What if feedback only counted if the learner used it?

by Steve Draper, University of Glasgow


Abstract

Most teachers give written feedback because it is a required deliverable, like a checkout assistant handing every customer the printed receipt, even though few use them. This paper explores taking the inverse position. Instead of assuming that feedback is good, assume that the only thing that matters is the actual use of feedback by learners. Symptoms from both practice and theory suggest this might be an important shift in perspective. Three types of implication are identified and developed, and specific examples of each are described. Firstly: we need to research what goals do learners actually want to use feedback for. Secondly, we should study cases of success, where feedback has definitely been used by learners. Thirdly, converting feedback into learning (into changes in knowledge) involves active, constructive, effortful mental labour. How can this be prompted and supported?

1. Introduction

This paper concerns feedback (from teachers to learners) in Higher Education (HE). It explores the implications of a shift of perspective from focusing on the supply of feedback to focusing on its consumption: on whether and how it is used by learners (and therefore could contribute to learning). This shift is like the difference between a rain god, proud of her broadcast provision across the planet, and especially in winter and on mountain tops (when and where it is least needed even by plants); and a disaster manager, focussed on emergency provision of safe drinking water to humans, where it makes a vital, life-saving difference, but only when it reaches and is actually drunk by people. The enormous difference between the total rainfall over a country, and the volume of bottled water actually drunk might make us wonder for a moment how much of the feedback available is actually productive of learning. The paper starts with some reasons for such a shift of perspective: the frequent disregard of feedback, particularly comments, by students, the outdated teacher-centredness of viewing feedback as a one-way action by teachers on learners, and the overlooking of the active work by students required to process feedback into action. It then develops the new perspective by discussing cases where learners have definitely used feedback productively, how they have used it, and how they could or wish to use it. The relationship of the argument here with related points in the existing literature is briefly reviewed. It concludes by suggesting what was overlooked in past research on feedback, and by summarising the new perspective as a design approach for practitioners and innovators.

In this paper, I use "marks" to refer to a quantitative judgement of merit (whether a grade or a percentage or some other unit), and "comments" to refer to open-ended, natural language communication about a piece of student work. I use "feedback" to refer to either or both of these: to any potentially useful information a student receives about their work; and "assessment task" to refer to the task done by a learner which may receive feedback. I would count feedback as having been used (i.e. effective) when and only when the learner modifies or actively reappraises something specific as a result of it.

Bear in mind however that people usually tacitly assume that "feedback" means "individualised comments from another person", or in education "written comments from a teacher", unless the context strongly imposes another meaning. If you ask a footballer practising set-piece kicks whether she got feedback on an attempt, she is unlikely to say "yes, I could see that my shot went over the goal". If you ask a schoolboy doing mathematics homework whether he got feedback, he's not very likely to say "yes, I checked my answer against those in the back of the book". If you ask a student discussing a key concept with friends whether that was useful she might say yes, but if you ask her if she got feedback she is unlikely to say that the differing views gave her diagnostic feedback on what she did and didn't understand adequately.

If we are to understand what best promotes learning, however, we don't want to presuppose that the solution is individualised comments from another person, but to enquire what sources (under what circumstances) can actually be shown to do that. This difference in what is understood as "feedback" renders much data on feedback ambiguous: whether it is the National Student Survey (NSS), or some of the interviews with students mentioned below.
2. Why change our perspective on feedback?

2.1 Symptoms of the problem from practice

What follows is a mixture of reports from my own experience with feedback and students, which I expect will find echoes in many readers' experience, and published accounts of similar phenomena.

According to surveys such as the NSS, students often complain about not getting enough feedback, although this is sometimes contested by staff. Either way, however, it suggests that feedback is not helping them adequately with the goals for which they wish to employ it.

Students often don't collect feedback, which suggests they don't expect to use it. Even more common is that on picking it up, they look at the mark and ignore the comments (Hounsell, 1987; Orsmond et al. 2005). Thus often feedback comments are not used (because not seen) by students even when supplied and put in front of students: a problem of processing not of supply.

I have had students read my feedback (because I adopted advice from the literature, in light of the above issues, to force them to read it in my presence) but say that they do not see how it would help with future work. Since this includes students who went on to be ranked near the top of a competitive class, it is unlikely that this is something wrong with the students. Increasingly staff complain that students will not do formative work unless it is summative: i.e. required. This indicates that many students have no expectation of feedback benefitting their learning, perhaps because they have no experience, or at least awareness, of it doing so.

These symptoms do not directly tell us what to do, but they do suggest that simply supplying feedback is a failed approach in practice. To understand more, we should study cases where it actually is used by learners.

2.2 Symptoms of the problem from theory

Considering feedback as something that is done by teachers to students is a profoundly teacher-centered view, and completely incompatible with a constructivist approach. Considering what and when feedback is used brings learners back into the consideration they receive in all other current theory of learning and teaching.

Similarly, transmitting individual written feedback is at odds both with social aspects of learning (peer interaction, communities of learning), and even more basically, with any ideas of the central importance to feedback of dialogue, as explored in Nicol (2010). Hattie (2009, p.173) goes so far as to say that he now sees that feedback is most powerful when it is from learners to teachers: opening a dialogue, and permitting contingent (as opposed to monologic) communication from the teacher.

2.3 Learning from feedback requires significant mental processing

If you are using feedback for work, e.g. polishing a document for a customer, then the most effective feedback will dictate to you how to revise the product, and hopefully requires little effort to understand. Marking up proofs for a printer is like this. If however you want feedback to cause learning rather than work productivity, then things are different: and this is too seldom acknowledged, and the implications followed up.

Consider as an example the case of feedback that corrects the spelling of a word. If the reader thinks this is trivial and beneath the notice of a teacher in HE, then remember that a theory which cannot even cover this case is certainly inadequate for more important feedback. What, ideally, are the consequences for the learner of such feedback?

a) If the current document is to be resubmitted, then the word's spelling should be corrected in this document. And probably all other occurrences of the same word should also be checked (an inference from the specific case to the general case).

b) A correction to the student's internal spelling-generator needs to be made, so that they don't make the same error in future. I.e. this is about learning, rather than getting the work done but learning nothing. For example, in papers on feedback such as this one, a landmark paper is Black & Wiliam (1998) and you need to learn that this Wiliam is spelled with only one 'l': and many authors get it wrong. If you continue to write in this area, you need to learn this by heart.

C) A modification to the student's internal "proof-reader" or bug detector, so that they notice the error if it recurs in either their own work or someone else's. You can't be said to know a spelling if you can't correct it in a colleague's work; and your own as well. This is perceptual learning, rather than action learning.

Even this simple case begins to show how learning from feedback requires inference, generalisation, and making multiple separate mental changes triggered by a single indication (perhaps one instance of the
spelling being corrected). If the feedback had instead been "you cannot divide by zero" or "your essay introduction must both outline what is to come and locate your essay in the wider literature" then how long would it take a student to work through the implications, which impinge on the strategic level of how the whole task is tackled? How likely is it that this processing will be done in the time it takes to glance over the comments? What are the conditions under which this mental work, and so the learning from some feedback, will be done at all?

2.4 Implications of the new approach
At this point in the argument we may make three points about the proposition that feedback only has value when it leads to changes in the learner:
1. Unless learners are interested in using feedback, it isn't likely to be used. So we need to find out what they want to use feedback for. (See the next section.)
2. We should study cases of success, where feedback has led to learner actions, as opposed to studying feedback as a thing in itself, a deliverable independent of customers. (See the section after next.)
3. Converting feedback into learning (changes in knowledge) must involve active, constructive mental work by the learner. How can this be arranged and promoted? Where is it timetabled? (See the previous section.)

3. What are learners trying to change (regulate)?
The first consequence of thinking about what students might do with feedback is to consider their aims. (This has been discussed in Draper, 2009b.) Almost the whole of the literature on feedback presupposes that students' only use of feedback is to improve their technical knowledge and skill at the subject. Asking my students about feedback quickly disabused me of this.

When I asked them why they looked at the mark before the comments, and what they would do differently (if anything) if the mark had been much higher or lower, their replies essentially expressed that they were checking if they were on track for the level of marks they wanted, and if not they would increase or lower their effort accordingly. Just like most professionals, they have many goals to attain and limited time and attention. Therefore they must divert effort if necessary but otherwise economise on how they divide their time. This is required of them by the institution, which gives them multiple courses (and deadlines) at the same time. This is effort self-regulation, and what students most often use the mark for. We'll see an example of this later.

On another occasion when I asked a student what she thought of the feedback she had got in a previous year, she said (implying that she hadn't got useful feedback for this) that all she wanted to know was whether she was really a psychologist, or whether she should switch subjects. The student goal here is making career decisions within the university. The education system requires students to make such choices, and the major information they have is the marks they get. This is another important student goal, imposed on them by the institution, and which they use feedback for, but which is generally ignored by the literature on feedback.

Another common student response is that they would like to know how to get a higher grade next time. This at least seems to be partly about technical knowledge: but note that it is actually not phrased as a question about understanding the subject, but about the criterion-based marking system. Teachers relatively seldom phrase their feedback in these terms, and so fail to match the goal the student is trying to use feedback to satisfy.

The above three correspond to common student responses, but some more sophisticated students use feedback to draw other inferences such as the validity of some markers compared to others, the degree of randomness in the process and so on. These other interpretations of feedback by students (and by staff e.g. when receiving peer review comments from journals) are discussed in Draper (2009b).

All three of the student goals mentioned here, and readily apparent from talking to them about how they do or would like to use feedback, are in fact imposed by the education system: they are not due to idiosyncratic or perverse aims. Yet most literature and practice fails to address these issues, and only seeks to increase skill and understanding of the topic: which is an aim typical of academics, but ignores actual student needs which in turn mainly derive from the situation imposed on them by the institution.
4. Cases of students successfully using feedback

The second implication of developing the new approach is that we should study success cases, where feedback to students actually is used. In fact some do exist, and they suggest ways forward.

I will briefly describe three such successes. Between them, they will cover feedback in both a calculation-based discipline, and an essay-based one. Furthermore they will illustrate three different student goals: regulation of effort (should I work harder or less hard on this subject?); regulation of understanding (what bits of this subject don't I understand properly?); and improving technical skill (how to write an essay; or how to do calculation problems in this discipline). Additionally they will illustrate how three different types of information may each alone support action by students: marks, comments, and the existence of a problem (with no information about how to act on it).

4.1 "3-D feedback": making marks comprehensible and usable

A colleague, Eric Yao, has more than doubled the pass rate on a first year physics course he teaches. It was around 40% historically and the first year he took it over, when he essentially repeated the existing course design. Three years later he achieved a 95% pass rate. Because multiple things were changed, we cannot be certain which were the important factors, but one interpretation is as follows. He required students once a fortnight to answer a set of MCQs from an online bank associated with the textbook, which were of course automatically marked. However he then emailed each student individually to comment on the mark they had got. This personal attention is likely in itself to have an important effect by embodying "teacher monitoring" (Draper, 2008): showing that each student and their work is individually noticed by the teacher. This is a significant amount of work, but feasible given that the setting and marking of the test questions required no teacher effort; and although the emails were individual, they were composed in a fairly routine pattern which could probably be rapidly generated.

Marks, like all measurements, are only meaningful in comparison to other already known measurements on the same scale: a contrast set. If I tell you a volume in minims, a weight in scruples or a temperature in degrees Réaumur it is only theoretically, not practically, informative. But what comparisons are useful for and actionable by students? Yao typically commented on a student's mark in two ways. One was in comparison to that student's last mark: giving the mark an "ipsative" reference point (about whether the student was improving or dropping off) see Hughes (2011). The other was in comparison to the marks the class as a whole had got on the same set of questions (a "normative" reference point): e.g. this was a difficult set for the class, so your lower mark actually suggests you are doing as well as before.

A third such contrast set might also be very useful to many students: what might be called "predictive" reference points i.e. information about what their current mark suggests is most likely to be their eventual grade for the whole course or degree programme. If the assessment tasks haven't changed, then data from previous years make this easy for a teacher to calculate, complete with confidence limits of some kind e.g. "in previous years, half of those with this mark went on to an eventual B grade, while a quarter got As, and a quarter got lower overall grades". Students usually cannot calculate this, so raw grades cannot give them meaningful self-regulatory information. Grades for one assignment generally have no simple predictive power and therefore are not meaningfully "criterion based" from a student's viewpoint for reasons including: a) Coursework generally gets higher marks than exam work (Elton, 1998; Bridges et al., 2002); b) If approximately the same marking standards are applied over one or two years, then average student marks within a cohort will increase over that time as they develop i.e. become better learners, better students (which is after all a chief aim of education). c) Unless all assessments on a course are of identical type, then they will not measure the same thing, and marks for one will have an unknown (to students) relationship to overall marks. Snyder (1971) describes how different students align themselves with different measures (i.e. assignment types) on a given course, giving different long term outcomes; although he implies that perhaps this is a student choice. However surely teachers should provide students with information on the meaning (predictive value) of each measure.

I interviewed some students from Yao's course. One said he didn't like the 9am lectures and if he missed one he felt he had caught up by reading the slides etc. on line. But he noticed that the quiz marks he got didn't support this feeling and so he increased his effort to keep up attendance. This student finally achieved an A grade. However it looks as if his conscientiousness would not have been enough without the feedback from Yao's system of self-testing plus teacher interpretation of the results: he would not have realised his actual level of comprehension until it was too late to adjust.

Thus this case shows how regular marks without comments (i.e. without any technical "constructive" / directive content on the subject matter) can be useful feedback, which students do use formatively to self-regulate their performance: in this case, their effort. This appeared to make a dramatic difference to eventual
overall outcomes on this course. However it also suggests that raw marks alone may not do it, without translating them against contrast sets that are informative in connection with students' self-regulatory goals. The learner goal here is the self-regulation of effort, and uses marks as the input, but needs interpretative help to translate marks into usable information for this goal.

4.2 Prompting the processing of feedback

A success I myself have had has been achieved by prompting my students to process my written feedback into plans for action. The context is weekly tutorial groups of 5-6 which (among other things) cover the supervision of a large piece of coursework (critical reviews). In these groups I also organise reciprocal peer critiquing, which students value, and which also sets up a good peer atmosphere for discussion. Nevertheless my own feedback to them seemed less successful even though I:

- Provided the written feedback in typed not hand-written form
- Gave both positive and negative comments
- Suggested specific changes that could have been made
- Used elective feedback i.e. had each student preface their submission with 1-3 questions which they specifically wanted feedback on
- Give them all the feedback for all of them so that they can learn vicariously as well as directly.
- Require them to collect the feedback from me, and read it on the spot.
- Promote discussion of the feedback with me and with each other.

I generally failed to get substantial discussion flowing, and if asked then even the best would say they didn't see how these comments, although interesting, would help them with future assignments. It seemed that their memory of their original work had faded, and that even reading extensive feedback was not enough to make them think about it actively. Then, adopting a suggestion from David Nicol, I added a simple five minute exercise, getting each student to fill out a prompt form essentially asking them to write out briefly how they would apply the feedback in future. Following that, discussion flowed freely when I went round the group asking what kind of thing they had written down; and all found either the discussion or the exercise or both valuable. If we may take written intentions as indicative of future actions, then this seems to have transformed whether my feedback was actually used.

This case suggests that written comments require significant mental processing; that this does not happen naturally even when the students definitely read the comments and have a carefully provided opportunity for both peer and tutor discussion of them; yet a short and simple exercise can change this. This concerns the learner goal of improving their technical skill or knowledge, using written, open-ended comments, but may need an extra prompt for the necessary conversion of comments into learning.

4.3 Catalytic assessment (questions, not answers)

Hunt (1982) showed that simply requiring learners to state a confidence level after answering each question on tests, increased overall learning speed. This shows that technical content is not necessary for feedback to benefit learning. Becoming aware of where knowledge needs to be improved can lead to self-remediation. In Hunt's case it was self-evaluation. Other work has shown that other sources are effective in this way at prompting self-regulation of knowledge and understanding. Piaget argued that peer interaction was better than teacher interaction for this, because authoritative sources tend to be accepted without thought, while disagreements with peers cannot reasonably be either dismissed or accepted without thinking. Thus peer discussion is good both for identifying gaps in secure understanding, and in promoting the establishment of reasons for a belief i.e. deeper learning that simply remembering a fact. This line of argument is given in greater detail in Draper (2009a), and has led to impressive learning gains in HE contexts: see Crouch & Mazur (2001), Smith et al. (2009). The latter work also revolves around "brain teasers": carefully designed questions which elicit uncertainty by having two or more answers which seem plausible. Thus self-ratings of confidence, brain teasers, and peer discussion are three forms of "catalytic assessment" where the feedback amounts to showing the learner that they do not understand a specific topic properly, and they have a demonstrated ability to prompt self-remediation without further input from a teacher, and so to cause learning.

In this they are similar to learning mathematics by doing examples in a textbook and checking with the answer in the back of the book: again, all the learner is directly told is whether or not they need to do further work to master this topic. This is the essence of Mastery Learning (Bloom 1984; Block, 1971), where a key feature is regular diagnostic, formative testing, followed by learners focussing on the specific parts of the topic they have not yet mastered, and usually self-remediating from resources other than their teachers e.g. small groups, alternative textbooks. This is consistent with the work of Roediger & Karpicke (2006) showing that self-testing is a much more effective learning tool than simply re-reading material.
These are all similar in that the feedback in effect diagnoses what needs remedial work, but does not carry any of the remedial content itself. Overall the evidence suggests that we (learners) are not very good at estimating what we know or understand. However once our attention has been drawn to a problem in our understanding, we often actively remedy it. These are cases of the learner goal of self-regulation of understanding, but where stimuli such as formative tests are a key resource for triggering it.

4.4 Other success cases worth examining

Elective feedback, where learners ask for the feedback they want, is obviously very promising from this perspective since it seems much more likely that students will act on feedback they actively sought, than feedback that was just dumped on them. Bloxham & Campbell (2010) report trials of this approach, which they call "interactive cover sheets". It can easily be combined with other techniques, as it is in my own practice (although it is also interesting to consider only giving feedback that is asked for). My impression is that it is important for about one in three students, who are concerned about some aspect of their work which in fact they are doing well enough that it never draws a comment: but they would like an explicit discussion of the issue to put an end to uncertainty (and perhaps improve slightly).

The patchwork text design (Scoggins & Winter, 1999) is interesting in a number of ways. It is designed to serve an unusual learning goal: how to pick a personal topic for a large assignment, and to use feedback to help make this decision. It is thus the only design I'm aware of for providing feedback and scaffolding for "creativity". The design provides peer ipsative feedback, by having the same small group provide each other feedback each week over many small assignments. Thus this is a case of identifying a special learner need; of attempting to design a system of feedback to serve this need; and creating an original course design to provide that feedback, and so serve the learner goal.

Feedback calendars are tables, published to students on a course, that list not only the task submission deadlines for students, but the date and type of feedback which will be returned by staff to students. Their main value is to make the cycle of student work and returned feedback more visible to both students and staff. As mentioned above, there are indications that both staff and students tend not to have any real view of feedback as part of learning, and that receiving (and thinking about it) is an important learning activity. More details and examples may be found at [URL].

5. Discussion

This paper explored the proposition that there is no point in producing feedback for a learner unless the learner uses it. Various elements of the points made are present in previous literature.

There are other papers that have pointed out that, if any benefit is to accrue, the learner must be not a passive recipient but active in often complex processing of feedback (e.g. Butler & Winne, 1995). There are other papers that analyse the overlooked complexity of what a learner's thought processes might be when they process feedback into new learning (e.g. Nicol, 2013). There are papers that point out that to understand the benefits of feedback we must take a more longitudinal view of the learning process: of repeated attempts and repeated feedback and feedforward (e.g. Molloy & Boud, 2013). Price at al. (2010) have begun the attempt to measure its effectiveness, but found it hard to do.

This paper set out to frame such papers as simply and clearly as possible, and to sharpen their emerging lessons, by replacing the presupposition that feedback must be necessary, and the further presupposition that feedback is only to be used to fulfil the teachers' rather than the learners' goals, with the question: how different would the world of feedback be, both theoretically and practically, if feedback only counted if the learner actually used it? While Gibbs & Simpson (2004) listed 10 conditions for assessment and feedback to promote learning, they put only as the last two that feedback must be received, attended to, and acted upon. This paper suggests reversing that and making them the first and only conditions; and in fact that it is the last (the learner acting on feedback) which is the only essential one from which others may, if necessary, be derived. This is a complete change of perspective, rather than of content.
5.1 Implications for research

Implications for research on feedback are:

- Researchers should measure and report what learning or other action results from feedback. With proper direct measures of the use and effect of various types of feedback, differences in the usage of the term "feedback" will become relatively unimportant, and it will matter less whether or not students are aware of the effect of the information sources which actually influence their learning.
- I have had several (very high achieving) students assert that feedback has never contributed to their learning. This is in direct contradiction to the assertions and assumptions of many researchers. Collecting data is essential to resolve this contradiction.
- It may be that those students took "feedback" to mean "written comments". If that is the explanation, then abandoning the production of written feedback by staff would be a large cost saving: finding out if this is sensible is clearly important.
- We should study what learners in fact use feedback (in a wider sense) for; and what they say they want to use it for.
- Studying success cases, as sketched above, seems a promising research priority to begin understanding the conditions under which feedback does increase learning.
- Some well known reviews, such as Black & Wiliam (1998) and Hattie & Timperley (2007), concluded that feedback was an important causal factor in learning. Perhaps these should be re-examined. For instance, it could be suggested that in all the studies they review, it wasn't just the presence or absence of feedback that was varied: other features of the teaching methods varied as well. How do we know that feedback was the cause of the variations in learning outcomes just because the reviewers and perhaps the original researchers used the term "feedback" to label a complex learning design used in an experiment?

Bloom's mastery learning, which according to Bloom (1984) produced an effect size of 1.0 (larger than the reviews typically report for "feedback"). However it is described by its practitioners not so much as consisting of formative testing, as of retraining students on how to interpret and act on the feedback; to interpret it not as a measure of their ability but as information about how to spend the timetabled self-remediation sessions, and then to demonstrate to themselves by re-testing that this interpretation is correct and the usual ability-measuring interpretation is wrong in their own case as well as in general. (See for example, Block, 1971.)

Hattie & Timperley (2007) make it clear that feedback quite often reduces learning: a powerful factor, but not always positive. Hattie (2009) shows a still more complex view of feedback, perhaps consistent with the simple framework presented in this paper. The general message is that to understand the evidence on feedback, we need to move far beyond the simple assumptions still prevalent.

5.2 Implications for practice

These might be expressed as follows. The topic could be defined as "feedback that is used by learners". Here the aspiration, and measure of success, is that feedback is only worthwhile to the extent that the learner uses it by modifying or actively reappraising something specific as a result.

Two design principles suggested by the cases discussed above are:

1) Ensure there is something which triggers the learner into processing any feedback into future actions.
2) Ensure marks are expressed on scales, or with anchor points, that are already understood by the learner, and which are relevant to their purposes (their self-regulation of learning actions). Expect that several different scales or comparisons need to be provided for each mark.

Acknowledgements

This paper owes much to the author's long collaboration with David Nicol.

Some of the work underlying this paper was supported by Glasgow University's Learning and Teaching Development Fund.
References


Bloom, B.S. (1984) "The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring" *Educational researcher* vol.13 no.6 pp.4-16


Draper, S.W. (2009a) "Catalytic assessment: understanding how MCQs and EVS can foster deep learning" *British Journal of Educational Technology* vol.40 no.2 pp.285-293

Draper, S.W. (2009b) "What are learners actually regulating when given feedback?" *British Journal of Educational Technology* vol.40 no.2 pp.306-315

Elton, L. (1998) "Are UK degree standards going up, down or sideways?" *Studies in Higher Education* vol.23 no.1 pp.35-42

Gibbs, G. & Simpson, C. (2004) "Conditions under which assessment supports students' learning" *Learning and teaching in higher education* no.1 pp.3-31


Hughes, G. (2011) "Towards a personal best: a case for introducing ipsative assessment in higher education" *Studies in Higher Education* vol.36 no.3 pp.353-367


Nicol, D. (2013) "Resituating feedback from the reactive to the proactive" ch.3 in D. Boud & E. Molloy (eds.) Feedback in higher and professional education: Understanding it and doing it well (London: Routledge)


