



# Kent Academic Repository

**Winstone, Naomi and Pitt, Edd (2025) *Approaches to feedback on examination performance: research, policy, and practice*. Assessment and Evaluation in Higher Education, 50 (6). pp. 876-896. ISSN 0260-2938.**

## Downloaded from

<https://kar.kent.ac.uk/109339/> The University of Kent's Academic Repository KAR

## The version of record is available from

<https://doi.org/10.1080/02602938.2025.2476622>

## This document version

Publisher pdf

## DOI for this version

## Licence for this version

CC BY-NC-ND (Attribution-NonCommercial-NoDerivatives)

## Additional information

## Versions of research works

### Versions of Record

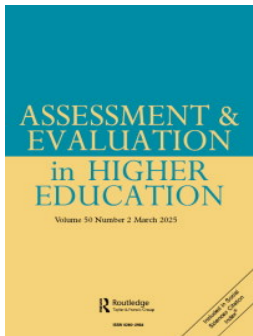
If this version is the version of record, it is the same as the published version available on the publisher's web site. Cite as the published version.

### Author Accepted Manuscripts

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding. Cite as Surname, Initial. (Year) 'Title of article'. To be published in **Title of Journal**, Volume and issue numbers [peer-reviewed accepted version]. Available at: DOI or URL (Accessed: date).

## Enquiries

If you have questions about this document contact [ResearchSupport@kent.ac.uk](mailto:ResearchSupport@kent.ac.uk). Please include the URL of the record in KAR. If you believe that your, or a third party's rights have been compromised through this document please see our [Take Down policy](https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies) (available from <https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies>).



## Approaches to feedback on examination performance: research, policy, and practice

Naomi E. Winstone & Edd Pitt

**To cite this article:** Naomi E. Winstone & Edd Pitt (20 Mar 2025): Approaches to feedback on examination performance: research, policy, and practice, Assessment & Evaluation in Higher Education, DOI: [10.1080/02602938.2025.2476622](https://doi.org/10.1080/02602938.2025.2476622)

**To link to this article:** <https://doi.org/10.1080/02602938.2025.2476622>



© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 20 Mar 2025.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)

# Approaches to feedback on examination performance: research, policy, and practice

Naomi E. Winstone<sup>a</sup>  and Edd Pitt<sup>b</sup> 

<sup>a</sup>Surrey Institute of Education, University of Surrey, Surrey, UK; <sup>b</sup>Centre for the Study of Higher Education, University of Kent, Canterbury, UK

## ABSTRACT



Examinations are viewed as one of the least inclusive forms of assessment. Attainment gaps between different student groups (e.g. based on ethnicity) are more pronounced in examinations than in coursework, perhaps because feedback on performance is commonplace for coursework yet rare for examinations. We developed, and explored evidence for, a novel taxonomy of approaches to examination feedback through a systematic literature review, policy analysis, and survey of UK academics. Our synthetic analysis indicated a disjuncture between research and policy/practice: there is very little research focusing on the most commonly-implemented forms of examination feedback in higher education, whilst the majority of research evidence focuses on the approaches to examination feedback that are seen as the most challenging to implement in practice. Our analysis also uncovered policies that encouraged exclusionary rather than inclusive approaches to examination feedback. We argue for a renewed research focus on examination feedback and differential attainment.

## KEYWORDS

Assessment; examination; feedback; multiple-choice testing

Few forms of assessment in higher education are subject to as much debate as examinations. Whilst their use in higher education remains commonplace (Buckley 2024; French, Dickerson, and Mulder 2023), their pedagogic value has long been the subject of critique (e.g. Biggs 2001; Gibbs 1992). In the current era of digital transformation, the perceived need to ‘secure’ assessment against the rise of Generative AI tools means that examinations continue to be lauded as the most robust, rigorous, and appropriate form of assessment in higher education (e.g. Sims 2023). Add the concerns over the rising cost of teaching and growing student numbers, and there is a risk of the unprincipled use of examinations becoming more widespread (French, Dickerson, and Mulder 2023).

Pragmatic arguments for the use of examinations are not new; it has long been argued that exams are easier (and more cost-effective) to mark than coursework assignments (Rowntree 1987; Richardson 2015). However, feedback on examinations is rare (Gibbs and Lucas 1997; Richardson 2015). Indeed, examinations have been described as a ‘feedback desert’ and presenting a ‘feedback gap’ (Scoles, Huxham, and McArthur 2013). On this basis, our motivation for the present study was to systematically explore the prevalence and practice of feedback on examinations, as represented in the literature, in policy, and in educators’ practice. In the sections

**CONTACT** Edd Pitt  [e.pitt@kent.ac.uk](mailto:e.pitt@kent.ac.uk)  Centre for the Study of Higher Education, University of Kent, Canterbury, UK.  
© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

that follow, we first present a summary of literature on the value of exams as a form of assessment, before exploring the role of feedback on examination performance. This leads to the rationale for our study: to develop a taxonomy of approaches to exam feedback, explore barriers to the implementation of exam feedback, and review evidence for the effectiveness of exam feedback.

### ***Examining examinations as a form of assessment***

Examinations are often cast as a form of assessment fraught with challenges. For example, in a recent review (French, Dickerson, and Mulder 2023), the authors outline drawbacks of exams as a choice of assessment including: encouraging surface learning in students and teaching to the test or curriculum narrowing for teachers; limited authenticity; limited reliability/validity; and creating stress and anxiety. There are two challenges with examinations highlighted by French, Dickerson, and Mulder (2023) which we see as particularly pertinent: examinations are often a site of inequity in the performance of student groups; and examinations are not commonly a source of feedback information for students in ways that could help them improve future performance. We believe that these two issues are interrelated, but first we consider issues relating to the inclusivity of examinations as a form of assessment, before turning to consider the challenges created by limited feedback opportunities.

Examinations have been described as one of the least inclusive forms of assessment (Tomas and Jessop 2019). Student performance is commonly lower on exams in comparison to coursework (Bridges et al. 2002; Simonite 2003), and exams can privilege those with prior experience of exams and well-developed exam technique (Blair et al. 2014). In international higher education systems, there are well-established, yet largely unexplained, gaps in attainment between Black and White students (e.g. Stegers-Jager, Brommet, and Themmen 2016; Farmer 2017). Examinations are a significant driver of these attainment gaps: the extent to which White students outperform Black students is more pronounced in examinations than in coursework assessment (e.g. Cramer 2021; Shaw and Tranter 2021), and examinations can disadvantage marginalised groups (French, Dickerson, and Mulder 2023).

Unlike coursework, students seldom receive feedback information on examination performance and areas for improvement (Brown and Glasner 2003; Weaver 2006). Thus, the challenge of differential attainment in education, and in particular the more prominent attainment gaps for examinations than coursework, may well be exacerbated by limited feedback on examinations.

### ***Examinations and the role of feedback information***

Despite the potential for examinations to serve a formative purpose, a lack of feedback on examinations means that the potential for learning is often inhibited by the limited opportunity to actively engage with and implement feedback (Wieman, Rieger, and Heiner 2014; Winstone and Carless 2019). The relevance and value of exams are questioned by students, especially when compared to other assessments such as essays, where comprehensive feedback is more common. In a study of UK undergraduate students, just 28% of students reported that they received written feedback information on their examination performance, with 26% reporting that they had the opportunity to gain verbal feedback on their performance (Blair et al. 2014). Challenges to examination feedback include the need to rapidly mark papers resulting in limited comments or annotations, the common end-of-semester timing of examinations making the collection of feedback less likely, and the limited opportunity to apply feedback to a subsequent examination (Blair et al. 2014). Without feedback, students can remain unsure about how to perform well in examinations, further diminishing students' confidence in the value of this form of assessment (Blair et al. 2014).

Recognising the importance of improving examination feedback, Blair et al. (2014) advocate for the establishment of a 'norm' of feedback on examinations. They argue that '...improving exam feedback is an important element in attempts to encourage better student engagement and 'rescue' the exam as a valuable form of assessment which is judged more favourably by students' (Blair et al. 2014, 1046). This call to action emphasises the need for an equitable approach to constructive feedback information on examination performance.

### ***The present study***

In response to the identified challenges with limited feedback on examinations and calls for more widespread adoption of feedback on examinations, we wanted to first understand the strength of evidence for the impact of feedback on examinations, how examination feedback is commonly framed within university policies, and to hear from educators themselves about their own experiences and reservations about the use of examination feedback. The latter focus is particularly pertinent given recent evidence that research on examinations rarely connects with the constraints faced by practitioners (Buckley 2024). Thus, we used a systematic literature review, policy analysis, and a survey to develop a taxonomy of examination feedback formats in higher education. We also sought to identify barriers to the implementation of examination feedback, and to synthesise evidence for the impact of examination feedback on students' learning.

## **Methods**

### ***Compiling the data corpus***

Our data corpus consisted of 63 research articles on examination feedback obtained through database searches and citation chasing; feedback policies from 100 UK universities; and survey data from 116 UK academics.

### ***Systematic literature review***

We used the database Scopus to identify literature published between January 1980 and June 2023 that contained within the title or abstract the keywords 'feedback' and 'examination'/'exam'/'test'. We limited our search to peer-reviewed journal articles published in English, and literature focused on higher education. Our inclusion/exclusion criteria required articles that related to: 1) feedback, formative or summative, as provided to students by teachers, peers, or computer systems, not feedback from students to teachers (i.e. evaluations of teaching); 2) undergraduate or master's education, not doctoral education; 3) feedback on exams or tests, not coursework assignments; 4) examinations or tests taken as part of a higher education course, not examinations with the purpose of determining admission or entrance to higher education courses. Because our focus was on feedback on examinations, we excluded articles where exam scores/grades were only used as an outcome measure.

An initial phase of title and abstract screening reduced the total number of hits from 1058 to 164 (see Figure 1). Following a subsequent phase of full text screening, we excluded a further 115 articles (see Figure 1 for exclusion decisions) and also conducted a phase of backward and forward citation chasing, resulting in identification of a further 40 potential articles, of which 26 were excluded, leaving a final pool of 63 articles for inclusion in the review. The articles represented research from across the world as indicated by the institutional affiliation of the first author: North America (35); Europe (16); Asia (6); Oceania (4); and Africa (2).

### ***Policies***

The data corpus consisted of extracts from policy documentation (either learning and teaching strategy or policies/guidelines for assessment and feedback) published on the websites of 120 UK

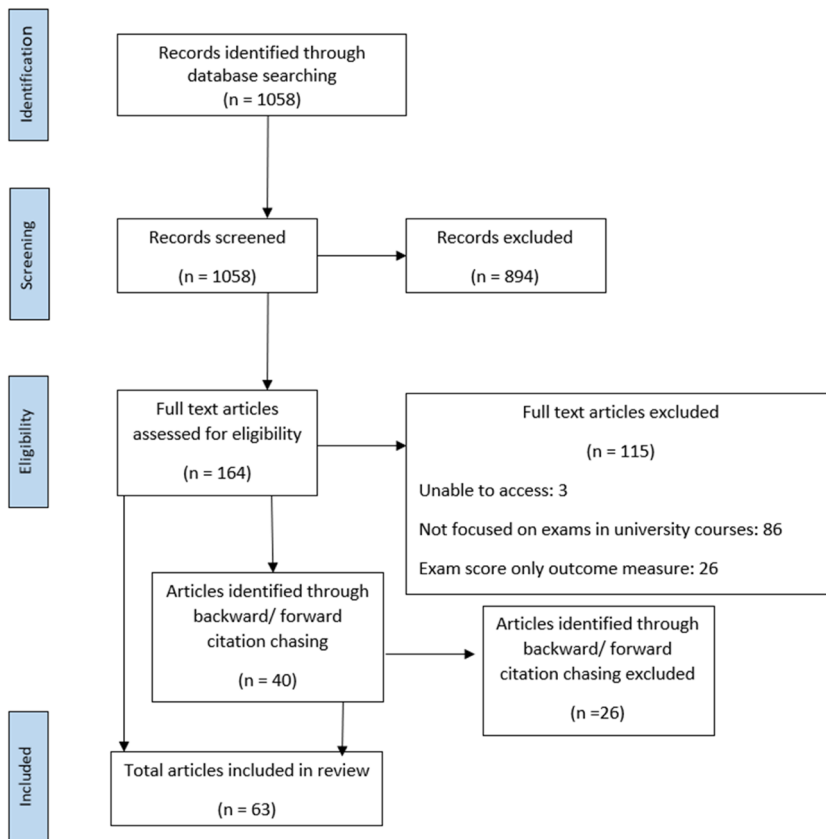


Figure 1. PRISMA Diagram.

Universities: 30 research-intensive (RI) Universities incorporating the Russell Group; 68 ‘Post-1992’ teaching-intensive (TI) Universities; and 22 dual-intensive (DI) Universities established prior to 1992. A total of 7 Universities (1 RI; 6 TI) had no policy/guidance documentation available on their website, and the documentation for 13 Universities made no reference to exam feedback (1 RI; 9 TI; 3 DI). Thus, the final sample consisted of extracts from documentation for 100 Universities (28 RI; 53 TI; 19 DI). A research assistant was trained to extract sections of the documentation that pertained to exam feedback. A total of 266 statements were extracted.

### Survey

Data were collected from 116 UK-based academic staff recruited *via* national mailing lists for learning and teaching (Staff and Educational Development Association (SEDA), and broad External Examiners UK email network lists). Of the 116 participants, 60 identified as female, 54 identified as male, and 2 identified as ‘other’. Participants had worked in higher education for between 3 to 43 years ( $M=18.86$ ;  $SD=8.56$ ). Most participants ( $N=59$ ) reported working in a TI University; the remaining participants reported working in an RI University ( $N=19$ ) or a DI University ( $N=33$ ). This question was not answered by 5 participants. Participants represented a range of disciplinary areas, categorized using the Becher-Biglan typology (Becher and Trowler 2001): Hard Pure ( $N=11$ ); Hard Applied ( $N=24$ ); Soft Pure ( $N=20$ ); and Soft Applied ( $N=61$ ). The Hard Pure category included disciplines from the natural sciences (e.g. Physics, Chemistry, Mathematics). Hard Applied disciplines were primarily from fields focused on practical application of scientific knowledge (e.g. Engineering, Computer Science, Medicine). Soft Pure disciplines represented the humanities

and social sciences (e.g. History, Philosophy, Anthropology), while Soft Applied disciplines represented practice-oriented fields within the social sciences (e.g. Education, Business Studies, Law, Social Work).

Participants were asked whether, in their own practice, they provide feedback information to students on examinations. If they answered 'yes' to this question, they were asked to describe how they typically provide feedback on examination performance (e.g. in person, *via* email, within a feedback proforma). If they answered 'no' to this question, they were asked to describe why they do not typically provide feedback on examination performance. All participants were then asked the same four open ended questions: 1) What is the perception of your students regarding feedback on examinations?; 2) What do you perceive as the benefits of providing feedback on examinations?; 3) What do you perceive as the challenges with providing feedback on examinations; and 4) Is there anything else you would like to say about the practice of providing feedback on examinations?

### ***Coding and analysis***

Initial inductive coding across the data corpus identified 24 distinct approaches to examination feedback. These codes were refined into a two-dimensional classification system, distinguishing between individual (specific to a student) and generic (group performance) feedback on one axis and provided (given routinely) and requested (available on demand) comments on the second axis. The resulting taxonomy (Individual Provided, Individual Requested, Generic Provided, and Generic Requested) was applied to the full dataset in a second round of deductive coding. A second coder independently coded 20% of the data. Overall agreement was 93%, with Cohen's Kappa of 0.94, 0.93, and 0.86 for policies, survey, and literature, respectively (all indicative of 'very good' agreement).

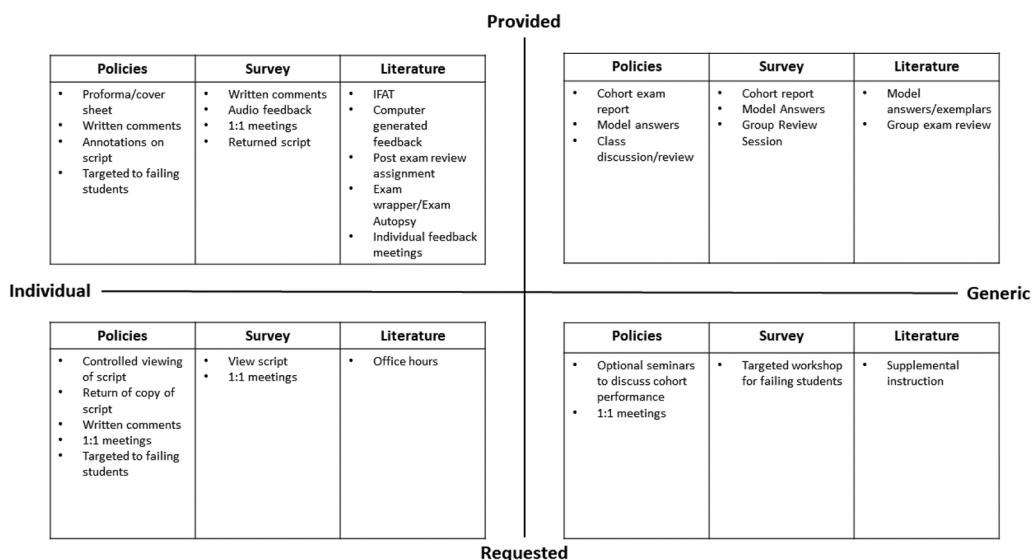
We analysed additional survey using an approach broadly following the principles of reflexive thematic analysis (Braun and Clarke 2021), involving data immersion, code generation, followed by the development and refinement of themes. The purpose of this approach was to provide a descriptive picture of participants' responses, rather than constituting a full qualitative analysis in its own right. Thus, we addressed cautions about the suitability of short free-text survey responses for rich qualitative analysis (LaDonna, Taylor, and Lingard 2018) by following recommendations to use these data to supplement our main analysis and to address an a priori research question.

## **Findings**

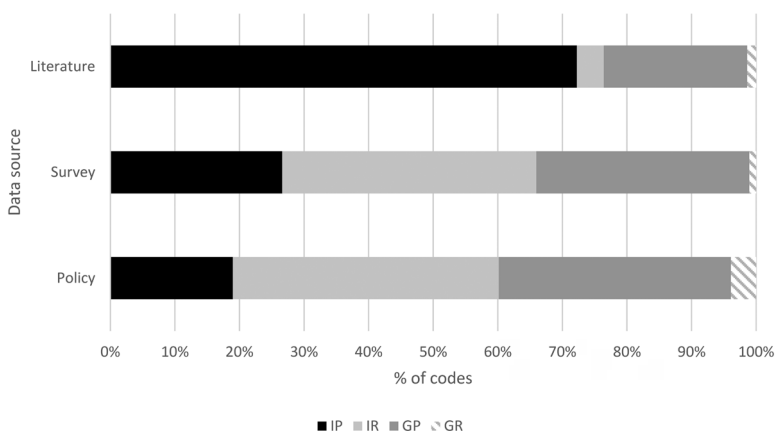
### ***The feedback taxonomy***

Within the four broad categories of examination Feedback (Individual Provided, IP; Individual Requested, IR; Generic Provided, GP; Generic Requested, GR), we identified 24 separate approaches to examination feedback (see Figure 2). Some of these approaches were unique to a particular category of examination feedback (e.g. Exam wrapper/Exam Autopsy, IP), whilst others were identified in more than one category (e.g. written comments, IP, IR).

Because we deductively coded policies, literature, and survey responses using the same coding scheme (IP, IR, GP, GR), we also compared the prevalence of each category of feedback as represented within each source of data (see Figure 3). This analysis revealed a stark difference in the focus of the literature on one hand, and policy and practice on the other. Most studies in the literature focused on Individual Provided feedback; in contrast, this form of feedback on examinations was the least frequently mentioned in policies and reports of practice by practitioners.



**Figure 2.** Examination feedback taxonomy.



**Figure 3.** Frequency of codes (IP, IR, GP, GR) within each data source (policies, survey, literature).

Below we present a synthetic analysis of data pertaining to each of the four categories in turn, integrating insights from the literature, policies, and reports of practice. First, we briefly outline findings from the survey which identified the reasons why feedback on examinations is not given.

### **Survey data: barriers to examination feedback**

Of our 116 participants, 87 (75%) reported that they provide feedback information to students on examination performance. Of the 29 participants who reported that they do not provide feedback information on examinations, 22 provided an explanation (see [Appendix A](#) for a breakdown of the prevalence of examination feedback by institution type and discipline area). Within these explanations, we identified three themes ([Table 1](#)).

All of these reasons represent what are perceived to be contextual constraints, rather than agentic decisions on the part of educators. In fact, participants expressed how they did not necessarily agree with policies that precluded examination feedback, and how they found the



**Table 1.** Themes representing barriers to examination feedback.

Theme	Description	Exemplar quotation
Status Quo	Feedback on Examinations is not given because that is the way it has always been	<i>there is no feedback given unless the student makes a complaint....it's a cultural thing 'we've always done it like that'....frustrating!</i> (Survey, P6, Business, DI)
Policy	University policy precludes the provision of feedback on examinations	<i>the University's policy around exams only allows for us to provide a mark for each student with no scope or opportunity to provide any feedback</i> (Survey, P25, Education, TI)
Workload	It is not possible to provide feedback on examinations because of student numbers	<i>There is a limited period in which to turn around exam marks so it is almost impossible to write feedback given the number of students who sit exams</i> (Survey, P10, Computer Science, DI)

maintenance of the status quo to be 'frustrating'. Having briefly considered the reasons why examination feedback may not be common practice, we now turn to our analysis of approaches to examination feedback that are practiced, against each quadrant of our taxonomy in turn.

### **Synthesis of findings: individual provided (IP) feedback**

Various approaches to providing individualised feedback on exams are detailed in the literature. A simple method includes performance breakdowns, such as scores for each question/topic or comparisons with class averages (e.g. Krippendorff, Bolender, and Kolesari 2008). Over 90% of medical students in Malau-Aduli et al. (2014) study found topic-specific feedback helpful, with 80% planning to use it for future revisions. In the context of a UK Pharmacy programme, Ellis and Barber (2016) developed 'Smallvoice' for personalised analysis of performance across different question types and topics, resulting in a 10–20% improvement in UK Pharmacy students' scores.

Essay-based exams, where answers are open-ended, benefit from specific feedback. For example, detailed feedback following an essay-based examination, either from an instructor or computer-generated, was associated with improvements in subsequent scores on a similar task (Lipnevich and Smith 2009). Students report positive perceptions of individualised feedback on examination performance, including computer-generated (Barker 2011), audio (Jordan 2004), or handwritten (Lee and Cha 2022) comments. Face-to-face discussions on examination performance can lead to gains in performance (Yopp and Rehberger 2009).

The majority of research on Individual Provided feedback focuses on multiple choice examinations (MCQs). MCQ items include distractors which, by exposing students to incorrect information, can lead students to acquire inaccurate knowledge. Feedback is important because it can reinforce correct knowledge whilst enabling errors and misunderstandings to be corrected. For example, Butler and Roediger (2008) demonstrated that feedback on a MCQ test increased correct responses and decreased number of intrusions on a later cued recall test. Feedback on MCQs also serves a metacognitive function (Butler, Karpicke, and Roediger 2008). As measured using a cued recall test, retention of correct (but low confidence) responses was doubled through provision of feedback in comparison to no feedback, thus demonstrating that feedback corrects metacognitive errors (i.e. when students possess correct knowledge in the absence of confidence in this knowledge). Simply marking answers as correct provides verification feedback that can maintain material in memory (Marsh et al. 2012). Nevertheless, grades and correct answers alone do not guarantee that students will review material and correct misunderstandings, especially if feedback is delayed (Douglas, Wilson, and Ennis 2012). Rostaminezhad (2019) found that immediate feedback could lead to stress, as knowledge of poor performance can discourage future efforts. Other studies found no benefit from immediate feedback on MCQs, even though students reported a preference for it and believed it improved their understanding (Badyal et al. 2019).

Some MCQ feedback practices require students to actively engage with feedback. For example, Fyfe et al. (2014) used a post-test reflection survey (PTRS) to encourage students to reflect on strengths, weaknesses, and preparation strategies based on feedback reports. Fyfe et al. found that older students, females, and those with more experience and higher performance were more likely to engage in reflection. Rozell et al. (2017) introduced a 'regrade' practice, where students could correct mistakes for partial credit. Students have one week to write responses to incorrectly answered questions, giving a thorough explanation of the reasoning behind the correct solution. Students have to describe why their original response was incorrect and why the correct answer is right. Their analysis demonstrated a correlation between performance on the regrade and performance on a subsequent exam.

The Immediate Feedback Answer Technique (IFAT; also termed 'answer-until-correct') is a well-established form of individual provided feedback. Students receive immediate feedback on their MCQ responses through a scratch-off answer form, transforming the test into a learning opportunity (Brosvic and Epstein 2007). In some uses of IFAT, students have the opportunity to select another answer, receiving partial credit if correct. Studies show IFAT can improve long-term retention and lead to higher subsequent exam scores (Brosvic et al. 2005; DiBattista et al. 2009). However, some studies indicate that a preference for IFAT does not necessarily lead to significant performance improvements (Bowman and Laurent 2011; Shoulders, Johnson, and Wiedenmann 2017). Persky and Pollack (2008) noted that students new to the format reported increased anxiety, although this was unrelated to overall performance.

Studies on the effects of the IFAT demonstrate evidence of positive impacts on student learning; for example, higher scores in subsequent examinations (Brosvic et al. 2005; DiBattista et al. 2009), superior long-term retention (Brosvic and Epstein 2007), and self-reported perceptions of performance improvements (Arkin and Schumann 1984; DiBattista, Mitterer, and Gosse 2004). Gaining feedback in real-time counters the 'Negative Suggestion Effect' in MCQ testing, which represents the reinforcement of incorrect knowledge when students leave an examination believing incorrectly answered items were answered correctly (Bowman and Laurent 2011).

Further mechanisms through which immediate feedback leads to performance improvements include correcting misconceptions (Shoulders, Johnson, and Wiedenmann 2017); helping students to understand their performance (Slepkov 2013); increasing confidence (Shoulders, Johnson, and Wiedenmann 2017); and providing a more engaging examination experience (DiBattista, Mitterer, and Gosse 2004; Slepkov 2013). Evidence also suggests that students hold more positive perceptions of courses when IFAT has been used in comparison to more traditional MCQ response formats (DiBattista, Mitterer, and Gosse 2004; DiBattista and Gosse 2006; Persky and Pollack 2008; Bowman and Laurent 2011). These preferences do not always translate into significant advantages in subsequent exam performance (e.g. Bowman and Laurent 2011; Shoulders, Johnson, and Wiedenmann 2017; Wood, Klausz, and MacNeil 2022), and the unfamiliarity of the IFAT approach can lead to student anxiety in the absence of sufficient practice (Persky and Pollack 2008). However, DiBattista and Gosse (2006) demonstrated that students with test anxiety were not disadvantaged by IFAT in comparison to standard response formats.

Another common approach to Individual Provided examination feedback is the use of post-exam review activities, which incorporates specific techniques such as the 'exam wrapper', the 'exam autopsy', and two-stage exams. Exam wrappers (e.g. Clark 2019) involve students reflecting upon their examination performance and areas of misunderstanding, correction of errors, as well as their preparation strategies for the examination and how they might go about preparation in the future (hence reflection on the stages 'wrapping' around the examination itself). For example, in a study with US Chemistry Undergraduates (Barnard and Sweeder 2020), following notification of their overall exam score, class time was used to provide students with their exam scripts, free of any indication of errors or correct answers. Through group discussion, students identified errors and collaboratively developed a better response to each question. Then students completed an Exam Wrapper exercise, reflecting on how they prepared for the exam,

their understanding of what each question was asking, where their misunderstanding might have arisen, any themes in their errors, and how they would change their exam preparation techniques for the future. In a similar activity designed to promote metacognition, students were given one week after their graded exam was returned to complete a post-exam review assignment (Andaya et al. 2017). The first section ('correcting and reflecting') involved students correcting errors and explaining why these items were answered incorrectly. The second section ('examining study strategies and behaviours') required students to choose from a list of study tools and strategies (e.g. concept maps, flashcards) which they used and which they thought were most beneficial. The final section ('additional reflection') posed open-ended questions to give students the opportunity to reflect on how accurate they felt their exam grade was in representing their understanding and their revision efforts, how they could improve their learning before another exam, and what their teacher could do to help them learn. Andaya et al. (2017) demonstrated a significant positive correlation between the quality of students' explanations for their incorrect answers and grades on a subsequent exam, but quality of explanations did not predict an improvement in scores between the first and second exams. 82% of students reported that the exam review was beneficial for their learning, helping them to correct misunderstandings, evaluate their exam strategies and study techniques, and to support their preparation for future exams. An LMS-embedded exam wrapper activity was reported by Carpenter, Beall, and Hodges (2020); however, uptake of the activity was low, and it was students with higher levels of prior attainment that chose to take part in the activity.

A variation on the Exam Wrapper is the Exam Autopsy (Havis 2019) which targets metacognition through student-generated analysis of the causes of exam performance and action planning for future improvement. An Exam Autopsy involves students using class time to see correct and incorrect answers to all exam questions, and self-evaluating their performance. Students reflect upon the discrepancy between their expected and achieved grade, preparation strategies, and goal-setting for improvement. Following peer feedback, teachers provided feedback on self-evaluations, which was discussed with each student during a short face-to-face meeting. The final stage involved students identifying what they will do differently in preparation for the next test. Compared to no feedback and a typical exam wrapper activity, students experiencing the exam autopsy demonstrated significant improvements in performance on the subsequent exam. However, teachers expressed concerns about the amount of time that the process takes.

The use of class time for analysis of exam performance is also reflected in the Two-Stage Exam technique. After completing an exam individually, students then work on the same questions in small groups. The group discussion generates personalised feedback for students on the answers they gave in the individual portion of the exam, and enables students to gain extra credit from the group completion of the exam. Students are generally very positive about the learning experience afforded by two-stage examinations where, for example, 87% of Canadian Physics students recommended using the two-stage examination format, reporting benefits including the opportunity to learn from mistakes and understand peers' reasoning (Wieman, Rieger, and Heiner 2014). There is also evidence that the two-stage examination approach leads to gains in subsequent academic performance (Khong and Tanner 2021, p.76). However, these studies also demonstrate that two-stage examinations do not reduce students' test anxiety (Khong and Tanner 2021), and that some students report an increase in anxiety when the group discussion identifies that they had made many mistakes in the individual part of the exam (Wieman, Rieger, and Heiner 2014).

The two-stage examination approach also highlights the importance of inner feedback arising from comparison processes during group discussions. Nicol and Selvairetnam (2022) demonstrated that this collaborative reflection allowed students to generate solutions to weaknesses identified during the exam. They highlight two external sources of comparison at the group discussion stage: the collaborative written answer, and the discussion between peers about how to generate that answer. In a study involving a two-stage examination with UK Social Sciences

students, Nicol and Selvairetnam (2022) demonstrated that by reflecting upon the group discussion process, students generated inner feedback on content, the writing process, and self-regulatory feedback. Having identified weaknesses in their work at the individual stage, working with peers enabled students to generate solutions to these weaknesses. The analysis also demonstrated that student-generated feedback was more elaborate than teacher-provided comments on performance.

Despite the wealth of research evidence on individual provided examination feedback, this approach was only present in 19% of feedback policies we reviewed. These policies argued that examinations should not be treated differently to any other form of assessment when it comes to feedback:

Students should be provided with appropriate, accessible and timely feedback on all forms of assessed work, including all forms of examination, in ways that promote learning and facilitate improvement. (Policy, DI)

Feedback should be given on all examinations; All exam scripts should be returned to students in line with the existing procedures for other types of assessed coursework; Feedback on examinations should have the same turnaround timescales and conditions as for other assessed work. (Policy, TI)

The importance of treating examinations in a similar way to all other types of assessment was also evident in the survey data:

I do not think there should be any assessment tasks that do not provide feedback. (Survey, P60, History, RI)

I cannot see why some forms of assessment consider feedback essential whilst for others it would be considered ineffective. (Survey, P9, Built Environment, DI)

Approaches to individual provided feedback on examinations identified within policies on examinations included annotations on the exam script itself, and written comments on a pro-forma or cover sheet. Some policies indicated that the provision of individual feedback on exams should be reserved for those students who have failed or performed poorly in the examination:

It is good practice to offer feedback on summative examination work to students attaining consistently low marks in examinations and those who have performed uncharacteristically badly in one or more examination. (Policy, RI)

Individual provided feedback on exam performance was clearly on the forefront of respondents' minds in the survey, as when identifying the challenges of exam feedback, or describing why they do not provide exam feedback, the workload associated with writing individual feedback within time constraints for all students was cited as the most common concern:

We just don't have the staffing/time/resources to give individualised tailored feedback on all exams for every student (Survey, P79, Pharmacy, TI)

There is a limited period in which to turn around exam marks, so it is almost impossible to write feedback given the number of students who sit exams (Survey, P10, Computer Science, DI)

One survey participant indicated that the rationale for using examinations as a form of assessment was based on the pragmatic purpose of saving time rather than for pedagogic reasons, meaning that the workload associated with providing individual feedback on exams 'defeats the purpose of using an exam in the first place' (Survey, P78, Philosophy, RI). Another challenge raised by participants was a concern that the time spent writing individual feedback would be wasted because students may not engage with the information:

Most of them are just interested in the mark, rather than how they can improve going into the next round of examinations. (Survey, P72, Biosciences, RI)

I'm sure [exam feedback] can be very well used but suspect that not much of it is read or properly used. (Survey, P42, Classics, DI)

Within the literature, concerns about student engagement were not relevant where examination feedback activities took place in class time (e.g. Two-Stage Examinations, Exam Wrappers, Exam Autopsy). With other approaches, studies did suggest that more able students might be more likely than weaker students to engage with examination feedback (Yin, Sims, and Cothran 2012; Fyfe et al. 2014; Carpenter, Beall, and Hodges 2020). There may also be disciplinary differences; when provided with elaborated feedback on test performance, 61% of Commerce, 56% of Law, and 70% of Health students from The Netherlands opened the file containing the feedback information (Timmers and Veldkamp 2011).

### ***Synthesis of findings: individual requested (IR) feedback***

Just 4% of articles included in our literature review referenced IR approaches to examination feedback. This is in stark contrast to representation of this form of feedback in policy and practice, where 41% and 39% of policies and survey responses, respectively, identified this approach as one that is implemented. This disjuncture is an important finding which we return to explore in depth in our discussion. Articles coded under this category deployed IR feedback as a supplement to either IP (Fyfe et al. 2014; Havis 2019) or GP feedback (Drouin 2010). This suggests that IR feedback has not been explored in depth, with no relevant evidence pertaining to its impact on student performance, confidence, or experience.

In policies, IR feedback was often reported as supplementary to other forms of feedback, but was referenced in isolation in 18 policies. Similarly, in the survey data, a commonly-reported practice was the combination of IR feedback with GP, for example:

Individual feedback is not provided routinely. Instead, class distributions are shared, and generic feedback is provided when possible. Should a student wish to go through their exam individually they are welcome to do so at a scheduled drop-in session, or by making an appointment with the module organiser. (Survey, P64, Psychology, DI)

This approach, where students can gain individualised feedback on demand by viewing their script or meeting with the assessor, was commonplace in policies. Where students were able to view their script, most policies had a strict requirement for students to do so under controlled conditions, where students could not take the paper away. For example:

Students who wish to view their exam scripts following the provision of general feedback should be facilitated to do so under supervision of staff. (Policy, DI)

For the purpose of providing feedback on examination performance, candidates may be given access to examination scripts in the presence of a member of academic staff. Candidates are not permitted to retain examination scripts. (Policy, DI)

As well as viewing their script, other policies extended the feedback process to include individual face-to-face meetings with assessors through which to better understand their performance and how to improve. For example:

Candidates cannot have their scripts or copies of them returned. Candidates may, however, receive feedback on their performance if they wish and this will be in the form of a consultation with a tutor where the script will be viewed and the decisions of the Examiners explained. (Policy, RI)

Students may be offered the opportunity to discuss their examination scripts and receive personal feedback (Policy, TI)

Whilst some policies inferred that data protection legislation precluded the return of examination scripts, others did not view this as a barrier: 'Students can ask for their examination scripts to be returned to them. In these cases, the original examination script should be retained by the faculty and a scanned copy provided to the student' (Policy, TI). In other policies, the opportunity

to request individual feedback on examination performance was only available to certain students, typically those who had failed or performed poorly:

Students may be given access to their examination scripts e.g. in cases of substantial concern about individual performance, a tutor may give detailed feedback to the individual student which may include reviewing the examination script. (Policy, RI)

Where a student has not done well in an exam, there may be an opportunity to get face-to-face individualised comments from the lecturer or tutor. (Policy, RI)

Survey respondents identified barriers to IR exam feedback, most notably, logistical challenges in locating scripts for students to view:

Need to retrieve the paper when the student books in to see you for feedback. It may no longer be stored locally. (Survey, P14, Sport, Exercise and Health, DI)

If the exam papers are marked in hard copy, then there are logistical difficulties related to storing them and returning them to students (not to mention making copies if they are required for record keeping purposes). (Survey, P78, Philosophy, RI)

A further perceived barrier was identified where survey respondents were sceptical about student engagement with IR feedback: 'It is disheartening to know that 90% of the students won't come and see the feedback, so I am basically writing this for the External examiner' (Survey, P36, Biosciences, RI). However, if students do not take opportunities to engage with IR feedback, this may not necessarily reflect poor engagement. In a focus group study with humanities undergraduates in three UK Universities, students expressed uncertainty about how to contact tutors for advice, and felt uncomfortable and anxious about trying to catch lecturers before or after class. Office hours were also seen as challenging to make use of (Blair et al. 2014). Coupled with evidence that some groups of students, for example ethnic minority students, may feel less confident to engage in feedback seeking (e.g. Campbell, Hawkins, and Osman 2021; Morrison, Machado, and Blackburn 2024), IR feedback may lead to exclusionary rather than inclusive practices.

### ***Synthesis of findings: generic provided (GP) feedback***

A total of 16 articles referenced Generic Provided (GP) feedback, with most ( $N=9$ ) exploring GP feedback as a standalone practice. Five articles examined GP feedback alongside IP feedback. This combination contrasts with data extracted from policies and survey responses, where only 19% of 47 policies that referenced GP feedback treated it as a stand-alone practice, while 79% described it as supplemental to IR feedback. Similarly, many survey respondents reported combining GP with IR feedback in their practices.

Several specific approaches to GP feedback were discussed in the literature, with limited evidence of impact on student performance. Approaches included posting correct answers to an examination on a notice board, followed by discussion of common errors (Carrillo-de-la-Peña et al. 2009), and return of a spreadsheet showing students their scores for each question against anonymised scores for all students (Ellis and Barber 2016). Students report positive perceptions of the utility of seeing where they might have gained or lost marks (Ellis and Barber 2016). Similarly, 91% of anatomy and physiology students strongly agreed or agreed that a scheduled 'feedback lecture', in which the distribution of students' responses to the ten most difficult questions in an MCQ exam was discussed, was useful (Weston-Green and Wallace 2016).

A similar method, Formative Summative Assessment (FSA), was proposed by Wininger (2005), allowing students to review their performance and discuss their answers in class. Students received their grades before the class, where the teacher went through the paper and facilitated discussions about the correct and incorrect answers. Wininger's study with US Psychology



students found that students using FSA significantly outperformed those who merely received correct answers. Drouin (2010) implemented an FSA approach with individual and group exam review activities and observed that group-based review led to better gains in performance compared to individual review.

The 'immediate feedback, no return' process (Friedman 1987) involves lecturers discussing correct answers and common misconceptions immediately after students have handed in MCQ papers. This process was implemented by Smith and Wight (1988) in a study with Psychology students in an American University. Students rated the technique as 9.69 on a scale from 1 (poor) to 10 (great), where students commented that they felt this technique greatly aided their learning.

Despite evidence of positive impact of GP exam feedback approaches, the literature also raises some cautions. Ellery (2008) questions whether the generic rather than personalised nature of comments might limit student engagement:

Providing 'global' feedback to a whole group can serve as a double-edged sword. On the one hand students tend to take the comments less personally and are therefore less defensive about them, allowing them to engage more positively with the feedback. On the other hand...many believe the comments do not apply specifically to themselves and tend to ignore them. (Ellery 2008, 427)

The use of GP exam feedback was strongly represented in feedback policies, often framed as a 'minimum standard' to be enacted by programme teams. For example:

As a minimum requirement all Schools will provide a Generic Feedback Report on examination performance for each examination paper. (Policy, DI)

All examinations will be followed by feedback to students. As a minimum, this will be in the form of a general presentation to students indicating common strengths and weaknesses exhibited in papers, advising upon how performance could be improved. (Policy, TI)

The most common approach to GP feedback represented in policies was the provision of a cohort exam report or summary of overall performance, common errors, and guidance for improvement:

Feedback on exams may well take the form of 'whole-class comments'...which are presented verbally in a class or emailed. (Policy, RI)

The whole cohort may be sent generic feedback which identifies common mistakes, gives examples of excellent work and encourages students to develop particular skills for future assessments. (Policy, DI)

In contrast to the more active and dialogic forms of GP feedback discussed in the literature, these examples from policies represent a more transactional approach to feedback, through the mere provision of a report that is 'delivered' to students. Weston-Green and Wallace (2016, 394) caution against this approach, suggesting that the mere 'delivery' of feedback 'may not engage the student or facilitate peer and student-teacher dialogues, nor does it seek to understand and correct common misconceptions'. Just one more active and dialogic example of GP exam feedback was identified in the policies, where it was recommended that exam performance could be reviewed in a related module/unit at the start of the next academic year, providing an opportunity for students 'to go over the exam questions and produce model answers for discussion in class' (Policy, TI).

### ***Synthesis of findings: generic requested (GR) feedback***

Generic feedback made available on request was the least represented category across literature, policy, and practice. Institutional policies mentioned opportunities to gain cohort level feedback information through optional sessions, for example obtaining feedback information 'by attending an event at which generic feedback is provided' (Policy, TI), or through 'optional seminars to discuss recent exams' (Policy, DI). Only one policy referred to the opportunity for students to request

written generic feedback information: 'Generic feedback on written examinations will be available to students if requested, via a face-to-face meeting with an appropriate member of the Module team' (Policy, TI).

Survey responses indicated that, in some cases, only those students who had failed an examination could request access to generic feedback information by attending an optional session: 'Students who have failed and have the option of a re-sit are invited by email to attend a workshop which debriefs on omissions and errors of commission in the past exam' (Survey P52, Business & Management, DI).

Only one example within the literature represented generic requested feedback. Through a practice termed 'supplemental instruction', students were able to view answers to a test on the VLE/LMS, and then attend a voluntary session to work through test answers with teachers (Erasmus 2017). Students reported that this helped them develop their test-taking skills, and supplemental instruction also led to greater gains in subsequent test performance (Erasmus 2017).

## Discussion

The primary aim of the present study was to build a taxonomy of approaches to examination feedback, whilst also exploring barriers to examination feedback and examining the strength of evidence for the impact of examination feedback. We now turn to a discussion of our findings against each of these aims in turn.

### *The taxonomy of examination feedback*

Our synthetic analysis across literature, policies, and reports of practice led us to identify 24 different examination feedback practices. We reduced these individual practices into a framework with two dimensions (individual-generic; provided-requested). This resulted in a four-quadrant taxonomy representing four broad approaches to feedback on examinations: Individual Provided (IP); Individual Requested (IR); Generic Provided (GP); Generic Requested (GR).

One of the most striking findings from our analysis was the uneven prevalence of these four approaches across each of our data sources. In our survey data, Generic Provided and Individual Requested feedback were the most common practices. These two approaches were often used in conjunction; for example, students would receive a cohort performance report and could request a meeting with the assessor to further discuss their performance. Both survey data and policy analysis highlighted the rarity of Individual Provided and Generic Requested feedback.

Research evidence, as per our systematic review, had a different emphasis. Most articles focused on Individual Provided feedback, including specific approaches such as IFAT (e.g. Brosvic et al. 2005; DiBattista et al. 2009) and exam wrappers. (e.g. Clark 2019). The misalignment between the emphasis within policy and practice on the one hand, and research on the other, is problematic for several reasons. First, the approach to examination feedback with the most research evidence to support its efficacy (IP), is the least prevalent in policy and practice. This means that the literature is unlikely to be of relevance to many educators, and does not address the constraints or challenges experienced by educators in practice (Buckley 2024). Second, there is very little evidence to support the efficacy of the approaches that are most common in policy and practice (GP and IR). Consequently, it remains uncertain whether these prevalent forms of feedback benefit student learning or potentially exacerbate attainment gaps (based, for example, on evidence that ethnic minority students are less likely to actively seek feedback, e.g. Campbell, Hawkins, and Osman 2021; Morrison, Machado, and Blackburn 2024). Furthermore, our analysis of policies identified examples of exclusionary rather than inclusive approaches to examination feedback, only offering these opportunities to certain groups of students (e.g. those who failed



or performed poorly). Whilst it could be argued that these students might stand to benefit the most from examination feedback, an inclusive approach to education requires that these opportunities are open to all.

### ***Barriers to examination feedback***

Despite previous reports that the provision of feedback information on examination performance is rare (e.g. Brown and Glasner 2003; Weaver 2006; Blair et al. 2014), in our survey, 75% of respondents reported utilising some form of examination feedback practice. This may indicate that exam feedback has become more common since these papers were published, or could indicate that those people responding to the survey were somehow more likely to use examination feedback. We identified three perceived barriers to the use of examination feedback: workload, deviation from the 'status quo', and university policies that do not mandate it as common reasons.

Providing detailed and individualised feedback information on student performance on any form of assessment is time-consuming (e.g. Black, DeGrassi, and Sweet 2021). Examinations often have to be graded in time to meet deadlines for examination boards and the release of final results. However, workload challenges are particularly problematic for IP feedback, and are probably one of the reasons why IP is rare in practice, with GP and IR approaches aiming to provide a manageable solution. Below we offer some tentative suggestions as to how workload challenges might be overcome.

The challenges of status quo and policy mandates are more complex. We suspect that there might be conflation between the two, where the status quo position of 'we don't do examination feedback' is maintained because of a belief that this is not the policy of the University, without this being the case in reality. Other authors have highlighted the prevalence of misconceptions around what is 'allowed' in examination feedback practice. For example, Scoles, Huxham, and McArthur (2013, p.632) argued that 'Few students are given access to their exam scripts, either because of the real or mythologised barriers to this practice'.

### ***The impact of examination feedback***

In the IP domain, there is ample evidence that providing some form of feedback information on examination feedback is perceived positively by students (e.g. DiBattista, Mitterer, and Gosse 2004; Jordan 2004; DiBattista and Gosse 2006; Barker 2011; Bowman and Laurent 2011; Persky and Pollack 2008; Malau-Aduli et al. 2014; Wieman, Rieger, and Heiner 2014; Andaya et al. 2017; Lee and Cha 2022;). There is also clear evidence that IP examination feedback can also lead to self-reported (Arkin and Schumann 1984; DiBattista, Mitterer, and Gosse 2004) and measured (e.g. Lipnevich and Smith 2009; Yopp and Rehberger 2009; Ellis and Barber 2016; Rozell et al. 2017; Havis 2019; Khong and Tanner 2021) gains in subsequent performance. However, there is limited evidence that such feedback can reduce established attainment gaps in examination performance.

Beyond the IP domain, there was evidence of performance improvements for GP (e.g. Wininger 2005; Drouin 2010) and GR (Erasmus 2017) approaches. However, as previously discussed, the lack of evidence for the efficacy of IR examination feedback is concerning, given that this is a common approach in policy and practice, and that these approaches could stand to advantage some students over others.

Not all evidence of the impact of examination feedback is positive. For example, our literature review highlighted that immediate feedback on MCQ performance does not always lead to improved performance (Badyal et al. 2019) and can exacerbate students' experience of stress by making errors salient (Rostaminezhad, 2019). There are also examples of where students report

preferences for approaches such as IFAT that are not translated into performance improvements (e.g. Bowman and Laurent 2011; Shoulders, Johnson, and Wiedenmann 2017; Wood, Klausz, and MacNeil 2022).

### ***Limitations and future research directions***

Whilst we addressed our aims using a data corpus consisting of literature, policy documents, and a practitioner survey, the latter two sources of data were only representative of UK universities and practitioners. Exploring approaches to examination feedback in a wider range of international settings would add to our understanding of its challenges. Our analysis of the impact of examination feedback relied solely on a qualitative synthesis of findings. A quantitative meta-analysis would enable us to know how meaningful the impact of examination feedback is in standardised terms, and how various factors related to the design of the feedback approaches might mediate the size of the effect on performance. Another key direction for future research is to engage in more focused study of 'signature' feedback processes (e.g. Quinlan and Pitt 2021; Carless et al. 2023) in some disciplines in the context of specific types of examinations, for example the Objective Structured Clinical Examination (OSCE) in health professions education.

Our synthetic analysis across policy, practice, and research evidence suggests the need for further research on the impact of commonly-used examination feedback approaches (Generic Provided and Individual Requested), as well as further research on the impact of Individual Provided examination feedback on differential attainment across student groups.

### ***Implications for policy, practice, and research***

Our analysis has identified that IR feedback is very common in policy and practice, yet lacks support from empirical evidence. This suggests that it is imperative to know more about the nature of students who choose to take up the opportunity to request further feedback, and the impact on future attainment, to be able to understand whether this common feedback practice might be inadvertently exacerbating established attainment gaps. Similarly, the literature indicates that when optional examination feedback activities are offered, more able students (Yin, Sims, and Cothran 2012; Fyfe et al. 2014; Carpenter, Beall, and Hodges 2020), older students, students with more experience, and female students (Fyfe et al. 2014) are more likely to engage. This indicates that feedback opportunities after examinations might be most effective if embedded into scheduled class time so that all students are able to benefit, without advantaging some over others.

Our survey data indicate that a common barrier to the use of examination feedback is concern about workload. Yet our literature review offers fruitful suggestions for approaches which are effective in supporting learning without adding further burden to educators. One such approach is the use of automated feedback information such as computer-generated comments (Barker 2011; Ellis and Barber 2016). There is also a key role for approaches that give students agency in reflecting upon or generating information about their performance, such as using a post-test reflection survey (Fyfe et al. 2014), Exam Wrapper (Clark 2019), post exam review assignments (Andaya et al. 2017), or the exam autopsy (Havis 2019). Approaches to examination feedback can also be embedded within the assessment process, through the opportunity to gain credit for engaging with and applying information about performance. This approach is seen in two-stage exams (Wieman, Rieger, and Heiner 2014; Nicol and Selvaetnam 2022) and designs where students can regrade incorrect answers for partial credit (Rozell et al. 2017).

We advocate a reframing of thinking about examination feedback away from how to 'provide' feedback information, which places all of the burden on educators and is challenging in terms

of timing and workload (e.g. Blair et al. 2014), towards the design of 'feedback spaces' (see Wood 2021) where students take responsibility for generating feedback on their own performance through meaningful class and assessment activities. This approach offers one potential solution to the conundrum of establishing a 'norm' of feedback on examinations (Blair et al. 2014) whilst minimising the impact on educators' workloads, whilst also aligning with recent shifts towards a 'new paradigm' of feedback processes which emphasise the role of the student in taking responsibility for engaging with and enacting feedback information (Winstone and Carless 2019). Also important is the development of policies and practices challenging the status quo and addressing workload challenges in providing impactful feedback information on student examination performance.

Finally, we believe that the examination feedback taxonomy is of theoretical and practical significance. In policy and practice, the taxonomy could be used to facilitate an audit of current approaches to examination feedback that are used in an institution and could also enable individual educators to align their own approaches with research evidence. For researchers, the taxonomy could be used to align approaches to categorising examination feedback approaches which would support future research syntheses.

## Acknowledgements

The authors thank Lauren McGuiness and Sarah Hack for their assistance with data extraction and coding.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Funding

The research reported in this article was supported by AdvanceHE through the award of a National Teaching Fellowship to the first author, and by an internal Teaching and Learning award to the second author.

## Notes on contributors

**Professor Naomi E. Winstone** is Professor of Educational Psychology and Director of the Surrey Institute of Education at the University of Surrey. Naomi is also an Honorary Professor in the Centre for Research in Assessment and Digital Learning (CRADLE) at Deakin University, Australia. She is a Principal Fellow of the Higher Education Academy and a UK National Teaching Fellow. Naomi is a cognitive psychologist specialising in the processing and impact of instructional feedback, and the influence of dominant discourses of assessment and feedback in policy and practice on the positioning of educators and students in feedback processes. Her 2019 book co-authored with David Carless, *Designing Effective Feedback Processes in Higher Education: A Learning-focused approach* set the agenda for new approaches to feedback policy and practice. Using her research and institutional leadership experience, Naomi works with schools, universities and organisations across the world to embed learning-focused approaches to feedback into practice.

**Dr Edd Pitt** is a Reader in Higher Education and Academic Practice in the Centre for the Study of Higher Education at the University of Kent, UK. Edd is also an Honorary Associate Professor in the Centre for Research in Assessment and Digital Learning (CRADLE) at Deakin University, Australia. He is a Principal Fellow of the Higher Education Academy. His principal research field is Assessment and Feedback with a particular focus upon the development of staff and student feedback literacy.

## ORCID

Naomi E. Winstone  <http://orcid.org/0000-0001-8157-8274>

Edd Pitt  <http://orcid.org/0000-0002-7475-0299>

## References

- \*Paper included in our systematic review\* Andaya, G., V. D. Hrabak, S. T. Reyes, R. E. Diaz, and K. K. McDonald. 2017. "Examining the Effectiveness of a Postexam Review Activity to Promote Self-Regulation in Introductory Biology Students." *Journal of College Science Teaching* 46 (4): 84–92. doi:10.2505/4/jcst17\_046\_04\_84.
- \*Arkin, R. M., and D. W. Schumann. 1984. "Effect of Corrective Testing: An Extension." *Journal of Educational Psychology* 76 (5): 835–843. doi:10.1037/0022-0663.76.5.835.
- \*Badyal, D., S. Bala, T. Singh, and G. Gulrez. 2019. "Impact of Immediate Feedback on the Learning of Medical Students in Pharmacology." *Journal of Advances in Medical Education & Professionalism* 7 (1): 1–6. doi:10.30476/JAMP.2019.41036.
- \*Barker, T. 2011. "An Automated Individual Feedback and Marking System: An Empirical Study." *Electronic Journal of E-Learning* 9 (1): 1–14.
- \*Barnard, R. A., and R. D. Sweeder. 2020. "Using Online Grading to Stagger Midterm Exam Feedback and Create Space for Meaningful Student Reflection." *College Teaching* 68 (2): 60–61. doi:10.1080/87567555.2020.1713041.
- Becher, T., and P. R. Trowler. 2001. *Academic Tribes and Territories* (2nd ed.). Buckingham: Society for Research into Higher Education & Open University Press.
- Biggs, J. 2001. "The Reflective Institution: Assuring and Enhancing the Quality of Teaching and Learning." *Higher Education* 41 (3): 221–238. [Mismatch ] doi:10.1023/A:1004181331049.
- Black, S. L., S. W. DeGrassi, and K. M. Sweet. 2021. "Multisource Feedback as an Experiential Learning Enabler in Large-Format Management Classes." *Journal of Management Education* 45 (3): 479–517. doi:10.1177/1052562920987292.
- \*Blair, A., A. Wyburn-Powell, M. Goodwin, and S. Shields. 2014. "Can Dialogue Help to Improve Feedback on Examinations?" *Studies in Higher Education* 39 (6): 1039–1054. doi:10.1080/03075079.2013.777404.
- \*Bowman, T. G., and T. Laurent. 2011. "Immediate Feedback and Learning in Athletic Training Education." *Athletic Training Education Journal* 6 (4): 202–207. doi:10.4085/1947-380X-6.4.202.
- Braun, V., and V. Clarke. 2021. "Can I Use TA? Should I Use TA? Should I Not Use TA? Comparing Reflexive Thematic Analysis and Other Pattern-Based Qualitative Analytic Approaches." *Counselling and Psychotherapy Research* 21 (1): 37–47. doi:10.1002/capr.12360.
- Bridges, Paul, Angela Cooper, Peter Evanson, Chris Haines, Don Jenkins, David Scurry, Harvey Woolf, and Mantz Yorke. 2002. "Coursework Marks High, Examination Marks Low: Discuss." *Assessment & Evaluation in Higher Education* 27 (1): 35–48. doi:10.1080/02602930120105045.
- \*Brosvic, G. M., and M. L. Epstein. 2007. "Enhancing Learning in the Introductory Course." *The Psychological Record* 57 (3): 391–408. doi:10.1007/BF03395584.
- \*Brosvic, G. M., M. L. Epstein, M. J. Cook, and R. E. Dihoff. 2005. "Efficacy of Error for the Correction of Initially Incorrect Assumptions and of Feedback for the Affirmation of Correct Responding: Learning in the Classroom." *The Psychological Record* 55 (3): 401–418. doi:10.1007/BF03395518.
- Brown, G., and A. Glasner. 2003. *Assessment Matters in Higher Education: Choosing and Using Diverse Approaches*. Buckingham: Open University Press.
- Buckley, A. 2024. "Are we Answering the Question That Has Been Set? Exploring the Gap between Research and Practice around Examinations in Higher Education." *Studies in Higher Education* 49 (11): 1928–1944. doi:10.1080/03075079.2023.2283784.
- \*Butler, A. C., J. D. Karpicke, and H. L. Roediger. III. 2008. "Correcting a Metacognitive Error: Feedback Increases Retention of Low-Confidence Correct Responses." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 34 (4): 918–928. doi:10.1037/0278-7393.34.4.918.
- Butler, A. C., and H. L. Roediger. 2008. "Feedback Enhances the Positive Effects and Reduces the Negative Effects of Multiple-Choice Testing." *Memory & Cognition* 36 (3): 604–616. doi:10.3758/mc.36.3.604.
- Campbell, D., C. D. Hawkins, and S. Osman. 2021. "Tackling Racial Inequalities in Assessment in Higher Education: A Multi-Disciplinary Case Study." University of Leicester. <https://le.ac.uk/-/media/uol/docs/news/tackling-racial-inequalities-in-assessment-in-he-may-21.pdf>
- Carless, D., J. To, C. Kwan, and J. Kwok. 2023. "Disciplinary Perspectives on Feedback Processes: Towards Signature Feedback Practices." *Teaching in Higher Education* 28 (6): 1158–1172. doi:10.1080/13562517.2020.1863355.
- \*Carpenter, T. S., L. C. Beall, and L. C. Hodges. 2020. "Using a Learning Management System for Exam Wrapper Feedback to Prompt Metacognitive Awareness in Large Courses." *Journal of Teaching and Learning with Technology* 9 (1): 79–91. doi:10.14434/jotlt.v9i1.29156.
- \*Carrillo-de-la-Peña, M. T., E. Bailles, X. Caseras, À. Martínez, G. Ortet-Fabregat, and J. Pérez. 2009. "Formative Assessment and Academic Achievement in Pre-Graduate Students of Health Sciences." *Advances in Health Sciences Education: Theory and Practice* 14 (1): 61–67. doi:10.1007/s10459-007-9086-y.
- \*Clark, K. R. 2019. "A Postexam Review Assignment." *Radiologic Technology* 91 (2): 184–186.
- Cramer, L. 2021. "Alternative Strategies for Closing the Award Gap between White and Minority Ethnic Students." *eLife* 10: E 58971. doi:10.7554/eLife.58971.
- \*DiBattista, D., and L. Gosse. 2006. "Test Anxiety and the Immediate Feedback Assessment Technique." *The Journal of Experimental Education* 74 (4): 311–328. doi:10.3200/JEXE.74.4.311-328.

- \*DiBattista, D., L. Gosse, J. A. Sinnige-Egger, B. Candale, and K. Sargeson. 2009. "Grading Scheme, Test Difficulty, and the Immediate Feedback Assessment Technique." *The Journal of Experimental Education* 77 (4): 311–338. doi:10.3200/JEXE.77.4.311-338.
- \*DiBattista, D., J. O. Mitterer, and L. Gosse. 2004. "Acceptance by Undergraduates of the Immediate Feedback Assessment Technique for Multiple-Choice Testing." *Teaching in Higher Education* 9 (1): 17–28. doi:10.1080/1356251032000155803.
- \*Douglas, M., J. Wilson, and S. Ennis. 2012. "Multiple-Choice Question Tests: A Convenient, Flexible and Effective Learning Tool? A Case Study." *Innovations in Education and Teaching International* 49 (2): 111–121. doi:10.1080/14703297.2012.677596.
- \*Drouin, M. A. 2010. "Group-Based Formative Summative Assessment Relates to Improved Student Performance and Satisfaction." *Teaching of Psychology* 37 (2): 114–118. doi:10.1080/00986281003626706.
- \*Ellery, K. 2008. "Assessment for Learning: A Case Study Using Feedback Effectively in an Essay-Style Test." *Assessment & Evaluation in Higher Education* 33 (4): 421–429. doi:10.1080/02602930701562981.
- \*Ellis, S., and J. Barber. 2016. "Expanding and Personalising Feedback in Online Assessment: A Case Study in a School of Pharmacy." *Practitioner Research in Higher Education* 10 (1): 121–129.
- \*Erasmus, M. 2017. "From Inky Pinky Ponky to Improving Student Understanding in Assessment: Exploring the Value of Supplemental Instruction in a Large First-Year Class." *Journal of Student Affairs in Africa* 5 (2): 33–53. doi:10.24085/jsaa.v5i2.2701.
- Farmer, L. 2017. *Success Rates and Examiner Bias in the Testing of International Medical Graduates on High-Stakes Postgraduate Clinical Examinations*. Australasian College for Emergency Medicine. [https://acem.org.au/getmedia/b87ed5ff-7210-450f-b676-9649bd86d6de/EFarmer\\_Report.aspx](https://acem.org.au/getmedia/b87ed5ff-7210-450f-b676-9649bd86d6de/EFarmer_Report.aspx)
- French, S., A. Dickerson, and R. A. Mulder. 2023. "A Review of the Benefits and Drawbacks of High-Stakes Final Examinations in Higher Education." *Higher Education*. Advance Online Publication.
- Friedman, H. 1987. "Immediate Feedback, no Return Test Procedure for Introductory Courses." *Teaching of Psychology* 14 (4): 244–244. doi:10.1207/s15328023top1404\_19.
- \*Fyfe, G., S. Fyfe, J. Meyer, M. Ziman, K. Sanders, and J. Hill. 2014. "Students Reflecting on Test Performance and Feedback: An on-Line Approach." *Assessment & Evaluation in Higher Education* 39 (2): 179–194. doi:10.1080/02602938.2013.801063.
- Gibbs, G. 1992. "Improving the Quality of Student Learning through Course Design." In *Learning to Effect*, edited by R. Barnett, 149–165. Milton Keynes, UK: SRHE & Open University Press.
- Gibbs, G., and L. Lucas. 1997. "Coursework Assessment, Class Size and Student Performance: 1984–94." *Journal of Further and Higher Education* 21 (2): 183–192. doi:10.1080/0309877970210204.
- \*Havis, L. R. 2019. "The Exam Autopsy: An Integrated Post-Exam Assessment Model." *International Journal for the Scholarship of Teaching and Learning* 13 (1): 1–8.
- \*Jordan, J. 2004. "The Use of Orally Recorded Exam Feedback as a Supplement to Written Comments." *Journal of Statistics Education* 12 (1): 1–6. doi:10.1080/10691898.2004.11910716.
- \*Khong, M. L., and J. A. Tanner. 2021. "A Collaborative Two-Stage Examination in Biomedical Sciences: Positive Impact on Feedback and Peer Collaboration." *Biochemistry and Molecular Biology Education: A Bimonthly Publication of the International Union of Biochemistry and Molecular Biology* 49 (1): 69–79. doi:10.1002/bmb.21392.
- \*Krippendorff, B. B., D. L. Bolender, and G. L. Kolesari. 2008. "Computerized Grading of Anatomy Laboratory Practical Examinations." *Anatomical Sciences Education* 1 (5): 220–223. doi:10.1002/ase.46.
- LaDonna, K. A., T. Taylor, and L. Lingard. 2018. "Why Open-Ended Survey Questions Are Unlikely to Support Rigorous Qualitative Insights." *Academic Medicine: Journal of the Association of American Medical Colleges* 93 (3): 347–349. doi:10.1097/ACM.0000000000002088.
- \*Lee, H. W., and Y. M. Cha. 2022. "The Effects of Digital Written Feedback on Paper-Based Tests for College Students." *The Asia-Pacific Education Researcher* 31 (4): 489–497. doi:10.1007/s40299-021-00592-8.
- \*Lipnevich, A. A., and J. K. Smith. 2009. "Effects of Differential Feedback on Students' Examination Performance." *Journal of Experimental Psychology: Applied* 15 (4): 319–333. doi:10.1037/a0017841.
- \*Malau-Aduli, B. S., D. Assenheimer, D. Choi-Lundberg, and C. Zimitat. 2014. "Using Computer-Based Technology to Improve Feedback to Staff and Students on MCQ Assessments." *Innovations in Education and Teaching International* 51 (5): 510–522. doi:10.1080/14703297.2013.796711.
- \*Marsh, E. J., J. P. Lozito, S. Umanath, E. L. Bjork, and R. A. Bjork. 2012. "Using Verification Feedback to Correct Errors Made on a Multiple-Choice Test." *Memory (Hove, England)* 20 (6): 645–653. doi:10.1080/09658211.2012.684882.
- Morrison, N., M. Machado, and C. Blackburn. 2024. "Bridging the Gap: Understanding the Barriers and Facilitators to Performance for Black, Asian and Minority Ethnic Medical Students in the United Kingdom." *Medical Education* 58 (4): 443–456. doi:10.1111/medu.15246.
- \*Nicol, D., and G. Selvaethnam. 2022. "Making Internal Feedback Explicit: Harnessing the Comparisons Students Make during Two-Stage Exams." *Assessment & Evaluation in Higher Education* 47 (4): 507–522. doi:10.1080/02602938.2021.1934653.
- \*Persky, A. M., and G. M. Pollack. 2008. "Using Answer-Until-Correct Examinations to Provide Immediate Feedback to Students in a Pharmacokinetics Course." *American Journal of Pharmaceutical Education* 72 (4): 83. doi:10.1016/S0002-9459(24)00440-6.



- Quinlan, K. M., and E. Pitt. 2021. "Towards Signature Assessment and Feedback Practices: A Taxonomy of Discipline-Specific Elements of Assessment for Learning." *Assessment in Education: Principles, Policy & Practice* 28 (2): 191–207. doi:[10.1080/0969594X.2021.1930447](https://doi.org/10.1080/0969594X.2021.1930447).
- Richardson, J. T. 2015. "Coursework versus Examinations in End-of-Module Assessment: A Literature Review." *Assessment & Evaluation in Higher Education* 40 (3): 439–455. doi:[10.1080/02602938.2014.919628](https://doi.org/10.1080/02602938.2014.919628).
- \*Rostaminezhad, M. A. 2019. "Students' Perceptions of the Strengths and Limitations of Electronic Tests Focusing on Instant Feedback." *Journal of Information Technology Education* 18: 59–71.
- Rowntree, D. 1987. *Assessing Students: How Shall we Know Them?* London: Routledge.
- \*Rozell, T. G., J. Johnson, A. Sexten, and A. E. Rhodes. 2017. "Correcting Missed Exam Questions as a Learning Tool in a Physiology Course." *Journal of College Science Teaching* 46 (5): 56–63. doi:[10.2505/4/jcst17\\_046\\_05\\_56](https://doi.org/10.2505/4/jcst17_046_05_56).
- Scoles, J., M. Huxham, and J. McArthur. 2013. "No Longer Exempt from Good Practice: Using Exemplars to Close the Feedback Gap for Exams." *Assessment & Evaluation in Higher Education* 38 (6): 631–645. doi:[10.1080/02602938.2012.674485](https://doi.org/10.1080/02602938.2012.674485).
- Shaw, L. M., and M. R. Tranter. 2021. "Levelling the Playing Field in Assessment: An Analysis of Attainment Gaps for Widening Participation, Black and Minority Ethnic Mathematics Undergraduates before and after the COVID-19 Lockdown." *Teaching Mathematics and Its Applications: An International Journal of the IMA* 40 (4): 497–505. doi:[10.1093/teamat/hrab024](https://doi.org/10.1093/teamat/hrab024).
- \*Shoulders, C. W., D. M. Johnson, and R. N. Wiedenmann. 2017. "Evaluation of Two Quiz Response Formats: Effects on Student Perceptions and Achievement." *NACTA Journal* 61 (2): 157–161.
- Simonite, V. 2003. "A Longitudinal Study of Achievement in a Modular First-Degree Course." *Studies in Higher Education* 28 (3): 293–302. doi:[10.1080/03075070309291](https://doi.org/10.1080/03075070309291).
- Sims, A. 2023. "ChatGPT and the Future of University Assessment." *Times Higher Education*, 9<sup>th</sup> March 2023.
- \*Slepkov, A. D. 2013. "Integrated Testlets and the Immediate Feedback Assessment Technique." *American Journal of Physics* 81 (10): 782–791. doi:[10.1119/1.4820241](https://doi.org/10.1119/1.4820241).
- \*Smith, R. A., and R. Wight. 1988. "Student Evaluation of Friedman's Immediate Feedback, no Return Test Procedure for Introductory Psychology." *Teaching of Psychology* 15 (4): 209–210. doi:[10.1207/s15328023top1504\\_12](https://doi.org/10.1207/s15328023top1504_12).
- Stegers-Jager, K. M., F. N. Brommet, and A. P. N. Themmen. 2016. "Ethnic and Social Disparities in Different Types of Examinations in Undergraduate Pre-Clinical Training." *Advances in Health Sciences Education: Theory and Practice* 21 (5): 1023–1046. doi:[10.1007/s10459-016-9676-7](https://doi.org/10.1007/s10459-016-9676-7).
- Tomas, C., and T. Jessop. 2019. "Struggling and Juggling: A Comparison of Student Assessment Loads across Research and Teaching-Intensive Universities." *Assessment & Evaluation in Higher Education* 44 (1): 1–10. doi:[10.1080/02602938.2018.1463355](https://doi.org/10.1080/02602938.2018.1463355).
- \*Timmers, C., and B. Veldkamp. 2011. "Attention Paid to Feedback Provided by a Computer-Based Assessment for Learning on Information Literacy." *Computers & Education* 56 (3): 923–930. doi:[10.1016/j.compedu.2010.11.007](https://doi.org/10.1016/j.compedu.2010.11.007).
- Weaver, M. R. 2006. "Do Students Value Feedback? Student Perceptions of Tutors' Written Responses." *Assessment & Evaluation in Higher Education* 31 (3): 379–394. doi:[10.1080/02602930500353061](https://doi.org/10.1080/02602930500353061).
- \*Weston-Green, K., and M. Wallace. 2016. "A Method of Providing Engaging Formative Feedback to Large Cohort first-Year Physiology and Anatomy Students." *Advances in Physiology Education* 40 (3): 393–397. doi:[10.1152/advan.00174.2015](https://doi.org/10.1152/advan.00174.2015).
- \*Wieman, C. E., G. W. Rieger, and C. E. Heiner. 2014. "Physics Exams That Promote Collaborative Learning." *The Physics Teacher* 52 (1): 51–53. doi:[10.1119/1.4849159](https://doi.org/10.1119/1.4849159).
- \*Wininger, S. R. 2005. "Using Your Tests to Teach: Formative Summative Assessment." *Teaching of Psychology* 32 (3): 164–166. doi:[10.1207/s15328023top3203\\_7](https://doi.org/10.1207/s15328023top3203_7).
- Winstone, N., and D. Carless. 2019. *Designing Effective Feedback Processes in Higher Education: A Learning-Focused Approach*. Abingdon: Routledge.
- Wood, J. 2021. "A Dialogic Technology-Mediated Model of Feedback Uptake and Literacy." *Assessment & Evaluation in Higher Education* 46 (8): 1173–1190. doi:[10.1080/02602938.2020.1852174](https://doi.org/10.1080/02602938.2020.1852174).
- \*Wood, E., N. Klausz, and S. MacNeil. 2022. "Examining the Influence of Multiple-Choice Test Formats on Student Performance." *Innovative Higher Education* 47 (3): 515–531. doi:[10.1007/s10755-021-09581-7](https://doi.org/10.1007/s10755-021-09581-7).
- \*Yin, M., J. Sims, and D. Cothran. 2012. "Scratching Where They Itch: Evaluation of Feedback on a Diagnostic English Grammar Test for Taiwanese University Students." *Language Assessment Quarterly* 9 (1): 78–104. doi:[10.1080/15434303.2010.516043](https://doi.org/10.1080/15434303.2010.516043).
- \*Yopp, D., and R. Rehberger. 2009. "A Curriculum Focus Intervention's Effects on Prealgebra Achievement." *Journal of Developmental Education* 33 (2): 28–38.

## Appendix A. Prevalence of the provision of examination feedback by institution type and discipline area

Institution Type					
Research Intensive		Teaching Intensive		Dual Intensive	
Provide feedback to students on examination performance?					
Yes	No	Yes	No	Yes	No
16 (84%)	3 (16%)	44 (75%)	15 (25%)	25 (76%)	8 (24%)

Discipline Area							
Hard Pure		Hard Applied		Soft Pure		Soft Applied	
Provide feedback to students on examination performance?							
Yes	No	Yes	No	Yes	No	Yes	No
10 (91%)	1 (9%)	21 (88%)	3 (12%)	15 (75%)	5 (25%)	41 (67%)	20 (33%)