* 54 (2): 225-252. Available at: doi.org/10.1177/0735633115615588

Azmat, G, Bagues, M, Cabrales, A and Iriberri, N (2019) 'What you don't know…can't hurt you? A natural field experiment on relative performance feedback in higher education', *Management Science,* 65 (8): 3714-3736. Available at: doi.org/10.1287/mnsc.2018.3131

Babo, R, Babo, L V, Suhonen, J T and Tukiainen, M (2020) 'E-assessment with multiple-choice questions: A 5-year study of students' opinions and experience', *Journal of Information Technology Education: Innovations in Practice,* 19: 1-29. Available at: doi.org/10.28945/4491

Bacchus, R, Colvin, E, Knight, E B and Ritter, L (2020) 'When rubrics aren't enough: Exploring exemplars and student rubric co-construction', *Journal of Curriculum and Pedagogy,* 17 (1): 48-61. Available at: doi.org/10.1080/15505170.2019.1627617

Bajada, C, Singh, S, Jarvis, W and Trayler, R (2021) 'The use of threshold concepts to support student learning through assessment – a case for renewing public trust in business education and qualifications', *Higher Education Research and Development,* 40 (4): 676-691. Available at: doi.org/10.1080/07294360.2020.1775558

Bangera, G and Brownell, S E (2014) 'Course-based undergraduate research experiences make research more inclusive', *CBE – Life Sciences Education,* 13 (4): 602-606. Available at: doi.org/10.1187/cbe.14-06-0099

Basith, A, Rosmaiyadi, S, Triani, N and Fitri, F (2020) 'Investigation of online learning satisfaction during Covid 19: In relation to academic achievement', *Journal of Educational Science and Technology,* 6 (3): 265-275. Available at: https://ojs.unm.ac.id/JEST/article/view/14803Bearman

M, Boud, D and Ajjawi, R (2020) 'New directions for assessment in a digital world', in Bearman, M, Dawson, P, Ajjawi, R, Tai, J and Boud, D (eds) *Re-Imagining University Assessment in a Digital World.* Cham: Springer, pp 7-18. doi.org/10.1007/978-3-030-41956-1_2

Bearman, M, Dawson, P, Ajjawi, R, Tai, J and Boud, D (eds) (2020) *Re-Imagining University Assessment in a Digital World. The Enabling Power of Assessment.* Cham: Springer. Available at: link.springer.com/book/10.1007/978-3-030-41956-1Beccaria, L, Kek, M and Huijser, H (2019) 'Using 'just in time' online feedback to improve first year undergraduate nursing students' essay writing performance', *Journal of University Teaching and Learning Practice,* 16 (4). Available at: doi.org/10.53761/1.16.4.7

Belboukhaddaoui, I and van Ginkel, S (2019) 'Fostering oral presentation skills by the timing of feedback: An exploratory study in virtual reality', *Research on Education and Media,* 11 (1): 25-31. Available at: doi.org/10.2478/rem-2019-0005

Bijami, M, Kashef, S H and Nejad, M S (2013) 'Peer feedback in learning English writing: Advantages and disadvantages', *Journal of Studies in Education,* 3 (4): 91-97. Available at: doi.org/10.5296/jse.v3i4.4314

Black, P and Wiliam, D (2018) 'Classroom assessment and pedagogy', *Assessment in Education: Principles, Policy and Practice,* 25 (6): 551-575. Available at: doi.org/10.1080/0969594X.2018.1441807

Bloxham, S and Campbell, L (2010) 'Generating dialogue in assessment feedback: Exploring the use of interactive cover sheets', *Assessment and Evaluation in Higher Education,* 35 (3): 291-300. Available at: doi.org/10.1080/02602931003650045

Boud, D and Soler, R (2016) 'Sustainable assessment revisited', *Assessment and Evaluation in Higher Education,* 41 (3): 400-413. Available at: doi.org/10.1080/02602938.2015.1018133

Brannick, M T, Erol-Korkmaz, H and Prewett, M T (2011) 'A systematic review of the reliability of objective structured clinical examination scores', *Medical Education,* 45 (12): 1181-1189. Available at: doi.org/10.1111/j.1365-2923.2011.04075.x

Brazhkin, V and Zimmerman, H (2019) 'Students' perceptions of learning in an online multiround business simulation game: What can we learn from them?', *Decision Sciences Journal of Innovative Education,* 17 (4): 363-386. Available at: doi.org/10.1111/dsji.12189

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

Bruwer, A and Ontong, J M (2020) 'Early assessment as a predictor of academic performance: An analysis of the interaction between early assessment and academic performance by first-year accounting students at a South African university', *South African Journal of Higher Education,* 34 (4): 11-26. Available at: doi.org/10.20853/34-4-3607

Budner, D (2020) 'The design and regular use of integrated problems within a general chemistry II course', *Journal of Chemical Education,* 97 (8): 2140-2146. Available at: doi.org/10.1021/acs.jchemed.9b01160

Buitrago-Flórez, F, Reyes, J C, Rincón, R, Hernández, C, Galvis, F A and Ángel, C (2020) 'Engaging in homework development: TARSIS platform as an innovative learning methodology', *Australasian Journal of Educational Technology,* 36 (3): 147-162. Available at: doi.org/10.14742/ajet.5865

Bulut, O, Cutumisu, M, Aquilina, A M and Singh, D (2019) 'Effects of digital score reporting and feedback on students' learning in higher education', *Frontiers in Education,* 4 (65): 1-16. Available at: doi.org/10.3389/feduc.2019.00065

Butler, A C and Woodward, N R (2018) 'Toward consilience in the use of feedback to promote learning: A review of the literature', *Psychology of Learning and Motivation,* 69 (1). Available at: doi.org/10.1016/bs.plm.2018.09.001

Butler-Henderson, K and Crawford, J (2020) 'A systematic review of online examinations: A pedagogical innovation for scalable authentication and integrity', *Computers and Education,* 159. Available at: doi.org/10.1016/j.compedu.2020.104024

Cantero-Chinchilla, F, Díaz-Martín, C, García-Marín, A P and Estévez, J (2020) 'Innovative student response system methodologies for civil engineering practical lectures', *Technology, Knowledge and Learning,* 25: 835-852. Available at: doi.org/10.1007/s10758-019-09410-z

Cao, Z H, Yu, S L and Huang, J (2019) 'A qualitative inquiry into undergraduates' learning from giving and receiving peer feedback in L2 writing: Insights from a case study', *Studies in Educational Evaluation,* 63: 102-112. Available at: doi.org/10.1016/j.stueduc.2019.08.001

Carless, D (2006) 'Differing perceptions in the feedback process', *Studies in Higher Education,* 31 (2): 219-233. Available at: doi.org/10.1080/03075070600572132

Carless, D (2015) *Excellence In University Assessment. Learning From Award Winning Practice.* London: Routledge. Available at: doi.org/10.4324/9781315740621

Carless, D and Boud, D (2018) 'The development of student feedback literacy: Enabling uptake of feedback', *Assessment and Evaluation in Higher Education,* 43 (8): 1315-1325. Available at: doi.org/10.1080/02602938.2018.1463354

Carless, D and Chan, K K H (2017) 'Managing dialogic use of exemplars', *Assessment and Evaluation in Higher Education,* 42 (6): 930-941. Available at: doi.org/10.1080/02602938.2016.1211246

Carless, D, Salter, D, Yang, M and Lam, J (2011) 'Developing sustainable feedback practices', *Studies in Higher Education,* 36 (4): 395-407. Available at: doi.org/10.1080/03075071003642449

Carmichael, M C, St Clair, C, Edwards, A M, Barrett, P, McFerrin, H, Davenport, I, Awad, M, Kundu, A and Ireland, S K (2016) 'Increasing URM undergraduate student success through assessment-driven interventions: A multiyear study using freshman-level general biology as a model system', *CBE – Life Sciences Education,* 15 (3). Available at: doi.org/10.1187/cbe.16-01-0078

Carter, A (2020) 'How well do vocational qualifications help prepare technician engineering students for undergraduate study?', *European Journal of Engineering Education,* 45 (6): 854-878. Available at: doi.org/10.1080/03043797.2019.1677561

Cavaleri, M, Kawaguchi, S, Di Biase, B and Power, C (2019) 'How recorded audio-visual feedback can improve academic language support', *Journal of University Teaching and Learning Practice,* 16 (4). Available at: doi.org/10.53761/1.16.4.6

Chalmers, C, Mowat, E and Chapman, M (2018) 'Marking and providing feedback face-to-face: Staff and student perspectives', *Active Learning in Higher Education,* 19 (1): 35-45. Available at: doi.org/10.1177/1469787417721363

Chang, K (2016) 'Exploring teaching/learning activities for sight translation: Effectiveness from students' perspectives', *Taiwan Journal Of TESOL,* 13 (2): 71-98.

Chen, I-C, Hwang, G-J, Lai, C-L and Wang, W-C (2020) 'From design to reflection: Effects of peer-scoring and comments on students' behavioural patterns and learning outcomes in musical theater performance', *Computers and Education,* 150: 103856. Available at: doi.org/10.1016/j.compedu.2020.103856

Chong, S W (2021) 'Reconsidering student feedback literacy from an ecological perspective', *Assessment and Evaluation in Higher Education,* 46 (1): 92-104. Available at: doi.org/10.1080/02602938.2020.1730765

Christie, H and Morris, N (2021) 'Assessment and emotion in higher education: The allure of blogging', *Research in Post-Compulsory Education,* 26 (2): 148-163. Available at: doi.org/10.1080/13596748.2021.1909922

Coates, H (2016) 'Assessing student learning outcomes internationally: Insights and frontiers', *Assessment and Evaluation in Higher Education,* 41 (5): 662-676. Available at: doi.org/10.1080/02602938.2016.1160273

Coates, H and Zlatkin-Troitschanskaia, O (2019) 'The governance, policy and strategy of learning outcomes assessment in higher education', *Higher Education Policy,* 32 (4): 507-512. Available at: doi.org/10.1057/s41307-019-00161-1

Cohen, A and Singh, D (2020) 'Effective student feedback as a marker for student success', *South African Journal of Higher Education,* 34 (5): 151-165. Available at: doi.org/10.20853/34-5-4259

Conejo, R, Garcia-Viñas, J I, Gastón, A and Barros, B (2016) 'Technology-enhanced formative assessment of plant identification', *Journal of Science Education and Technology,* 25 (2): 203-221. Available at: doi.org/10.1007/s10956-015-9586-0

Consorti F, Mancuso, R, Nocioni, M and Piccolo, A (2012) 'Efficacy of virtual patients in medical education: A meta-analysis of randomized studies', *Computers in Education,* 59 (3): 1001-1008. Available at: doi.org/10.1016/j.compedu.2012.04.017

Cook, B R and Babon, A (2017) 'Active learning through online quizzes: Better learning and less (busy) work', *Journal of Geography in Higher Education,* 41 (1): 24-38. Available at: doi.org/10.1080/03098265.2016.1185772

Cooke, B J (2016) 'Embedding digital (information) literacy in the first year curriculum: An intervention', *Journal of Perspectives in Applied Academic Practice,* 4 (2): 40-45. Available at: jpaap.napier.ac.uk/index.php/JPAAP/article/view/204

Cox, B E, Reason, R D, Tobolowsky, B F, Brower, R L, Patterson, S, Luczyk, S and Roberts, K (2017) 'Lip service or actionable insights? Linking student experiences to institutional assessment and data-driven decision making in higher education', *Journal of Higher Education,* 88 (6): 835-862. Available at: doi.org/10.1080/00221546.2016.1272320

Cummings, R and Maddux, C D (2010) 'The use of e-portfolios as a component of assessment and accreditation in higher education', in Buzzetto-More, N (ed) *The e-portfolio paradigm: Informing, educating, assessing and managing with e-portfolios.* Santa Rosa, CA: Informing Science Press, pp 207-223.

Davies, G R, Proops, H and Carolan, C M (2020) 'The development and use of a multiple-choice question (MCQ) assessment to foster deeper learning: An exploratory web-based qualitative investigation', *Journal of Teaching and Learning,* 14 (1): 1-12. Available at: doi.org/10.22329/jtl.v14i1.6300

Dawson, P, Henderson, M, Mahoney, P, Phillips, M, Ryan, T, Boud, D and Molloy, E (2019) 'What makes for effective feedback: Staff and student perspectives', *Assessment and Evaluation in Higher Education,* 44 (1): 25-36. Available at: doi.org/10.1080/02602938.2018.1467877

Deneen, C C, Brown, G and Carless, D (2018) 'Students' conceptions of e-portfolios as assessment and technology', *Innovations in Education and Teaching International,* 55 (4): 487-496. Available at: doi.org/10.1080/14703297.2017.1281752

Dibbs, R (2019) 'Forged in failure: Engagement patterns for successful students repeating calculus', *Educational Studies in Mathematics,* 101 (1): 35-50. Available at: doi.org/10.1007/s10649-019-9877-0

Dijks, M A, Brummer, L and Kostons, D (2018) 'The anonymous reviewer: The relationship between perceived expertise and the perceptions of peer feedback in higher education', *Assessment and Evaluation in Higher Education,* 43 (8): 1258-1271. Available at: doi.org/10.1080/02602938.2018.1447645

Double, K S, McGrane, J A and Hopfenbeck, T N (2019) 'The impact of peer assessment on academic performance: A meta-analysis of control group studies', *Educational Psychology Review,* 32: 481-509. Available at: doi.org/10.1007/s10648-019-09510-3

Douglas, T, Salter, S, Iglesias, M, Dowlman, M and Eri, R (2016) 'The feedback process: Perspectives of first- and second-year undergraduate students in the disciplines of education, health science and nursing', *Journal of University Teaching and Learning Practice,* 13 (1). Available at: doi.org/10.53761/1.13.1.3

Doyle, E and Buckley, P (2020) 'The impact of co-creation: An analysis of the effectiveness of student authored multiple choice questions on achievement of learning outcomes', *Interactive Learning Environments.* Available at: doi.org/10.1080/10494820.2020.1777166

Doyle, E, Buckley, P and Whelan, J (2019) 'Assessment co-creation: An exploratory analysis of opportunities and challenges based on student and instructor perspectives', *Teaching In Higher Education,* 24 (6): 739-754. Available at: doi.org/10.1080/13562517.2018.1498077

Egelandsdal, K and Krumsvik, R J (2019) 'Clicker interventions at university lectures and the feedback gap', *Nordic Journal of Digital Literacy,* 14 (1): 70-87. Available at: doi.org/10.18261/issn.1891-943x-2019-01-02-06

Elkington, S (2020) 'Essential frameworks for enhancing student success: Transforming assessment in higher education', *A guide to the Advance HE Framework.* York: Advance HE.

El-Maaddawy, T and Deneen, C (2017) 'Outcomes-based assessment and learning: Trialling change in a postgraduate civil engineering course', *Journal of University Teaching and Learning Practice,* 14 (1). Available at: ro.uow.edu.au/jutlp/vol14/iss1/10/

Impacts of Higher Education Assessment and Feedback Policy and Practice on Students: A Review of the Literature 2016-2021
Edd Pitt and Kathleen M Quinlan

Elmahdi, I, Al-Hattami, A and Fawzi, H (2018) 'Using technology for formative assessment to improve students' learning', *Turkish Online Journal of Educational Technology,* 17 (2): 182-188.

Elola, I and Oskoz, A (2016) 'Supporting second language writing using multimodal feedback', *Foreign Language Annals,* 49 (1): 58-74. Available at: doi.org/10.1111/flan.12183

Eltahir, M E, Alsalhi, N R, Al-Qatawneh, S, AlQudah, H A and Jaradat, M (2021) 'The impact of game-based learning (GBL) on students' motivation, engagement and academic performance on an Arabic language grammar course in higher education', *Education and Information Technologies,* 26 (3): 3251-3278. Available at: doi.org/10.1007/s10639-020-10396-w

Esterhazy, R and Damşa, C (2019) 'Unpacking the feedback process: An analysis of undergraduate students' interactional meaning-making of feedback comments', *Studies in Higher Education,* 44 (2): 260-274. Available at: doi.org/10.1080/03075079.2017.1359249

Eyler, J, Giles, D E, Stenson, C M and Gray, C J (2001) *At-a-glance: What we know about the effects of service-learning on students, faculty, institutions and communities, 1993-2001.* Available at: digitalcommons.unomaha.edu/slcehighered/139/

Flournoy, E L and Bauman, L C (2021) 'Collaborative assessment: Using self-assessment and reflection for student learning and program development', *Canadian Journal for the Scholarship of Teaching and Learning,* 12 (1). Available at: doi.org/10.5206/cjsotlrcacea.2021.1.14207

Foss, D J and Pirozzolo, J W (2017) 'Four semesters investigating frequency of testing, the testing effect and transfer of training', *Journal of Educational Psychology,* 109 (8): 1067-1083. Available at: doi.org/10.1037/edu0000197

Francis, R A, Millington, J D A and Cederlöf, G (2019) 'Undergraduate student perceptions of assessment and feedback practice: Fostering agency and dialogue', *Journal of Geography in Higher Education,* 43 (4): 468-485. Available at: doi.org/10.1080/03098265.2019.1660867

Frank, B, Simper, N and Kaupp, J (2018) 'Formative feedback and scaffolding for developing complex problem solving and modelling outcomes', *European Journal of Engineering Education,* 43 (4): 552-568. Available at: doi.org/10.1080/03043797.2017.1299692

Gamlath, S (2021) 'Business undergraduates' progress and satisfaction with learning experiences: The role of group assessment', *Assessment and Evaluation In Higher Education,* 46 (3): 360-375. Available at: doi.org/10.1080/02602938.2020.1776839

García-Loro, F, Martín, S, Ruipérez-Valiente, J A, San Cristóbal, E and Castro, M (2020) 'Reviewing and analysing peer review inter-rater reliability in a MOOC platform', *Computers and Education,*154. Available at: doi.org/10.1016/j.compedu.2020.103894

Gezer-Templeton, P, Mayhew, E J, Korte, D S and Schmidt, S J (2017) 'Use of exam wrappers to enhance students' metacognitive skills in a large introductory food science and human nutrition course', *Journal of Food Science Education,* 16 (1): 28-36. Available at: doi.org/10.1111/1541-4329.12103

Gilewski, A, Mallory, E, Sandoval, M, Litvak, M and Ye, L (2019) 'Does linking help? Effects and student perceptions of a learner-centered assessment implemented in introductory chemistry', *Chemistry Education Research and Practice,* 20 (2): 399-411. Available at: doi.org/10.1039/C8RP00248G

Gopal, R, Singh, V and Aggarwal, A (2021) 'Impact of online classes on the satisfaction and performance of students during the pandemic period of Covid 19', *Education Information Technology,* 26: 6923-6947. Available at: doi.org/10.1007/s10639-021-10523-1

Gould, J and Day, P (2013) 'Hearing you loud and clear: Student perspectives of audio feedback in higher education', *Assessment and Evaluation in Higher Education,* 38 (5): 554-566. Available at: doi.org/10.1080/02602938.2012.660131

Gozzard, D R and Zadnik, M G (2021) 'Contribution of self-directed, naked-eye observations to students' conceptual understanding and attitudes towards astronomy', *Physical Review Physics Education Research,* 17 (1). Available at: doi.org/10.1103/physrevphyseducres.17.010134

Grainger, P R, Heck, D and Carey, M D (2018) 'Are assessment exemplars perceived to support self-regulated learning in teacher education?', *Frontiers in Education,* 3: 1-20. Available at: doi.org/10.3389/feduc.2018.00060

Gravett, K (2020) 'Feedback literacies as sociomaterial practice', *Critical Studies in Education.* Available at: doi.org/10.1080/17508487.2020.1747099

Grigoryan, A (2017) 'Feedback 2.0 in online writing instruction: Combining audio-visual and text-based commentary to enhance student revision and writing competency', *Journal of Computing in Higher Education,* 29: 451-476. Available at: doi.org/10.1007/s12528-017-9152-2

Guardia, J J, Del Olmo, J, Roa, I and Berlanga, V (2019) 'Innovation in the teaching-learning process: The case of Kahoot!', *On The Horizon,* 27 (1): 35-45. Available at: doi.org/10.1108/OTH-11-2018-0035

Guo, F and Shi, J (2016) 'The relationship between classroom assessment and undergraduates' learning within Chinese higher education system', *Studies in Higher Education,* 41 (4): 642-663. Available at: doi.org/10.1080/03075079.2014.942274

Gurer, M D (2020) 'Sense of community, peer feedback and course engagement as predictors of learning in blog environments', *Turkish Online Journal of Distance Education,* 21 (4): 237-250. Available at: doi.org/10.17718/tojde.803415

Hallenberg, K M and Haddow, C (2016) 'Beyond criminal justice: connecting justice and sustainability', *The Law Teacher,* 50 (3): 352- 370. Available at: doi.org/10.1080/03069400.2016.1262985

Hammad, S, Dimitriadis, C and Graham, T (2021) 'Using concepts maps in a foundation mathematics

Impacts of Higher Education Assessment and Feedback Policy and Practice on Students: A Review of the Literature 2016-2021
Edd Pitt and Kathleen M Quinlan

course: What have we learnt?', *Eurasia Journal of Mathematics, Science and Technology Education,* 17 (2). Available at: doi.org/10.29333/ejmste/9700

Hammond, K M and Brown, S (2021) 'Transitioning to learning outcomes at the coalface: An academic's quantitative evaluation at the course level', *Studies in Educational Evaluation*, 68. Available at: doi.org/10.1016/j.stueduc.2020.100961

Han, Y and Hyland, F (2019) 'Academic emotions in written corrective feedback situations', *Journal of English for Academic Purposes,* 38: 1-13. Available at: doi.org/10.1016/j.jeap.2018.12.003

Han, Y and Xu, Y (2021) 'Student feedback literacy and engagement with feedback: A case study of Chinese undergraduate students', *Teaching in Higher Education,* 26 (2): 181-196. Available at: doi.org/10.1080/13562517.2019.1648410

Hancock, D, Hare, N, Denny, P and Denyer, G (2018) 'Improving large class performance and engagement through student-generated question banks', *Biochemistry and Molecular Biology Education,* 46 (4): 306-317. Available at: doi.org/10.1002/bmb.21119

Hansen, M and Mendzheritskaya, J (2017) 'How university lecturers' display of emotion affects students' emotions, failure attributions and behavioural tendencies in Germany, Russia and the United States', *Journal of Cross-Cultural Psychology,* 48 (5): 734-753. Available at: doi.org/10.1177/0022022117697845

Hanum, S R,, Che-Ani, A, Johar, S, Ismail, K and Abd-Razak, M (2016) 'E-portfolio: A descriptive survey for contents and challenges', *International Journal Of Emerging Technologies in Learning,* 11 (1): 4-10. Available at: doi.org/10.3991/ijet.v11i01.4900

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', *Assessment and Evaluation in Higher Education,* 40 (4): 528-541. doi.org/10.1080/02602938.2014.931927

Hattie, J and Timperley, H (2007) 'The power of feedback', *Review of Educational Research,* 77 (1): 81-1.

Hawe, E and Dixon, H (2017) 'Assessment for learning: A catalyst for student self-regulation', *Assessment and Evaluation in Higher Education,* 42 (8): 1181-1192. Available at: doi.org/10.1080/02602938.2016.1236360

Hawe, E, Lightfoot, U and Dixon, H (2019) 'First-year students working with exemplars: Promoting self-efficacy, self-monitoring and self-regulation', *Journal of Further and Higher Education,* 43 (1): 30-44. Available at: doi.org/10.1080/0309877X.2017.1349894

Headrick, J, Harris-Reeves, B and Daly-Olm, T (2021) 'Collaborative testing in sport and exercise degrees: A comparison of first- and third-year students' perceptions', *Journal of University Teaching and Learning Practice,* 18 (6). Available at: https://doi.org/10.53761/1.18.6.11

Healey, M and Jenkins, A (2009) Linking discipline-based research and teaching through mainstreaming undergraduate research and inquiry

Heinicke, M R, Zuckerman, C K and Cravalho, D A (2017) 'An evaluation of readiness assessment tests in a college classroom: Exam performance, attendance and participation', *Behavior Analysis: Research and Practice,* 17 (2): 129-141. Available at: https://doi.org/10.1037/bar0000073

Heitkamp, S J, Rüttermann, S and Gerhardt-Szép, S (2018) 'Work shadowing in dental teaching practices: Evaluation results of a collaborative study between university and general dental practices', *BMC Medical Education,* 99 (18). Available at: doi.org/10.1186/s12909-018-1220-4

Henderson, M, Ajjawi, R, Boud, D and Molloy, E (2019) 'Identifying feedback that has impact', in Henderson, M, Ajjawi, R, Boud, D and Molloy, E (eds) *The Impact of Feedback in Higher Education: Improving Assessment Outcomes for Learners.* Cham: Palgrave Macmillan, pp 15-34.

Henderson, M, Molloy, E, Ajjawi, R and Boud, D (2019) 'Designing feedback for impact', in Henderson, M, Ajjawi, R, Boud, D and Molloy, E (eds) *The Impact Of Feedback In Higher Education: Improving Assessment Outcomes For Learners.* Cham: Palgrave Macmillan, pp 267-285. Available at: doi.org/10.1007/978-3-030-25112-3_15

Hendry, G D, White, P and Herbert, C (2016) 'Providing exemplar-based 'feedforward' before an assessment: The role of teacher explanation', *Active Learning in Higher Education,* 17 (2): 99-109. Available at: doi.org/10.1177/1469787416637479

Henri, M, Johnson, M D and Nepal, B (2017) 'A review of competency-based learning: Tools, assessments and recommendations', *Journal of Engineering Education,* 106 (4): 607-638. Available at: doi.org/10.1002/jee.20180

Herbert, K, Demskoi, D and Cullis, K (2019) 'Creating mathematics formative assessments using LaTeX, PDF forms and computer algebra', *Australasian Journal of Educational Technology,* 35 (5): 153-167. Available at: doi.org/10.14742/ajet.4539

Hill, J and West, H (2020) 'Improving the student learning experience through dialogic feed-forward assessment', *Assessment and Evaluation in Higher Education,* 45 (1): 82-97. Available at: doi.org/10.1080/02602938.2019.1608908

Hill, N, Denholm-Price, J, Atkins, N, Dourado, L T, Nimoh, O and Page, N (2016) 'Does group assessment impact BME attainment?', *New Directions in the Teaching of Physical Sciences,* 11 (1). Available at: doi.org/10.29311/ndtps.v0i11.579

Hirst, N (2016) 'Using Pecha Kucha as formative assessment in two undergraduate modules (re) conceptualising the right lines', *Practitioner Research in Higher Education,* 10 (1): 140-155.

Hsia, L-H, Huang, I and Hwang, G-J (2016) 'Effects of different online peer-feedback approaches on students' performance skills, motivation and self-efficacy in a dance course', *Computers and Education,* 96: 55-71. Available at: doi.org/10.1016/j.compedu.2016.02.004

Huang, S-C (2018) 'A GEARed peer feedback model and implications for learning facilitation',

Assessment and Evaluation in Higher Education, 43 (7): 1194-1210. Available at: doi.org/10.1080/02602938.2018.1439881

Huisman, B, Saab, N, van den Broek, P and van Driel, J (2019) 'The impact of formative peer feedback on higher education students' academic writing: A meta-analysis', *Assessment and Evaluation in Higher Education,* 44 (6): 863-880. Available at: doi.org/10.1080/02602938.2018.1545896

Huisman, B, Saab, N, van Driel, J and van den Broek, P (2018) 'Peer feedback on academic writing: Undergraduate students' peer feedback role, peer feedback perceptions and essay performance', *Assessment and Evaluation in Higher Education,* 43 (6): 955-968. Available at: doi.org/10.1080/02602938.2018.1424318

Hung, S-T A and Huang, H-T D (2016) 'Blogs as a learning and assessment instrument for English-speaking performance', *Interactive Learning Environments,* 24 (8): 1881-1894. Available at: doi.org/10.1080/10494820.2015.1057746

Hunukumbure, A D, Smith, S F and Das, S (2017) 'Holistic feedback approach with video and peer discussion under teacher supervision', *BMC Medical Education,* 17. Available at: doi.org/10.1186/s12909-017-1017-x

Ilgaz, H and Adanir, G A (2020) 'Providing online exams for online learners: Does it really matter for them?', *Education And Information Technologies,* 25 (2): 1255-1269. Available at: doi.org/10.1007/s10639-019-10020-6

Ion, G, Cano, E and Cabrera, N (2016) 'Competency assessment tool (CAT). The evaluation of an innovative competency-based assessment experience in higher education', *Technology, Pedagogy and Education,* 25 (5): 631-648. Available at: doi.org/10.1080/1475939X.2015.1134635

Jaap, A, Dewar, A, Duncan, C, Fairhurst, K, Hope, D and Kluth, D (2021) 'Effect of remote online exam delivery on student experience and performance in applied knowledge tests', *BMC Medical Education,* 21 (1). Available at: doi.org/10.1186/s12909-021-02521-1

James, L T and Casidy, R (2018) 'Authentic assessment in business education: Its effects on student satisfaction and promoting behaviour', *Studies in Higher Education,* 43 (3): 401-415. Available at: doi.org/10.1080/03075079.2016.1165659

Jeon, E Y (2018) 'The effect of learner centred EFL writing instruction on Korean university students' writing anxiety and perception', *TESOL International Journal,* 13 (3): 100-112.

Jeong, J S, González-Gómez, D and Prieto, F Y (2020) 'Sustainable and flipped STEM education: Formative assessment online interface for observing pre-service teachers' performance and motivation', *Education Sciences,* 10 (10). Available at: doi.org/10.3390/educsci10100283

Joh, J (2019) 'Impact of Peer Feedback on Learning: A case of EFL teacher trainees', *Journal of Asia TEFL,* 16 (4): 1103-1118. Available at: doi.org/10.18823/asiatefl.2019.16.4.3.1103

Jolly, B and Boud, D (2013) 'Written feedback: what is it good for and how can we do it well?', in Boud, D and Molloy, E (eds) *Feedback in higher and professional education: Understanding it and doing it well.* Abingdon: Routledge, pp 104-124.

Jorre de St Jorre, T, Boud, D and Johnson, E D (2021) 'Assessment for distinctiveness: Recognising diversity of accomplishments', *Studies in Higher Education,* 46 (7): 1371-1382. Available at: doi.org/10.1080/03075079.2019.1689385

Jurkowski, S (2018) 'Do question prompts support students in working with peer feedback?', *International Journal of Educational Research,* 92: 1-9. Available at: doi.org/10.1016/j.ijer.2018.07.003

Kahl, J (2017) 'Automatic, multiple assessment options in undergraduate meteorology education', *Assessment and Evaluation in Higher Education,* 42 (8): 1319-1325. Available at: doi.org/10.1080/02602938.2016.1249337

Kaider, F, Hains-Wesson, R and Young, K (2017) 'Practical typology of authentic work-integrated learning activities and assessments', *Asia-Pacific Journal of Cooperative Education,* 18 (2): 153-165.

Kang, D, Goico, S, Ghanbari, S, Bennallack, K C, Pontes, T, O'Brien, D H and Hargis, J (2019) 'Providing an oral examination as an authentic assessment in a large section, undergraduate diversity class', *International Journal for the Scholarship of Teaching and Learning,* 13 (2). Available at: doi.org/10.20429/ijsotl.2019.130210

Kaur, A, Noman, M and Nordin, H (2017) 'Inclusive assessment for linguistically diverse learners in higher education', *Assessment and Evaluation in Higher Education,* 42 (5): 756-771. Available at: doi.org/10.1080/02602938.2016.1187250

Kay, A E, Hardy, J and Galloway, R K (2020) 'Student use of PeerWise: A multi-institutional, multidisciplinary evaluation', *British Journal of Educational Technology,* 51 (1): 23-35. Available at: doi.org/10.1111/bjet.12754

Keshavarz, M H and Köseoğlu, M P (2021) 'A letter to my teacher: Interactive written teacher feedback', *TESOL Journal,* 12 (3). Available at: doi.org/10.1002/tesj.598

Kılıçkaya, F (2017) 'The effects of pre-lecture online quizzes on language students' perceived preparation and academic performance', *PASAA Journal,* 53: 59-84.

Kim, E J and Lee, K R (2019) 'Effects of an examiner's positive and negative feedback on self-assessment of skill performance, emotional response and self-efficacy in Korea: A quasi-experimental study', *BMC Medical Education,* 19. Available at: doi.org/10.1186/s12909-019-1595-x

Kinzie, J (2019) 'Taking stock of initiatives to improve learning quality in American higher education through assessment', *Higher Education Policy,* 32 (4): 577-595. Available at: doi.org/10.1057/s41307-019-00148-y

Kluger, A N and DeNisi, A (1996) 'The effects of feedback interventions on performance: A historical review, a meta-analysis and a preliminary feedback intervention theory', *Psychological Bulletin,* 119 (2): 254. Available at: doi.org/10.1037/0033-2909.119.2.254

Impacts of Higher Education Assessment and Feedback Policy and Practice on Students: A Review of the Literature 2016-2021
Edd Pitt and Kathleen M Quinlan

Knauf, H (2016) 'Reading, listening and feeling: Audio feedback as a component of an inclusive learning culture at universities', *Assessment and Evaluation in Higher Education,* 41 (3): 442-449. Available at: doi.org/10.1080/02602938.2015.1021664

Knight, P (2002) 'A systemic approach to professional development: Learning as practice', *Teaching and Teacher Education,* 18 (3): 229-241. Available at: doi.org/10.1016/S0742-051X(01)00066-X

Kuh, G D and O'Donnell, K (2013) *Ensuring quality and taking high-impact practices to scale.* Washington, DC: Association Of American Colleges And Universities.

Kuh, G, O'Donnell, K and Schneider, C G (2017) 'HIPs at ten', *Change, The Magazine Of Higher Learning,* 49 (5): 8-16. Available at: doi.org/10.1080/00091383.2017.1366805

Landers, R N (2015) 'Developing a theory of gamified learning: Linking serious games and gamification of learning', *Simulation and Gaming,* 45: 752-768. Available at: doi.org/10.1177/1046878114563660

Lavrysh, Y (2016) 'Peer and self-assessment at ESP classes: Case study', Advanced Education, 6: 60-68. Available at: doi.org/10.20535/2410-8286.85351

Leung, A, Pour, B H, Reynolds, D and Jerzak, S (2017) 'New assessment process in an introductory undergraduate physics laboratory: An exploration on collaborative learning', *Assessment and Evaluation in Higher Education,* 42 (2): 169-181. Available at: doi.org/10.1080/02602938.2015.1089977

Lewis, S, Heath, G, Lim, L and Roberts, R (2021) ''I'm not a number, I'm someone to them': Supporting commencing university students' through technology-mediated personalised communication', *Student Success,* 12 (1): 24-34. Available at: doi.org/10.5204/ssj.1623

Li, F and Curdt-Christiansen, X (2020) 'Teacher feedback in UK higher education: Affective and cognitive perceptions of Chinese postgraduate students', *International Journal of Educational Research,* 104: 101674. Available at: doi.org/10.1016/j.ijer.2020.101674

Li, H, Xiong, Y, Hunter, C V, Guo, X and Tywoniw, R (2020) 'Does peer assessment promote student learning? A meta-analysis', *Assessment and Evaluation in Higher Education,* 45 (2): 193-211. Available at: doi.org/10.1080/02602938.2019.1620679

Li, H, Zhao, C, Long, T, Huang, Y and Shu, F (2021) 'Exploring the reliability and its influencing factors of peer assessment in massive open online courses', *British Journal of Educational Technology,* 52 (6): 2263-2277. Available at: doi.org/10.1111/bjet.13143

Li, N, Marsh, V and Rienties, B (2016) 'Modelling and managing learner satisfaction: Use of learner feedback to enhance blended and online learning experience', *Decision Sciences – Journal of Innovative Education,* 14 (2): 216-242. Available at: doi.org/10.1111/dsji.12096

Liu, L, Du, X H, Zhang, Z and Zhou, J Y (2019) 'Effect of problem-based learning in pharmacology education: A meta-analysis', *Studies in Educational Evaluation,* 60: 43-58. Available at: doi.org/10.1016/j.stueduc.2018.11.004

Liu, X, Li, L and Zhang, Z (2018) 'Small group discussion as a key component in online assessment training for enhanced student learning in web-based peer assessment', *Assessment and Evaluation in Higher Education,* 43 (2): 207-222. Available at: doi.org/10.1080/02602938.2017.1324018

Lizzio, A and Wilson, K (2013) 'First-year students' appraisal of assessment tasks: Implications for efficacy, engagement and performance', *Assessment and Evaluation in Higher Education,* 38 (4): 389-406. Available at: doi.org/10.1080/02602938.2011.637156

López, G A, Sáenz, J, Leonardo, A and Gurtubay, I G (2016) 'Use of the Moodle platform to promote an ongoing learning when lecturing general physics in the physics, mathematics and electronic engineering programmes at the University of the Basque Country UPV/EHU', *Journal of Science Education and Technology,* 25 (4): 575-589. Available at: doi.org/10.1007/s10956-016-9614-8

Lu, V N, Scholz, B and Nguyen, L T V (2018) 'Work integrated learning in international marketing: Student insights', *Australasian Marketing Journal,* 26 (2): 132-139. Available at: doi.org/10.1016/j.ausmj.2018.05.002

Lunt, T and Curran, J (2010) ''Are you listening please?' The advantages of electronic audio feedback compared to written feedback', *Assessment and Evaluation in Higher Education,* 35 (7): 759-769. doi.org/10.1080/02602930902977772

Lynam, S and Cachia, M (2018) 'Students' perceptions of the role of assessments at higher education', *Assessment and Evaluation in Higher Education,* 43 (2): 223-234. Available at: doi.org/10.1080/02602938.2017.1329928

MacKay, J R D, Hughes, K, Marzetti, H, Lent, N and Rhind, S M (2019) 'Using National Student Survey (NSS) qualitative data and social identity theory to explore students' experiences of assessment and feedback', *Higher Education Pedagogies,* 4 (1): 315-330. Available at: doi.org/10.1080/23752696.2019.1601500

Mahoney, P, Macfarlane, S and Ajjawi, R (2019) 'A qualitative synthesis of video feedback in higher education', *Teaching in Higher Education,* 24 (2): 157-179. Available at: doi.org/10.1080/13562517.2018.1471457

Malecka, B, Boud, D and Carless, D (2020) 'Eliciting, processing and responding to feedback: Mechanisms for embedding feedback literacy within the curriculum', *Teaching in Higher Education.* Available at: doi.org/10.1080/13562517.2020.1754784

Mathisen, P (2012) 'Video feedback in higher education – a contribution to improving the quality of written feedback', *Nordic Journal of Digital Literacy,* 7 (2): 97-116. Available at: doi.org/10.18261/ISSN1891-943X-2012-02-02

Mayhew, E (2017) 'Playback feedback: the impact of screen-captured video feedback on student satisfaction, learning and attainment', *European Political Science,* 16 (2): 179-192. Available at: doi.org/10.1057/eps.2015.102

McArthur, J (2020) 'Student involvement in assessment: Involving the whole student in pursuit of social justice and the social good', *RELIEVE – Revista Electronica De Investigacion Y Evaluacion Educativa,* 26 (1). Available at: doi.org/10.7203/relieve.26.1.17089

McCarthy, J (2017) 'Enhancing feedback in higher education: Students' attitudes towards online and in-class formative assessment feedback models', *Active Learning In Higher Education,* 18 (2): 127-141. Available at: doi.org/10.1177/1469787417707615

McGahan, W, Ernst, H and Dyson, L E (2016) 'Individual learning strategies and choice in student-generated multimedia', *International Journal of Mobile and Blended Learning,* 8 (3): 1-18. Available at: doi.org/10.4018/IJMBL.2016070101

Meek, S E M, Blakemore, L and Marks, L (2017) 'Is peer review an appropriate form of assessment in a MOOC? Student participation and performance in formative peer review', *Assessment and Evaluation in Higher Education,* 42 (6): 1000-1013. Available at: doi.org/10.1080/02602938.2016.1221052

Melser, M C, Lettner, S, Bäwert, A, Puttinger, C and Holzinger, A (2020) 'Pursue today and assess tomorrow – how students' subjective perceptions influence their preference for self- and peer assessments', *BMC Medical Education,* 20 (1). Available at: doi.org/10.1186/s12909-020-02383-z

Milne, L, McCann, J, Bolton, K, Savage, J and Spence, A (2020) 'Student satisfaction with feedback in a third-year nutrition unit: A strategic approach', *Journal of University Teaching and Learning Practice,* 17 (5). Available at: doi.org/10.53761/1.17.5.5

Min, H-T (2016) 'Effect of teacher modelling and feedback on EFL students' peer review skills in peer review training', *Journal of Second Language Writing,* 31: 43-57. Available at: doi.org/10.1016/j.jslw.2016.01.004

Mir, K, Iqbal, M Z and Shams, J A (2019) 'An investigation of AIOU students' satisfaction about formative m-assessment using SMS technology', *Pakistan Journal of Distance and Online Learning,* 5 (2): 157-174.

Misiejuk, K, Wasson, B and Egelandsdal, K (2020) 'Using learning analytics to understand student perceptions of peer feedback', *Computers In Human Behavior,* 117: 106658. Available at: doi.org/10.1016/j.chb.2020.106658

Mohamadi, Z (2018) 'Comparative effect of online summative and formative assessment on EFL student writing ability', *Studies in Educational Evaluation,* 59: 29-40. Available at: doi.org/10.1016/j.stueduc.2018.02.003

Molloy, E, Boud, D and Henderson, M (2020) 'Developing a learning-centred framework for feedback literacy', *Assessment and Evaluation in Higher Education,* 45 (4): 527-540. Available at: doi.org/10.1080/02602938.2019.1667955

Molloy, E, Noble, C and Ajjawi, R (2019) 'Attending to emotion in feedback', in Henderson, M, Ajjawi, R, Boud, D and Molloy, E (eds) *The impact of feedback in higher education.* Cham: Palgrave

Macmillan. Available at: doi.org/10.1007/978-3-030-25112-3_6

Morley, C and Ablett, P (2017) 'Designing assessment to promote engagement among first year social work students', *E-Journal of Business Education and Scholarship of Teaching,* 11 (2): 1-14.

Morris, N J, Christie, H and Barber, J (2019) ''It's one of the first times I've felt fully engaged': Developing student engagement using blogging as a form of assessment', *Journal of Geography in Higher Education,* 43 (3): 343-361. Available at: doi.org/10.1080/03098265.2019.1612862

Mshayisa, V V (2020) 'Students' perceptions of Plickers and crossword puzzles in undergraduate studies', *Journal of Food Science Education,* 19 (2): 49-58. Available at: doi.org/10.1111/1541-4329.12179

Mujtaba, S M, Parkash, R and Nawaz, M W (2020) 'Do indirect coded corrective feedback and teachers short affective comments improve the writing performance and learners uptake?', *Reading and Writing Quarterly,* 36 (1): 34-47. Available at: doi.org/10.1080/10573569.2019.1616638

Mulliner, E and Tucker, M (2017) 'Feedback on feedback practice: Perceptions of students and academics', *Assessment and Evaluation in Higher Education,* 42 (2): 266-288. Available at: doi.org/10.1080/02602938.2015.1103365

Murillo-Zamorano, L R and Montanero, M (2018) 'Oral presentations in higher education: A comparison of the impact of peer and teacher feedback', *Assessment and Evaluation in Higher Education,* 43 (1): 138-150. Available at: doi.org/10.1080/02602938.2017.1303032

Nagda, B A, Gregerman, S R, Jonides, J, von Hippel, W and Lerner, J S (1998) 'Undergraduate student-faculty research partnerships affect student retention', *Review of Higher Education,* 22 (1): 55-72. Available at: doi.org/10.1353/rhe.1998.0016

Naghdipour, B (2017) 'Incorporating formative assessment in Iranian EFL writing: A case study', *Curriculum Journal,* 28 (2): 283-299. Available at: doi.org/10.1080/09585176.2016.1206479

Nardi, A and Ranieri, M (2019) 'Comparing paper-based and electronic multiple-choice examinations with personal devices: Impact on students' performance, sel-efficacy and satisfaction', *British Journal of Educational Technology,* 50 (3): 1495-1506. Available at: doi.org/10.1111/bjet.12644

Neal, K B, Amendola, N, Jeske, D M, Martorana, R, Smith, E G, Joyce, S J and Baudinette, S (2018) 'A team-based public history assessment for undergraduates: Rationale, design and implementation in a medieval history course', *International Journal of Teaching and Learning in Higher Education,* 30 (2): 334-349.

Neupane Bastola, M and Hu, G (2021) ''Commenting on your work is a waste of time only!': An appraisal-based study of evaluative language in supervisory feedback', *Studies in Educational Evaluation,* 68. Available at: doi.org/10.1016/J.Stueduc.2020.100962

Ngo, J and Ngadiman, A (2019) 'The impacts of Edmodo on students' performance in ESP classrooms', *Knowledge E Social Sciences,* 3 (10): 369-378. Available at: doi.org/10.18502/kss.

v3i10.3918

Nguyen, T T H and Walker, M (2016) 'Sustainable assessment for lifelong learning', *Assessment and Evaluation in Higher Education,* 41 (1): 97-111. Available at: doi.org/10.1080/02602938.2014.985632

Nicol, D (2021) 'The power of internal feedback: Exploiting natural comparison processes', *Assessment and Evaluation in Higher Education,* 46 (5): 756-778. Available at: doi.org/10.1080/02602938.2020.1823314

Nicol, D, Thomson, A and Breslin, C (2014) 'Rethinking feedback practices in higher education: A peer review perspective', *Assessment and Evaluation in Higher Education,* 39 (1): 102-122. Available at: doi.org/10.1080/02602938.2013.795518

Nikolaeva, S and Korol, T (2021) 'Prospective philologists' translation assessment triangulation: Screen video recording and think aloud protocol combination', *Advanced Education,* 18: 30-41. Available at: doi.org/10.20535/2410-8286.228550

Nicola-Richmond, K, Richards, K and Britt, K (2016) 'The impact of an authentic, simulated learning activity on student preparedness for work-integrated learning', *Asia-Pacific Journal of Cooperative Education,* 16 (4): 343-354.

Nkealah, N E (2019) 'Applying formative assessment strategies in the teaching of poetry: An experiment with third year English studies students at the University of Limpopo', *South African Journal of Higher Education,* 33 (1): 242-261. Available at: doi.org/10.20853/33-1-1373

Noble, C, Billet, S, Sly, C, Collier, L, Armit, L, Hilder, J and Molloy, E (2020) ''It's yours to take': Generating learner feedback literacy in the workplace', *Advances in Health Sciences Education,* 25: 55-74. Available at: doi.org/10.1007/s10459-019-09905-5

Noroozi, O and Mulder, M (2017) 'Design and evaluation of a digital module with guided peer feedback for student learning biotechnology and molecular life sciences, attitudinal change and satisfaction', *Biochemistry and Molecular Biology Education,* 45 (1): 31-39. Available at: doi.org/10.1002/bmb.20981

O'Brien, M and Freund, K (2018) 'Lessons learned from introducing social media use in undergraduate economics research', *International Journal of Education and Development using Information and Communication Technology,* 14 (1): 4-16.

O'Donovan, B (2017) 'How student beliefs about knowledge and knowing influence their satisfaction with assessment and feedback', *Higher Education,* 74 (4): 617-633. Available at: doi.org/10.1007/s10734-016-0068-y

Office for Students (2021) *'Standards of evidence and evaluation self-assessment tool'.* Bristol and London: Office for Students. Available at: www.officeforstudents.org.uk/advice-and-guidance/promoting-equal-opportunities/evaluation/standards-of-evidence-and-evaluation-self-assessment-tool/ [accessed 24 March 2022].

O'Malley, E, Scanlon, A M, Alpine, L and McMahon, S (2021) 'Enabling the feedback process in work-based learning: An evaluation of the 5-minute feedback form', *Assessment and Evaluation in Higher Education,* 46 (7): 1020-1034. Available at: doi.org/10.1080/02602938.2020.1842852

Omer, T (2016) 'Nursing students' perceptions of satisfaction and self-confidence with clinical simulation experience', *Journal of Education and Practice,* 7 (5): 131-138.

Padayachee, P, Wagner-Welsh, S and Johannes, H (2018) 'Online assessment in Moodle: A framework for supporting our students', *South African Journal of Higher Education,* 32 (5): 211-235. Available at: doi.org/10.20853/32-5-2599

Palmer, L, Levett-Jones, T and Smith, R (2018) 'First year students' perceptions of academic literacies preparedness and embedded diagnostic assessment', *Student Success,* 9 (2): 49-61. Available at: doi.org/10.5204/ssj.v9i2.417

Paloncy, K A (2020) 'Postdebriefing supervised practice improves clinical performance during simulation-based cardiopulmonary resuscitation encounter', *Athletic Training Education Journal,* 15 (2): 85-92. Available at: doi.org/10.4085/1947-380X-19-064

Panadero, E and Alqassab, M (2019) 'An empirical review of anonymity effects in peer assessment, peer feedback, peer review, peer evaluation and peer grading', *Assessment and Evaluation in Higher Education,* 44 (8): 1253-1278. Available at: doi.org/10.1080/02602938.2019.1600186

Panadero, E and Jonsson, A (2020) 'A critical review of the arguments against the use of rubrics', *Educational Research Review,* 30. Available at: doi.org/10.1016/j.edurev.2020.100329

Park, J (2018) 'Effectiveness of teacher and peer feedback: Through the lens of Korean tertiary writing classroom', *The Journal of Asia TEFL,* 15 (2): 429-444. Available at: doi.org/10.18823/asiatefl.2018.15.2.11.429

Park, K-Y, Park, H-K and Hwang, H-S (2019) 'Group randomized trial of teaching tobacco-cessation counselling to senior medical students: A peer role-play module versus a standardized patient module', *BMC Medical Education,* 19. Available at: doi.org/10.1186/s12909-019-1668-x

Paschalis, G (2017) 'A compound LAMS-Moodle environment to support collaborative project-based learning: A case study with the group investigation method', *Turkish Online Journal of Distance Education,* 18 (2): 134-150. Available at: doi.org/10.17718/tojde.306565

Patchan, M M, Schunn, C D and Correnti, R J (2016) 'The nature of feedback: How peer feedback features affect students' implementation rate and quality of revisions', *Journal Of Educational Psychology,* 108 (8): 1098-1120. Available at: doi.org/10.1037/edu0000103

Paul, M W, Torgerson, C, Tracz, S, Coy, K and Wahleithner, J (2020) 'Engaging the control-value theory: A new era of student response systems and formative assessment to improve student achievement', *Research in Learning Technology,* 28. Available at: doi.org/10.25304/rlt.v28.2454

Pearce, C (2018) 'It pays to prepare: The value of low-stakes tutorial preparation exercises to student performance', *Australian Journal of Adult Learning,* 58 (2): 246-265.

Pelkola, T, Rasila, A and Sangwin, C (2018) 'Investigating Bloom's Learning for Mastery in mathematics with online assessment', *Informatics in Education,* 17 (2): 363-380. Available at: doi.org/10.15388/infedu.2018.19

Petrović, J, Pale, P and Jeren, B (2017) 'Online formative assessments in a digital signal processing course: Effects of feedback type and content difficulty on students learning achievements', *Education and Information Technologies,* 22 (6): 3047-3061. Available at: doi.org/10.1007/s10639-016-9571-0

Phillips, T N and Wood, L I (2017) 'Teaching diversity through case competition', *Journal of Learning in Higher Education,* 13 (2): 25-31.

Pitt, E and Carless, D (2021) 'Signature feedback practices in the creative arts: Integrating feedback within the curriculum', *Assessment and Evaluation in Higher Education.* doi.org/10.1080/02602938.2021.1980769

Pitt, E and Norton, L (2017) ''Now that's the feedback I want!' Students' reactions to feedback on graded work and what they do with it', *Assessment and Evaluation in Higher Education,* 42 (4): 499-516. Available at: doi.org/10.1080/02602938.2016.1142500

Pitt, E and Winstone, N (2018) 'The impact of anonymous marking on students' perceptions of fairness, feedback and relationships with lecturers', *Assessment and Evaluation in Higher Education,* 43 (7): 1183-1193. Available at: doi.org/10.1080/02602938.2018.1437594

Popham, W J (2011) 'Assessment literacy overlooked: A teacher educator's confession', *The Teacher Educator,* 46 (4), 265-273. Available at: doi.org/10.1080/08878730.2011.605048

Postlethwait, A E (2016) 'Group projects in social work education: The influence of group characteristics and moderators on undergraduate student outcomes', *Journal of Teaching in Social Work,* 36 (3): 256-274. Available at: doi.org/10.1080/08841233.2016.1182610

Pruett, M K, Schepard, A, Cornett, L, Gerety, C and Kourlis, R L (2018) 'Law students on interdisciplinary, problem-solving teams: An empirical evaluation of educational outcomes at the University of Denver's Resource Center for Separating and Divorcing Families', *Family Court Review,* 56 (1): 100-118. Available at: doi.org/10.1111/fcre.12325

Quesada, V, Gómez Ruiz, M Á, Gallego Noche, M B and Cubero-Ibáñez, J (2019) 'Should I use co-assessment in higher education? Pros and cons from teachers and students' perspectives', *Assessment and Evaluation in Higher Education,* 44 (7): 987-1002. Available at: doi.org/10.1080/02602938.2018.1531970

Quinlan, K M (2011) *Developing the whole student: Leading higher education initiatives that integrate mind and heart.* York: Leadership Foundation for Higher Education. Available at: www.advance-he.ac.uk/knowledge-hub/developing-whole-student-leading-higher-education-initiatives-integrate-mind-and

Quinlan, K M (2016) 'Developing student character through disciplinary curricula: An analysis of UK QAA subject benchmark statements', *Studies in Higher Education,* 41: 1041-1054. Available at: doi.org/10.1080/03075079.2014.966069

Quinlan, K M and Berndtson, E (2013) 'The emerging European higher education area: Implications for instructional development', in Simon, E and Pleschová, G (eds) *Teacher Development in Higher Education: Existing Programs, Program Impact and Future Trends.* New York: Taylor and Francis, pp 129-150. Available at: doi.org/10.4324/9780203096826

Quinlan, K M and Pitt, E (2021) 'Towards signature assessment and feedback practices: A taxonomy of discipline-specific elements of assessment for learning', *Assessment in Education: Principles, Policy and Practice,* 28 (2): 191-207. Available at: doi.org/10.1080/0969594X.2021.1930447

Ramon-Casas, M, Nuño, N, Pons, F and Cunillera, T (2019) 'The different impact of a structured peer-assessment task in relation to university undergraduates' initial writing skills', *Assessment and Evaluation in Higher Education,* 44 (5): 653-663. Available at: doi.org/10.1080/02602938.2018.1525337

Ramsgaard, M B and Østergaard, S J (2018) 'An entrepreneurial learning approach to assessment of internships', *Education and Training,* 60 (7): 909-922. Available at: doi.org/10.1108/ET-11-2016-0164

Rasi, P and Vuojärvi, H, (2017) 'Toward personal and emotional connectivity in mobile higher education through asynchronous formative audio feedback', *British Journal of Educational Technology,* 49 (2): 292-304. Available at: doi.org/10.1111/bjet.12587

Raupach, T, de Temple, I, Middeke, A, Anders, S, Morton, C and Schuelper, N (2021) 'Effectiveness of a serious game addressing guideline adherence: Cohort study with 1.5-year follow-up', *BMC Medical Education,* 21 (1). Available at: doi.org/10.1186/s12909-021-02591-1

Reboredo, J C (2017) 'Do continuous assessment results affect final exam outcomes? Evidence from a microeconomics course', *Multidisciplinary Journal for Education, Social and Technological Sciences,* 4 (1): 88-101. Available at: doi.org/10.4995/muse.2017.6548

Reju, C O and Jita, L C (2020) 'A comparative investigation of assessment practices in distance and online learning undergraduate mathematics in Nigeria', *Journal of Educational Research and Practice,* 10 (1): 90-103. Available at: doi.org/10.5590/JERAP.2020.10.1.06

Rivero, C, del Campo, C, Urquia-Grande, E, Camacho-Miñano, M and Pascual-Ezama, D (2017) 'Analysing students' academic performance in higher education in Spain', Educade –Revista De *Educacion En Contabilidad Finanzas Y Administracion De Empresas,* 8: 3-19. Available at: doi.org/10.12795/EDUCADE.2017.i08.02

Robinson, K E, Allen, P J, Quail, M and Beilby, J (2020) 'Virtual patient clinical placements improve student communication competence', *Interactive Learning Environments,* 28 (6): 795-805. Available at: doi.org/10.1080/10494820.2018.1552869

Roediger, H L, III and Karpicke, J D (2006) 'The power of testing memory: Basic research and implications for educational practice', *Perspectives on Psychological Science,* 1: 181-210. Available at: doi.org/10.1111/j.1745-6916.2006.00012.x

Rotsaert, T, Panadero, E and Schellens, T (2018) 'Anonymity as an instructional scaffold in peer assessment: Its effects on peer feedback quality and evolution in students' perceptions about peer assessment skills', *European Journal of Psychology of Education,* 33: 75-99. Available at: doi.org/10.1007/s10212-017-0339-8

Rovagnati, V and Pitt, E (2021) 'Exploring intercultural dialogic interactions between individuals with diverse feedback literacies', *Assessment and Evaluation in Higher Education.* Available at: doi.org/10.1080/02602938.2021.2006601

Rovagnati, V, Pitt, E and Winstone, N (2021) 'Feedback cultures, histories and literacies: International postgraduate students' experiences', *Assessment and Evaluation in Higher Education.* Available at: doi.org/10.1080/02602938.2021.1916431

Rowe, A D (2017) 'Feelings about feedback: The role of emotions in assessment for learning', in Carless, D, Bridges, S, Chan, C K and Glofcheski, R (eds) *Scaling Up Assessment for Learning in Higher Education.* Singapore: Springer, pp 159-172. Available at: doi.org/10.1007/978-981-10-3045-1_11

Ryan, T and Henderson, M (2018) 'Feeling feedback: Students emotional responses to educator feedback', *Assessment and Evaluation in Higher Education,* 43 (6): 880–892. Available at: doi.org/10.1080/02602938.2017.1416456

Ryan, T, Henderson, M and Phillips, M (2019) 'Feedback modes matter: Comparing student perceptions of digital and non-digital feedback modes in higher education', *British Journal of Educational Technology,* 50 (3): 1507-1523. Available at: doi.org/10.1111/bjet.12749

Sadler, D R (2010) 'Beyond feedback: Developing student capability in complex appraisal', *Assessment and Evaluation In Higher Education,* 35: 535-550. Available at: doi.org/10.1080/02602930903541015

Saeed, M A, Ghazali, K, Sahuri, S S and Abdulrab, M (2018) 'Engaging EFL learners in online peer feedback on writing: What does it tell us?', *Journal of Information Technology Education: Research,* 17: 39-61. Available at: doi.org/10.28945/3980

Sailer, M and Homner, L (2020) 'The gamification of learning: A meta-analysis', *Educational Psychology Review,* 32: 77-112. Available at: doi.org/10.1007/s10648-019-09498-w

Sailer, M and Sailer, M (2021) 'Gamification of in-class activities in flipped classroom lectures', *British Journal of Educational Technology,* 52 (1): 75-90. Available at: doi.org/10.1111/bjet.12948

Sandquist, E J, Cervato, C and Ogilvie, C (2019) 'Positive affective and behavioural gains of first-year students in course-based research across disciplines', *Scholarship and Practice of Undergraduate Research,* 2 (4): 45-57. Available at: doi.org/10.18833/spur/2/4/9

Impacts of Higher Education Assessment and Feedback Policy and Practice on Students: A Review of the Literature 2016-2021
Edd Pitt and Kathleen M Quinlan

Sangster, S L, Loy, K L, Mills, S D and Lawson, K L (2016) 'Engaging first-year university students in research: Promise, potentials and pitfalls', *Canadian Journal for the Scholarship of Teaching and Learning,* 7 (1). Available at: doi.org/10.5206/cjsotl-rcacea.2016.1.3

Santoveña Casal, S (2019) 'Effects of continuous assessment on the academic performance of future teachers', *Croatian Journal of Education,* 21 (3): 777-822. Available at: doi.org/10.15516/cje.v21i3.3013

Sarka, S, Lijalem, T and Shibiru, T (2017) 'Improving academic achievement through continuous assessment methods: In the case of year two students of animal and range sciences department in Wolaita Sodo University, Ethiopia', *Journal of Education and Practice,* 8 (4): 1-5.

Saville, K-M, Birdi, G, Hayes, S, Higson, H and Eperjesi, F (2020) 'Using strength-based approaches to fulfil academic potential in degree apprenticeships', *Higher Education Skills and Work-Based Learning,* 10 (4): 659-671. Available at: doi.org/10.1108/HESWBL-02-2019-0024

Scalise, K, Douskey, M and Stacy, A (2018) 'Measuring learning gains and examining implications for student success in STEM', *Higher Education Pedagogies,* 3 (1): 183-195. Available at: doi.org/10.1080/23752696.2018.1425096

Schillings, M, Roebertsen, H, Savelberg, H, van Dijk, A and Dolmans, D (2021) 'Improving the understanding of written peer feedback through face-to-face peer dialogue: Students' perspective', *Higher Education Research and Development,* 40 (5): 1100-1116. Available at: doi.org/10.1080/07294360.2020.1798889

Schonell, S and Macklin, R (2019) 'Work integrated learning initiatives: Live case studies as a mainstream WIL assessment', *Studies in Higher Education,* 44 (7): 1197-1208. Available at: doi.org/10.1080/03075079.2018.1425986

Schunn, C, Godley, A and DeMartino, S (2016) 'The reliability and validity of peer review of writing in high school AP English classes', *Journal of Adolescent and Adult Literacy,* 60 (1), 13-23. Available at: doi.org/10.1002/jaal.525

Schwab, K, Moseley, B and Dustin, D (2018) 'Grading grades as a measure of student learning', *Journal of Leisure Studies and Recreation Education,* 33 (2): 87-95. Available at: doi.org/10.1080/1937156X.2018.1513276

Sedghi, G and Rushworth, E (2017) 'The relation between multi-cultural group work and the integration of home and international students', *New Directions in the Teaching of Physical Sciences,* 12 (1). Available at: doi.org/10.29311/ndtps.v0i12.857

Seifert, L B, Herrera-Vizcaino, C, Herguth, P, Sterz, J and Sader, R (2020) 'Comparison of different feedback modalities for the training of procedural skills in oral and maxillofacial surgery: A blinded, randomized and controlled study', *BMC Medical Education,* 20 (1). Available at: doi.org/10.1186/s12909-020-02222-1

Seifert, L B, Socolan, O, Sader, R, Rüsseler, M and Sterz, J (2019) 'Virtual patients versus small-group teaching in the training of oral and maxillofacial surgery: A randomized controlled trial', *BMC Medical Education,* 19 (1). Available at: doi.org/10.1186/s12909-019-1887-1

Seifert, T and Feliks, O (2019) 'Online self-assessment and peer-assessment as a tool to enhance student-teachers' assessment skills', *Assessment and Evaluation in Higher Education,* 44 (2): 169-185. Available at: doi.org/10.1080/02602938.2018.1487023

Selhorst, A L, Bao, M, Williams, L and Klein, E (2017) 'The effect of online discussion board frequency on student performance in adult learners', *Online Journal of Distance Learning Administration,* 20 (4).

Seymour, E, Hunter, A-B, Laursen, S and DeAntoni, T (2004) 'Establishing the benefits of research experiences for undergraduates: First findings from a three-year study', *Science Education,* 88: 493-594. Available at: doi.org/10.1002/sce.10131

Shafi, A, Hatley, J, Middleton, T, Millican, R and Templeton, S (2018) 'The role of assessment feedback in developing academic buoyancy', *Assessment and Evaluation in Higher Education,* 43 (3): 415-427. Available at: doi.org/10.1080/02602938.2017.1356265

Shahnia, F and Yengejeh, H H (2019) 'Various interactive and self-learning focused tutorial activities in the power electronic course', *Transactions on Education*, 62 (4): 246-255. Available at: doi.org/10.1109/TE.2019.2892672

Shahriari-Rad, A, Cox, M and Woolford, M (2017) 'Clinical skills acquisition: Rethinking assessment using a virtual haptic simulator', *Technology, Knowledge and Learning,* 22 (2): 185-197. Available at: doi.org/10.1007/s10758-017-9308-1

Simpson, G and Clifton, J (2015) 'Assessing postgraduate student perceptions and measures of learning in a peer review feedback process', *Assessment and Evaluation in Higher Education,* 41 (4): 501-514. Available at: doi.org/10.1080/02602938.2015.1026874

Simpson, J (2016) 'Authentic learning - does it improve pass rates and student satisfaction?', *Journal of Perspectives in Applied Academic Practice,* 4 (2): 62-70. Available at: jpaap.napier.ac.uk/index.php/JPAAP/article/view/172

Slavin, R E (2010) 'Co-operative learning: What makes group-work work', in Hanna, D, David, I and Francisco, B (eds) *The Nature of Learning: Using Research to Inspire Practice.* Chicago: OECD Publishing, pp 161-178. Available at: doi.org/10.1787/9789264086487-9-en

Smith, A C, Ralph, B C W, MacLeod, C M and Smilek, D (2019) 'Test feedback and learning: Student preferences and perceived influence', *Scholarship of Teaching and Learning in Psychology,* 5 (4): 255-264. Available at: doi.org/10.1037/stl0000140

Sobhanzadeh, M and Zizler, P (2021) 'Selective assessment in introductory physics labatorials', *Physics Teacher,* 59 (2): 114-116. Available at: doi.org/10.1119/10.0003465

Sokhanvar, Z, Salehi, K and Sokhanvar, F (2021) 'Advantages of authentic assessment for improving the learning experience and employability skills of higher education students: A systematic literature review', *Studies in Educational Evaluation,* 70. Available at: doi.org/10.1016/j.stueduc.2021.101030

Sparrow, A, Smith, S, Petronzi, D, Wilson, H, Roeschlaub, S and Smith, M (2020) 'Student autonomy of feedback format in higher education and perceived functional behaviours for academic development', *Journal of Pedagogical Research,* 4 (2): 98-111. Available at: doi.org/10.33902/JPR.2020060191

Sridharan, B, Tai, J and Boud, D (2019) 'Does the use of summative peer assessment in collaborative group work inhibit good judgement?', *Higher Education,* 77 (5): 853-870. doi.org/10.1007/s10734-018-0305-7

Sritrakarn, N (2018) 'A comparison of teacher's and senior students' feedback: Student attitudes and their writing improvement', *The Journal of Asia TEFL,* 15 (2): 329-348. Available at: doi.org/10.18823/asiatefl.2018.15.2.5.329

Stanford, J S, Rocheleau, S E, Smith, K P W and Mohan, J (2017) 'Early undergraduate research experiences lead to similar learning gains for STEM and non-STEM undergraduates', *Studies in Higher Education,* 42 (1): 115-129. Available at: doi.org/10.1080/03075079.2015.1035248

Stannard, R and Mann, S (2018) 'Using screen capture feedback to establish social presence and increase student engagement: A genuine innovation in feedback', in Xiang, C H (ed) *Cases on Audio-Visual Media in Language Education.* Hershey, PA: IGI Global, pp 93-116. Available at: doi.org/10.4018/978-1-5225-2724-4.ch005

Steen-Utheim, A and Wittek, A L (2017) 'Dialogic feedback and potentialities for student learning', *Learning, Culture and Social Interaction,* 15: 18-30. Available at: doi.org/10.1016/j.lcsi.2017.06.002

Stobart, G (2008) *Testing times: The uses and abuses of assessment.* London: Routledge. Available at: doi.org/10.4324/9780203930502

Stowe, R L and Cooper, M M (2017) 'Practicing what we preach: Assessing critical thinking in organic chemistry', *Journal of Chemical Education,* 94 (12): 1852-1859. Available at: doi.org/10.1021/acs.jchemed.7b00335

Strijbos, J-W, Narciss, S and Dünnebier, K (2010) 'Peer feedback content and sender's competence level in academic writing revision tasks: Are they critical for feedback perceptions and efficiency?', *Learning and Instruction,* 20 (4): 291-303. Available at: doi.org/10.1016/j.learninstruc.2009.08.008

Struyven, K, Dochy, F and Janssens, S (2005) 'Students' perceptions about evaluation and assessment in higher education: A review', *Assessment and Evaluation in Higher Education,* 30 (4): 325-341. Available at: doi.org/10.1080/02602930500099102

Sutherland, D, Warwick, P, Anderson, J and Learmonth, M (2018) 'How do quality of teaching, assessment and feedback drive undergraduate course satisfaction in UK business schools? A comparative analysis with non-business school courses using the UK National Student Survey', *Journal of Management Education,* 42 (5): 618-649. Available at: doi.org/10.1177/1052562918787849

Svirko, E and Mellanby, J (2017) 'Teaching neuroanatomy using computerşaided learning: What makes for successful outcomes?', *Anatomical Sciences Education,* 10 (6): 560-569. Available at: doi.org/10.1002/ase.1694

Taylor, A T S, Olofson, E L and Novak, W R P (2017) 'Enhancing student retention of prerequisite knowledge through pre-class activities and in-class reinforcement', *Biochemistry and Molecular Biology Education,* 45 (2): 97-104. Available at: doi.org/10.1002/bmb.20992

Taylor, S, Ryan, M and Elphinstone, L (2021) 'Generating genuine inclusion in higher education utilising an original, transferable and customisable model for teaching and assessing reflective learning', *Reflective Practice,* 22 (4): 531-549. Available at: doi.org/10.1080/14623943.2021.1933408

Teichgräber, U, Ingwersen, M, Bürckenmeyer, F, Malouhi, A, Arndt, C, Herzog, A, Franiel, T, Mentzel, H-J and Aschenbach, R (2021) 'Structured work-based learning in undergraduate clinical radiology immersion experience', *BMC Medical Education,* 21 (1). Available at: doi.org/10.1186/s12909-021-02592-0

Terry, R, Hing, W, Orr, R and Milne, N (2017) 'Do coursework summative assessments predict clinical performance? A systematic review', *BMC Medical Education,* 17. Available at: doi.org/10.1186/s12909-017-0878-3

Thai, N T T, De Wever, B and Valcke, M (2020) 'Feedback: An important key in the online environment of a flipped classroom setting', *Interactive Learning Environments,* 20: 1-14. Available at: doi.org/10.1080/10494820.2020.1815218

Thijssen, D H J, Hopman, M T E, van Wijngaarden, M T, Hoenderop, J G J, Bindels, R J M and Eijsvogels, T M H (2019) 'The impact of feedback during formative testing on study behaviour and performance of (bio)medical students: A randomised controlled study', *BMC Medical Education,* 19 (1): 97. Available at: doi.org/10.1186/s12909-019-1534-x

Thomas, L J, Parsons, M and Whitcombe, D (2019) 'Assessment in smart learning environments: Psychological factors affecting perceived learning', *Computers in Human Behaviour,* 95: 197-207. Available at: doi.org/10.1016/j.chb.2018.11.037

Thompson, J and Houston, D (2020) 'Programmatic assessment condensed: Introducing progress testing approaches to a single semester paramedic subject', *Journal of University Teaching and Learning Practice,* 17 (3). Available at: doi.org/10.53761/1.17.3.14

Tila, D and Levy, D (2020) 'Revising online assignments and the impact on student performance at a community college', *Community College Journal of Research and Practice,* 44 (3): 163-180. Available at: doi.org/10.1080/10668926.2018.1564089

To, J and Liu, Y (2018) 'Using peer and teacher-student exemplar dialogues to unpack assessment standards: Challenges and possibilities', *Assessment and Evaluation in Higher Education,* 43 (3): 449-460. Available at: doi.org/10.1080/02602938.2017.1356907

To, J, Panadero, E and Carless, D (2021) 'A systematic review of the educational uses and effects of exemplars', *Assessment and Evaluation in Higher Education.* Available at: doi.org/10.1080/02602938.2021.2011134

Torda, A (2020) 'CLASSIE teaching – using virtual reality to incorporate medical ethics into clinical decision making', *BMC Medical Education,* 20 (1). Available at: doi.org/10.1186/s12909-020-02217-y

Tucker, R and Abbasi, N (2016) 'Bad attitudes: Why design students dislike teamwork', *Journal of Learning Design,* 9 (1): 1-20. Available at: doi.org/10.5204/jld.v9i1.227

Turner, W and West, J (2013) 'Assessment for 'digital first language' speakers: Online video assessment and feedback in higher education', *International Journal of Teaching and Learning in Higher Education,* 25: 288-296.

Umar, A M (2018) 'The impact of assessment for learning on students' achievement in English for specific purposes', *English Language Teaching,* 11 (2): 16-25. Available at: doi.org/10.5539/elt.v11n2p15

United States Department of Education's What Works Clearinghouse in Education (2018) Available at: ies.ed.gov/ncee/wwc/Docs/referenceresources/WWC-Standards-Handbook-v4-1-508.pdf

Uribe, S N and Vaughan, M (2017) 'Facilitating student learning in distance education: A case study on the development and implementation of a multifaceted feedback system', Distance Education, 38 (3): 288-301. Available at: doi.org/10.1080/01587919.2017.1369005

Uymaz, E (2019) 'The effects of peer feedback on the essay writing performances of EFL students', *International Journal of Curriculum and Instruction,* 11 (2): 20-37.

Vaessen, B E, van den Beemt, A, van de Watering, G, van Meeuwen, L, Lemmens, L and den Brok, P (2017) 'Students' perception of frequent assessments and its relation to motivation and grades in a statistics course: A pilot study', *Assessment and Evaluation in Higher Education,* 42 (6): 872-886. Available at: doi.org/10.1080/02602938.2016.1204532

Valenzuela, L, Jerez, O M, Hasbun, B A, Pizarro, V, Valenzuela, G and Orsini, C A (2018) 'Closing the gap between business undergraduate education and the organisational environment: A Chilean case study applying experiential learning theory', *Innovations in Education and Teaching International,* 55 (5): 566-575.

van den Bos, A H and Tan, E (2019) 'Effects of anonymity on online peer review in second-language writing', *Computers and Education,* 142. Available at: https://doi.org/10.1016/j.compedu.2019.103638

van der Kleij, F M, Adie, L E, and Cumming, J J (2019). A meta-review of the student role in feedback. International Journal of Educational Research, 98, 303–323. doi.org/10.1016/j.ijer.2019.09.005

van Esch, P, von der Heidt, T, Frethey-Bentham, C and Northey, G (2020) 'The effect of marketing simulations on student engagement and academic outcomes', *Marketing Education Review,* 30 (1): 43-56. Available at: doi.org/10.1080/10528008.2020.1713003

van Popta, E, Kral, M, Camp, G, Martens, R L and Simons, P R-J (2017) 'Exploring the value of peer feedback in online learning for the provider', *Educational Research Review,* 20: 24-34. Available at: doi.org/10.1016/j.edurev.2016.10.003.

Vardi, I (2013) 'Effectively feeding forward from one written assessment task to the next', *Assessment and Evaluation in Higher Education,* 38 (5): 599-610. Available at: doi.org/10.1080/02602938.2012.670197

Vattøy, K-D, Gamlem, S M and Rogne, W M (2021) 'Examining students' feedback engagement and assessment experiences: A mixed study', *Studies in Higher Education,* 46 (11): 2325-2337. Available at: doi.org/10.1080/03075079.2020.1723523

Vattøy, K-D and Smith, K (2019) 'Students' perceptions of teachers' feedback practice in teaching English as a foreign language', *Teaching and Teacher Education,* 85: 260-268. Available at: doi.org/10.1016/j.tate.2019.06.024

Vidiella, M-C and Garcia, E C (2016) 'Regulating the writing process in the teacher training practicum guided by feedback at the University of Barcelona', *Cambridge Journal of Education,* 46 (1): 55-79. Available at: doi.org/10.1080/0305764X.2015.1011083

Villarroel, V, Bloxham, S, Bruna, D, Bruna, C and Herrera-Seda, C (2018) 'Authentic assessment: Creating a blueprint for course design', *Assessment and Evaluation in Higher Education,* 43 (5): 840-854. Available at: doi.org/10.1080/02602938.2017.1412396

Villarroel, V, Boud, D, Bloxham, S and Bruna, C (2019) 'Using principles of authentic assessment to redesign written examinations and tests', *Innovations in Education and Teaching International,* 57 (1): 38-49. Available at: doi.org/10.1080/14703297.2018.1564882

Walker, R and Handley, Z (2016) 'Designing for learner engagement with computer-based testing', *Research in Learning Technology,* 24. Available at: doi.org/10.3402/rlt.v24.30083

Way, K A, Burrell, L, D'Allura, L and Ashford-Rowe, K (2021) 'Empirical investigation of authentic assessment theory: An application in online courses using mimetic simulation created in university learning management ecosystems', *Assessment and Evaluation in Higher Education,* 46 (1): 17-35. Available at: doi.org/10.1080/02602938.2020.1740647

Wei, W, Sun, Y and Xu, X (2021) 'Investigating the impact of increased student feedback literacy level on their expectations on university teachers' feedback', *Assessment and Evaluation in Higher Education,* 46 (7): 1092-1103. Available at: doi.org/10.1080/02602938.2020.1846017

West, J and Turner, W (2016) 'Enhancing the assessment experience: Improving student perceptions, engagement and understanding using online video feedback', *Innovations in Education and Teaching International,* 53 (4): 400-410. Available at: doi.org/10.1080/14703297.2014.1003954

Wilkin, C L (2017) 'Enhancing critical thinking: Accounting students' perceptions', *Education and Training,* 59 (1): 15-30. Available at: doi.org/10.1108/ET-01-2015-0007

Wilkinson, K, Dafoulas, G, Garelick, H and Huyck, C (2020) 'Are quiz-games an effective revision tool in anatomical sciences for higher education and what do students think of them?', *British Journal of Educational Technology, 5*1 (3): 761-777.

Winstone, N and Boud, D (2019) 'Exploring cultures of feedback practice: The adoption of learning-focused feedback practices in the UK and Australia', *Higher Education Research and Development,* 38 (2): 411-425. Available at: doi.org/10.1080/07294360.2018.1532985

Winstone, N, Pitt, E and Nash, R (2021) 'Educators' perceptions of responsibility-sharing in feedback processes', *Assessment and Evaluation in Higher Education,* 46 (1): 118-131. Available at: doi.org/10.1080/02602938.2020.1748569

Winstone, N E and Carless, D (2019) *Designing effective feedback processes in higher education: A learning-focused approach.* London: Routledge.

Winstone, N E, Nash, R A, Rowntree, J and Menezes, R (2016) 'What do students want most from written feedback information? Distinguishing necessities from luxuries using a budgeting methodology', *Assessment and Evaluation in Higher Education,* 41 (8): 1237-1253. Available at: doi.org/10.1080/02602938.2015.1075956

Winstone, N E, Nash, R, Rowntree, J and Parker, M (2017) ''It'd be useful, but I wouldn't use it'. Barriers to university students' feedback seeking and recipience', *Studies in Higher Education,* 42 (11): 2026-2041.

Wisniewski, B, Zierer, K and Hattie, J (2020) 'The power of feedback revisited: A meta- analysis of educational feedback research', *Frontiers in Psychology,* 10: 1-14. Available at: doi.org/10.3389/fpsyg.2019.03087

Wolstencroft, P and de Main, L (2021) ''Why didn't you tell me that before?' Engaging undergraduate students in feedback and feedforward within UK higher education', *Journal of Further and Higher Education,* 45 (3): 312-323. Available at: doi.org/10.1080/0309877X.2020.1759517

Wu, Y and Schunn, C D (2021) 'From plans to actions: A process model for why feedback features influence feedback implementation', *Instructional Science,* 49: 365-394. Available at: doi.org/10.1007/s11251-021-09546-5

Yang, M, Tai, M and Lim, C P (2016) 'The role of e-portfolios in supporting productive learning', *British Journal of Educational Technology,* 47 (6): 1276-1286. Available at: doi.org/10.1111/bjet.12316

Young, S C, Colabroy, K L and Baar, M R (2016) 'Comparable educational benefits in half the time: An alternating organic chemistry laboratory sequence targeting prehealth students', *Journal of Chemical Education,* 93 (12): 2004-2011. Available at: doi.org/10.1021/acs.jchemed.6b00254

Zamberlan, L and Wilson, S E (2017) 'Conversation leading to progress: Student perceptions of peer tutors' contribution to enhancing creativity and collaboration in a first-year design studio', *Journal of Peer Learning,* 10: 59-75.

Zavhorodnia, V M, Slavko, A S, Degtyarev, S I and Polyakova, L G (2021) 'The use of practice-based assessments in preparing humanities and social sciences specialists: The case of Sumy State University (Ukraine)', *European Journal of Contemporary Education,* 10 (3): 768-782. Available at: doi.org/10.13187/ejced.2021.3.768

Zeglen, E and Rosendale, J A (2018) 'Increasing online information retention: Analysing the effects of visual hints and feedback in educational games', *Journal of Open, Flexible and Distance Learning,* 22 (1): 22-33.

Zhang, X and McEneaney, J E (2019) 'What is the influence of peer feedback and author response on Chinese university students' English writing performance?', *Reading Research Quarterly.* Available at: doi.org/10.1002/rrq.259

Zhao, K, Zhou, J M and Dawson, P (2021) 'Using student-instructor co-constructed rubrics in signature assessment for business students: Benefits and challenges', *Assessment in Education: Principles, Policy and Practice,* 28 (2): 170-190. Available at: doi.org/10.1080/096959 4X.2021.1908225

Zheng, L, Cui, P, Li, X and Huang, R (2018) 'Synchronous discussion between assessors and assessees in web-based peer assessment: Impact on writing performance, feedback quality, meta-cognitive awareness and self-efficacy', *Assessment and Evaluation in Higher Education,* 43 (3): 500-514. Available at: doi.org/10.1080/02602938.2017.1370533

Zheng, L, Zhang, X and Cui, P (2020) 'The role of technology-facilitated peer assessment and supporting strategies: A meta-analysis', *Assessment and Evaluation in Higher Education,* 45 (3): 372-386. Available at: doi.org/10.1080/02602938.2019.1644603

Zhu, Q and Carless, D (2018) 'Dialogue within peer feedback processes: Clarification and negotiation of meaning', *Higher Education Research and Development,* 37 (4): 883-897. Available at: doi.org/10.1 080/07294360.2018.1446417

Zong, Z, Schunn, C D and Wang, Y (2021) 'What aspects of online peer feedback robustly predict growth in students' task performance?', *Computers in Human Behaviour,* 124. Available at: doi.org/10.1016/j.chb.2021.106924

Zou, Y, Schunn, C D, Wang, Y and Zhang, F (2018) 'Student attitudes that predict participation in peer assessment', *Assessment and Evaluation in Higher Education,* 43 (5): 800–811.

Contact us

General enquiries

+44 (0) 3300 416201

enquiries@advance-he.ac.uk

www.advance-he.ac.uk

Media enquiries

+44 (0) 1904 717500

communications@advance-he.ac.uk

www.advance-he.ac.uk/contact-us

@AdvanceHE

Advance HE helps HE and research be the best they can be.

We are a member-led, sector-owned charity that works with institutions and higher education across the world to improve higher education for staff, students and society. We are experts in higher education with a particular focus on enhancing teaching and learning, effective governance, leadership development and tackling inequalities through our equality, diversity and inclusion (EDI) work.

Our strategic goals to enhance confidence and trust in HE, address systemic inequalities and advance education to meet the evolving needs of students and society, support the work of our members and the HE sector. We deliver our support through professional development programmes and events, Fellowships, awards, student surveys and research, providing strategic change and consultancy services and through membership (including accreditation of teaching and learning, equality charters, knowledge and resources).

Advance HE is a company limited by guarantee registered in England and Wales no. 04931031. Registered as a charity in England and Wales no. 1101607 Registered as a charity in Scotland no. SC043946. The Advance HE logo should not be used without our permission.

© 2022 Advance HE. All rights reserved.

The views expressed in this publication are those of the author and not necessarily those of Advance HE. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any storage and retrieval system without the written permission of the copyright owner. Such permission will normally be granted for non-commercial, educational purposes provided that due acknowledgement is given.

To request copies of this report in large print or in a different format, please contact the Marketing and Communications Team at Advance HE: +44 (0) 3300 416201 or publications@advance-he.ac.uk