CLASSIFYING AND EVALUATING ASSESSMENT FEEDBACK PRACTICES

Dominic Palmer-Brown\(^1\), Fang Fang Cai\(^2\), Preeti Patel\(^1\)

\(^1\) London Metropolitan University (UNITED KINGDOM)
\(^2\) Anglia Ruskin University (UNITED KINGDOM)

Abstract

The provision of assessment feedback to students is an area which has received much interest in modern education, particularly in the Higher Education context. As current pedagogic practices strongly encourage the provision of feedback and given also the advances in digital technology, feedback mechanisms are becoming ever more sophisticated. However, considering that a great deal of effort is expended on timely, actionable and constructive feedback by tutors, the student perception of the value of the feedback given to them is not as positive as it could be.

Currently a multitude of feedback practices have been developed and utilised, though with varying degrees of productiveness. Research in this area is understandably extremely broad as subject disciplines, use of technology, assessment types, methods and tools, educator preferences, student audience and peer and self-assessment capability all have a significant part to play. Given that the approaches to providing feedback are myriad, it is desirable to advance a systematic method of understanding the most constructive feedback types. This paper describes the development of a taxonomical classification which provides structure, order and frame to current popular practices that have evolved during the last decade. The taxonomy is then evaluated with the use of dimensions such as effectiveness/impact, satisfaction, adoption/engagement and quantity of feedback. The main finding of the taxonomical evaluation is the significance of developmental feed-forward guidance with which students are able to self-regulate and evaluate themselves. The paper concludes that this powerful combination should underpin further investigations into how assessment and feedback provision can be optimised for the experiential learning domain in general and to the work-based learning area in particular.

Keywords: Feedback, assessment, learning, self-regulation.

1 INTRODUCTION

Given the current profusion of research activity and ever-increasing output of publications describing and evaluating assessment feedback types, a degree of systematic and extensive study with the intention to classify can be purposive. As a starting point, and to illustrate the extent of literature available, a generic search of the terms ‘assessment’ and ‘feedback’ results in 5,639 articles from ERIC (Education Resources Information Center) of which 29% are over the last 5 years, 2,055 articles from IEEE Xplore of which 41% are over the last 5 years and 2,830 articles from the Web of Science database of which 43% are over the last 5 years. A recent substantial work [1] represents a monumental attempt at a systematic review of the literature where initial searches found 1,131 possible articles from 5 databases, of which 460 were considered.

The sheer volume of literature demands some structure to be assigned and in this current work, a comparatively broad range of research is accessed and considered in the context of particular themes. The themes are selected so as to represent the broadest areas of discourse whilst still helping to shape the initial review; they include subject disciplines, student audience, assessment types, methods and tools, use of technology, educator preferences and peer and self-assessment.

It can be argued that a broad divide exists between the assessment type requirements within Humanities and Sciences. A plethora of research has already been (and continues to be) conducted by educators in these two areas, however, there does not appear to be much overlap here. Research and case studies in the Humanities lean towards peer, self, group and counselling types of feedback, all within the context of essays, abstracts, reflections and reports [2], [3]. Research in the Sciences arena leans towards the use of technology-facilitated feedback for MCQ, online quizzes, adaptive e-learning and artificial intelligence. Language learning (particularly English) and its associated feedback provision have received much attention, with studies being developed to gauge the effectiveness of...
feedback for compositions, essays, portfolios and reports. The assessment and subsequent feedback of computer programming skills lends itself relatively easily to the online environment [4]. It is evident that discipline-based variations necessitate different forms of feedback and also that feedback which is predominantly used in one area may be beneficially deployed elsewhere.

When providing feedback, it is necessary to take account of the nature of the student cohort. Variations in the types of feedback provided for University and College students are generally different to those provided for primary and secondary School pupils [5]. There may also be cultural diversity, gender [6] and variations in study patterns of differing ethnic groupings [7] to consider. Individual ability and self-confidence will play a prominent role in the manner in which feedback is sought, digested, assimilated and applied. The hypothesis here is that an individual’s existing skill base will predetermine the use they are able to make of feedback [8], [9]. Students pursuing different disciplines will also add to the variation: science students may be assessed differently to humanities students and each will utilise feedback on the basis of their evaluative prowess gained elsewhere in their studies.

The type and method of assessment is fundamentally related to the type of feedback which can and should be provided. Educators can sometimes struggle with developing an assessment that serves multiple purposes. For example, it can be difficult to formulate an assessment that not only tests for grasp of factual knowledge but also emphasises understanding, creativity, inventiveness and real-world application. The predominant assessment types include diagnostic, formative, summative, norm-referenced, criterion-referenced, and interim or benchmark. Within the broad range of assessment method and tool categories available (for example unseen examination, coursework/assignment, class test, oral presentation, logbook/workbook and practical), there are several benefits of each category. Educators have experimented with various combinations of assessment methods and corresponding feedback types such as online [10], continuous [11], exemplars [12], virtual learning environment [13] and integrative [14].

Technology can improve assessment quality and it can also support diversity and accessibility. Recent years have seen an increase in the ways in which technology has been used to enhance the student experience of receiving feedback. Technological solutions to feedback provision include: adaptive eLearning environments [15], web-based environments [16] and [17], video technology [18], formative audio feedback [19], synchronous video [20], screencasts [21], digital pen [22], augmented reality [23] and online MCQ [24], feedback microblogs [25]. Sophisticated mobile and hand-held devices also open up alternative possibilities of feedback provision [26], [27], [28] which may be more readily acceptable to present-day students.

Educators have available to them a vast range of feedback mechanisms, but most prefer to use only a small subset of them. Preference choices are most likely based on experience, expertise, knowledge and confidence of technology and subject area. In attempting to categorise feedback types, the preferences of those giving the feedback needs to be observed [29]. What is evident from the diverse range of educator adoption levels of certain feedback practices is that it is challenging to form an overall judgement as to best practices. The novice educator therefore has to navigate the dense forest of current feedback practices in order to arrive at something which is workable for them.

The role of peer and self-assessment models and the subsequent feedback opportunities has been trialed in various contexts and particularly with online discussion forums [30], [31]. Self-assessment and self-evaluation can be valuable ways to enhance learning and this practice can be beneficial long after an individual has completed formal education. However, researchers find that students often significantly over-rate or under-rate themselves [32].

## 2 METHOD

Literature reviews ([1], [33], [34] and [10]) and seminal papers [35] play an important part in surveying scholarly output within a particular area of research and providing a critical and summary evaluation of what is currently available in the literature on feedback types. However, it is the role of the taxonomy to systematise a field so as to provide a useful framework for practitioners; the best example being Bloom’s Taxonomy of Education Objectives [36] which has been utilised by educationalists and researchers for decades. The rationale for developing a taxonomical classification of assessment feedback is to provide a systematic reference of the various types of feedback with associated criteria. The aim is for the taxonomy to aid in the highlighting of under-utilised or overlooked feedback types with a view to discerning any hybrid formats that may potentially work well. The taxonomy will also support the identification of weak or under-performing feedback types as well as distinguishing any
missing elements. Categorising feedback on student assessments can be problematic as much depends on the variability of educator preferences, student engagement and assessment types. In particular, educator preferences where tutors do not wish to modify practice or embrace new technologies can be a barrier and, on the other hand, student engagement with feedback can be erratic, with students often not understanding the feedback they receive. In addition, certain assessment types lend themselves to particular types of feedback opportunities whereas others do not.

In terms of the development of taxonomical structures in the assessment and feedback areas, recent examples include [37], [38], [39], [40] and [41]. However, these examples attempt to categorise very particular elements only, such as computer-based critiquing tools, adaptive feedback, group supervision, formative assessment for mathematical modelling problems and the characteristics of student peer mentors. Of particular note is the taxonomy of feedback as suggested by [42] which characterises feedback according to whether it relates to skills or content and whether it is retrospective or future-altering. Within this taxonomy, marking guides are used by tutors to give feedback on student submissions where retrospective feedback notes are made on-script and future-altering feedback is given on a separate summary sheet. Future-altering feedback points forward explicitly to future work, stating and justifying the skills that are to be developed through the course of assessment. The present taxonomy development will not restrict itself to specific elements, but will attempt to arrive at a general, higher level classification of all major feedback types, situations and their features.

2.1 Construction of the Taxonomy

A cross-sectional synopsis of the literature associated with feedback is utilised to formulate the taxonomy by investigating prevailing popular practices which have evolved during the last decade. The intention is to take an all-encompassing perspective, which does not disregard any subject disciplines, assessment methods or tools. The construction encompasses two main stages, namely taxonomy creation and taxonomy evaluation which includes taxonomy testing and validation. As a starting point, and to further emphasise the extremely broad and diverse nature of feedback practices, we begin by randomly enumerating, in Table 1, some possible feedback types and terms, as characterised in the current literature.

<table>
<thead>
<tr>
<th>TABLE 1. Illustrative Feedback Terminology</th>
</tr>
</thead>
<tbody>
<tr>
<td>peer-to-peer, learner-to-learner</td>
</tr>
<tr>
<td>self-feedback</td>
</tr>
<tr>
<td>group feedback</td>
</tr>
<tr>
<td>e-feedback</td>
</tr>
<tr>
<td>feedback on feedback</td>
</tr>
<tr>
<td>direct / indirect feedback</td>
</tr>
<tr>
<td>instant feedback</td>
</tr>
<tr>
<td>feedback in blended learning</td>
</tr>
<tr>
<td>technology-facilitated feedback</td>
</tr>
<tr>
<td>performance feedback</td>
</tr>
<tr>
<td>explicit / implicit feedback</td>
</tr>
<tr>
<td>feed up</td>
</tr>
<tr>
<td>evaluative / corrective feedback</td>
</tr>
<tr>
<td>discursive feedback</td>
</tr>
<tr>
<td>verbal feedback</td>
</tr>
<tr>
<td>written feedback</td>
</tr>
<tr>
<td>formative feedback</td>
</tr>
<tr>
<td>internal/external feedback</td>
</tr>
</tbody>
</table>
Taxonomising is hindered by the diversity, the overlap, the multiplicity and the contextual applicability of each feedback term. Certain feedback types can be immediately viewed as a category, for example, all those that make direct use of technology, such as web-based, adaptive, AI-assisted, VLE-based, audio and video; others however cross multiple groupings such as verbal feedback and corrective feedback. In addition, some feedback terms are synonymous whilst the meaning of others is dependent on context. In light of this apparent diversity of feedback terminology, seven main categories are formulated: assessment tool, context, feedback platform, feedback type, feedback situation, feedback method and purpose of feedback.

**Assessment tool.** The actual assessment tool or method has a huge impact on the nature of feedback provision available. Feedback styles on a written report submission, for example, differ vastly to the feedback given on a computer programming assignment, as do feedback for unseen examinations, portfolios, log books or oral presentations.

**Context.** The contextual environment has a bearing on the mechanisms available for providing feedback, for example, a feedback style adopted within a physical classroom-based activity may not necessarily be appropriate for a distance-learning environment.

**Feedback platform.** The platforms utilised for feedback can be broadly distinguished as either technology-facilitated or not. Technological solutions open up a wide variety of possibilities, although they still have to adhere to conventional best practices related to learning.

**Feedback type.** The envisaged intention of the feedback, whether it be diagnostic in nature or summative or formative is an important consideration which can dictate the timing that feedback can be given. Norm-referenced feedback compares a student’s performance with other students, whereas criterion-referenced feedback compares performance to a standard or criterion.

**Feedback situation.** As assessment situations can take many forms, from group to self to peer-to-peer, so the feedback situation should also ideally be represented in these groupings.

**Feedback method.** Marking proformas and annotations are the predominant ways in which to provide feedback on almost all categories of assessment.

**Purpose of feedback.** The two main purposes of feedback are to enhance performance and to aid learning and self development.

The resulting taxonomy (Fig. 1) depicts these seven main categories. Note that a taxonomical view is commonly hierarchical by nature, but in terms of the categorisation required here, a forced hierarchical perspective is not useful. Therefore, the classifications have been levelled as each is perceived to have equal contribution within a particular category.

**FIGURE 1.** Taxonomical Classification of Feedback
2.2 Taxonomy Evaluation: Selecting Dimensions

We now proceed with some evaluative work to further authenticate and validate the classifications. Evaluating feedback types requires the determining of dimensions which can be used to cross-reference and refine the understanding of the practice of certain types of feedback. In the current frame of reference, an initial determination of aspects such as effectiveness, adoption, satisfaction, quality, quantity, authenticity, context, impact, relevance, constructiveness and assessment type are considered. The choice of exactly which dimensions to include is governed by the requirement to appreciate, from the student’s point of view, exactly how effective and well-received a certain feedback type was. The challenge of choosing these particular dimensions is one of neutrality; the evaluation of feedback should progress in a non-biased manner where the goals of students and tutors coincide. The four dimensions, as shown in Fig. 2, of effectiveness/impact, satisfaction, adoption/engagement and quantity were selected to evaluate the taxonomy.

![FIGURE 2. Evaluation of the Taxonomy using Dimensions and Settings](image)

**Effectiveness/Impact:** As expected student perceptions in this dimension are two-fold: on the one hand feedback provides students with performance-related or grade information and on the other hand feedback facilitates task improvement and general development. The feedback would have a higher impact if students are able to use it to identify their strengths and weaknesses and as a consequence be able to improve performance.

**Satisfaction:** Feedback can be seen to be a factor in the quality of student experience with several positive aspects such as confidence-building, enhancing motivation and increasing self-esteem. However, feedback can also have unpredictable and negative effects where students feel demotivated and unable to make use of the feedback.

**Adoption/Engagement:** The level of engagement with feedback can be related to the student’s approach to learning (such as deep, surface, strategic or apathetic learning) and all learners will adopt feedback depending on their own learning style. Active engagement with feedback can enhance lifelong learning by enriching powers of reasoning and refining meta-cognitive skills. However, adoption of feedback is generally only increased if it is directly related to improvements in performance. In any case, feedback styles should seek to maximise adoption levels by ensuring clarity, accuracy, relevancy and also positivity.

**Quantity:** The granularity or size of feedback given can have an effect on the way it is consumed or assimilated by students. The ‘chunking’ of feedback is important to ensuring acceptable cognitive loads and allowing students to concentrate on crucial information.

2.3 Taxonomy Evaluation: Selecting Settings

Next, we move to a stage of systematically grouping elements of the taxonomical classification, as shown in Fig. 2, into a series of settings. Each setting is made up of 3 or 4 search terms which have been chosen to weave through the literature in a more unexpected way than was done in the initial review. Here we wish to elicit some added-value in the sense that settings can be made to be interesting by combining moderately unlikely feedback features together in a single setting. For example, a setting which is made up of the search terms ‘assessment’, ‘feedback’, ‘higher education’ and ‘essay’ yields a resulting 49 articles, which may give some insight into feedback practices for essay writing assignments that may be foreign to students entering higher education. A total of 10 settings were compiled and each setting’s terms were submitted as an advanced search to the Educational Resources Information Center (ERIC).
From the resulting list of articles, a subset (usually 3 to 4, although one setting warranted 6) was selected for scrutiny on the basis that:

1. The articles were published after January 2010.
2. Assessment feedback (rather than student feedback) was to be the focus.
3. The articles related findings primarily from empirical studies related to student perceptions.

There is the possibility that the articles’ content could be weighted or somehow quantitatively measured according to the dimension – for example, for any particular setting, a score (perhaps on a rating of 1 to 5) for each of the four dimensions ‘Effectiveness’, ‘Satisfaction’ ‘Adoption’, ‘Quantity’ could be given and then a final result score could be calculated by aggregation. However, there are problems of measuring in this way since there can be a degree of arbitrariness of weightings. Also the meaning of a resulting weighting would be elusive unless it was accompanied by further qualitative descriptions. Therefore, given the discursive nature of most of the findings and also given that the goal is to tease out and arrive at a series of properties or characteristics which could define effective feedback, a qualitative approach was taken to record key features within each article according to the four dimensions. Ten feedback settings were formulated, each with a minimum of 3 search terms – the first two of which in every case were to be ‘assessment’ and ‘feedback’. Table 2 shows the ten settings along with those articles that were selected for further investigation.

<table>
<thead>
<tr>
<th>Setting</th>
<th>Search Terms</th>
<th>Selected articles</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>assessment AND feedback AND higher education AND essay</td>
<td>[43], [44], [45], [46]</td>
</tr>
<tr>
<td>2</td>
<td>assessment AND feedback AND formative AND group</td>
<td>[47], [48], [49], [50], [51]</td>
</tr>
<tr>
<td>3</td>
<td>assessment AND feedback AND further education or school AND self</td>
<td>[52], [53], [54]</td>
</tr>
<tr>
<td>4</td>
<td>assessment AND feedback AND presentation</td>
<td>[55], [56], [57]</td>
</tr>
<tr>
<td>5</td>
<td>assessment AND feedback AND portfolio</td>
<td>[58], [59], [3], [60], [61], [62]</td>
</tr>
<tr>
<td>6</td>
<td>assessment AND feedback AND technology AND lecture</td>
<td>[63], [64]</td>
</tr>
<tr>
<td>7</td>
<td>assessment AND feedback AND work-related learning</td>
<td>[65]</td>
</tr>
<tr>
<td>8</td>
<td>assessment AND feedback AND distance learning AND technology</td>
<td>[42], [66], [67]</td>
</tr>
<tr>
<td>9</td>
<td>assessment AND feedback AND mathematic*</td>
<td>[68], [69], [70]</td>
</tr>
<tr>
<td>10</td>
<td>assessment AND feedback AND audio</td>
<td>[71], [72], [73], [74], [75]</td>
</tr>
</tbody>
</table>
3 RESULTS AND DISCUSSION

The following discussion is serviceably themed according to the original four dimensions as this allows for cross-referencing, cross-checking, comparison and summary.

Effectiveness/Impact: Several studies (for example, [43]) have found that a mixture of written and verbal feedback given without grades in the first instance improves the effectiveness of that feedback being taken seriously by students. Self-evaluation is deemed to be an impactful resource [54] where the student is required to assess his or work against detailed rubric [50]. Timing also plays a crucial part in the effectiveness of feedback assimilation [60], [64]. Several studies [71], [72], [73], [74] and [75] report a high degree of effectiveness and student positivity with audio feedback.

Satisfaction: A greater intrinsic motivation on the part of the student can be accomplished with more extensive and supportive tutor feedback [48]. [47] and [59] hypothesis that students with a strong sense of self-efficacy make better self-regulated learners who can plan, monitor and evaluate not only their learning but also the manner in which they assimilate and apply formative feedback. Positive feedback [52], gender-neutral strategies [53] and peer-assessment feedback [56] and [57] and interactive tasks with feedback [67] are all feedback types that can stimulate and motivate students.

Adoption/Engagement: Feedback styles should seek to maximise adoption levels by ensuring clarity, accuracy, relevancy and also positivity [46]. This transferability of achievement to alternative (future) assessments is an important facet of good feedback. The dual approach of providing detailed annotations and synoptic summary feedback, which many tutors take, can facilitate maximum adoption by students [45]. Many studies have attempted to maximise the engagement of students in the feedback process within an online [44] or MOOC [51] environment with peer review which facilitates a community learning approach. Timeliness of feedback [64] and the use of video technology [55] appear again to encourage engagement with formative feedback. Engagement levels can be increased by the use of developmental feedback [70] which highlights and promotes self-regulation [61]. Appreciation of individual learner styles and preferences with a more tailored approach to feedback [72] may also be necessary to increase adoption levels.

Quantity: There appear to be wide variations in the granularity or size of feedback given to students [49] which can have a detrimental effect to students ability to assimilate and utilise the feedback [43]. Continuous feedback which takes an iterative form [3] can be particularly helpful and can be more readily achievable in an online context [51].

4 CONCLUSIONS

The taxonomy evaluation has given further insights into the manner in which students perceive feedback. Recurring common themes around the most effective types of feedback centre on self-regulated learning, with much agreement in the literature that where such developmental and personalised feedback is made available, the chances of students being able to self-regulate themselves and genuinely assimilate and apply the feedback is much higher. It is acknowledged that self-regulation is not an easy learning skill to acquire as it requires meta-cognitive awareness, autonomy, strategic action in the form of planning, monitoring and evaluation, and above all else a motivation to learn. Developmental, continuous feedback cycles, as well as future-altering concepts appear to be the categories of feedback that can yield the most long-term benefits for students. Feeding forward is a concept that is synonymous with development as it involves students in determining how feedback relates to their own understanding and about ways to apply the feedback in new or unfamiliar situations. Another common theme is the provision of regular formative feedback which is motivational and digestible (in terms of quantity) as being the most functional. However, there is common agreement that students need to be active rather than passive recipients of feedback and one popular method of engagement is that of peer review and feedback.

The evaluation has highlighted an area where further attention is required, namely the type of feedback provision that is conducive to the experiential learning domain in general and to the work-related learning area in particular. The main findings of the taxonomical classification and evaluation developed here, namely the significance of developmental feed-forward guidance with which students are able to self-regulate and evaluate themselves can underpin further investigations into how assessment and feedback provision can be optimised for work-related learning contexts. The standard feedback process, with its focus on measurement of effectiveness of learning outcomes in particular and student performance in general, and its inability to propel and move forward to goal attainment incrementally, may be one important reason why students may not heed feedback comments as they
are not able to appreciate how that feedback will allow improvements in learning and ultimately performance in later assignments, or in this case, to workplace tasks. The concept of feed-forward may therefore be more valuable for work-related learning development than the more common feedback strategies. Students should ideally be exposed to setting their own targets and stretch targets within a more direct and formalised feed-forward environment; thereby developing a degree of maturity and self-confidence. It is evident that negotiated learning agreements as used typically in work-related learning initiatives need further enhancement of feedback practices to better align them to the performance measuring tools used in industry.

In summary, developmental feedback paired with self-regulation and self-evaluation appears to be a powerful combination that could prove to be effective in the work-based and work-related learning context.

REFERENCES


[64] Voelkel, S. (2013). Combining the formative with the summative: The development of a two-stage online test to encourage engagement and provide personal feedback in large classes. Research in Learning Technology, 21(1), 18.


