Beyond feedback: Developing student capability in complex appraisal

D. Royce Sadler
Griffith Institute for Higher Education, Griffith University, Brisbane

Abstract

A common practice is to give students detailed feedback about the strengths and weaknesses of their work, with suggestions for improvement. However, the impact of feedback often seems negligible, despite the investment of considerable time and effort put into its construction. With a view to increasing its effectiveness, extensive theoretical and empirical research has been carried out into structure, timing and other parameters. For students to be able to apply feedback, they need to understand the meaning of the feedback statements. They also need to identify, with near certainty, the particular aspects of their work that need attention. For this to occur, students must possess critical background knowledge. This article sets out the nature of this knowledge and how students can acquire it. They must appropriate for themselves three fundamental concepts, namely response genre, quality, and criteria, and in addition develop a cache of relevant tacit knowledge.

_____

Introduction

Numerous studies have been carried out into the effectiveness of teacher feedback for improving student learning. Early researchers defined feedback narrowly as knowledge of test results, and investigated its effectiveness from a behaviourist perspective. Knowledge of results was one element in the cycle of stimulus (test item), response (student answer), outcome (correctness) and reinforcement (credit). Later research took a broader view: feedback included teachers' verbal communications intended to help students improve. Two major reviews of research into feedback and formative assessment were carried out by Crooks (1988) and Black and Wiliam (1998). Both reviews supported the view that feedback can lead to improvement, but does not always. It has to be well constructed and certain other

conditions have to be satisfied. Research continues across all sectors of education, including higher education. The study by Higgins et al. (2002) is typical.

For this article, the point of departure from traditional research is the premise that, regardless of levels of motivation to learn, students cannot convert feedback statements into improvement without sufficient working knowledge of fundamental concepts that are routinely assumed by the teachers who compose the feedback. Unless this prerequisite knowledge is identified and addressed, the prospects for even the most thorough feedback will remain limited.

Scope and terminology

The assessment tasks of specific interest require students to demonstrate higher cognitive skills or forms of professional proficiency. Achievement in these is typically assessed using divergent tasks which require complex extended responses that involve analysis, synthesis, creativity, evaluation, or critical thinking. Such responses are common in a wide variety of disciplines and fields at both undergraduate and postgraduate levels, including languages, humanities, health and social sciences, visual and performing arts, the professions, and many areas of science and technology. How well a student response achieves the purpose set out in the task specifications is determined by making direct qualitative judgments.

Feedback is taken to include all comments a teacher provides to a student following appraisal of a student response. To simplify discussion, assume that the task specifications are given to students in written form, and that the feedback is conveyed in text, audio recording or other permanent form that allows for continuing access. The assessor is generally referred to as the teacher, although in practice it may be a tutor, teaching assistant or other suitably qualified person. Judgment and appraisal are used as synonyms. An assessment event is concluded when the teacher has evaluated the responses, advised students of the outcomes and provided any relevant information (such as feedback, marks or grades) to both students and the institution. This basic sequence of teaching – setting an assessment task, appraising student responses, and returning information about performance – is widespread practice in higher education.

Feedforward and feedback

The two main tools for steering students towards a work of the type the teacher requires are the assessment task specifications (with clear parameters within which to construct a response) and, possibly, information about how the work will be assessed. The actual design is left open. Using Bjorkman's (1978) terminology, these 'feedforward' communications are future oriented. At this point a caveat is necessary: the legitimacy of providing students with advance notice of the criteria and standards as a rubric or in some other format is contested

territory, as indicated by Sadler (2009b) and others. This issue is not taken further here because it is secondary to the main theme.

The first of two broad functions of feedback is to provide the teacher's assessment of the student response, maybe as a grade, with a rationale that indicates how the judgment took the strengths and weaknesses of the response into account. The second function is to provide advice or suggestions as to how a better response could have been constructed. Clearly, if feedback is to have a reasonable prospect of achieving its formative purpose, it has to be both specific (referring, as it necessarily does, to the work just appraised) and general (identifying a broader principle that could be applied to later works). So although feedback is mainly retrospective, it has a prospective orientation as well. Technically, this carry-forward component constitutes an additional element of feedforward.

General recommendations in the literature about the desirable properties of feedback include: complimenting students on the strengths of their works; telling them (gently) about deficiencies, where they occurred, and their nature; telling students what would have improved their submitted productions; and pointing them to what could be done next time they complete a related type of response. Throughout, feedback should aim to be constructive and supportive. For most teachers, providing feedback with these characteristics is labour intensive and cognitively demanding. They give careful thought to exactness in wording, because the feedback will later stand as a discrete communication that can be accessed multiple times. The volume of feedback for a particular work depends partly on the extent to which the work is deemed salvageable. For those that do seem salvageable, the teacher may provide considerable detail. For high quality work, there may not be much to be said, and for pathologically poor work, it may be difficult for the teacher to know where to begin. Furthermore, because the communication is asynchronous, the teacher has to anticipate how the student is likely to react to both the content and tone of the feedback, and this calls for a significant affective outlay on the teacher's part. Notwithstanding the limited effect feedback often seems to have, conscientious teachers continue to invest heavily in providing it to their students.

Feedforward and feedback share an important characteristic: as one-way messages from the teacher to the student, they are essentially about telling, or disclosure. Yet despite the teachers' best efforts to make disclosure full, objective and precise, many students do not understand it appropriately because, as is argued below, they are not equipped to decode the statements properly. By implication, teachers who are committed to providing high quality feedback want it to work for their students. This suggests that complementary attention should be directed to what students make of the feedback, rather than just its composition. Seen from the learner's perspective, this represents a shift in emphasis from disclosure (by the teacher) to visibility (to the student). How can the situation be improved for the student? To

start with, those parts of feedback that specifically deal with strengths, weaknesses and especially guidance for improving future works are more than mere conduits of information; for the most part, they are intentionally expository and didactic. The teacher wishes the student to learn from the assessment event, so the text of the feedback is an instructional medium. The next question is: How do humans learn from expository text? During the 1960s, this topic was investigated under the banner of 'meaningful verbal earning' by numerous researchers, among them Ausubel (1963) and Carroll (1968).

Obviously, feedback refers to a student work which is, in principle, equally accessible to both teacher and student, so it might be thought that the student's processes of interpreting and learning from it would be straightforward. To test this assumption, it is necessary to take a step back from the feedback itself and analyse the contexts in which, respectively, teachers compose (expository) feedback and students interpret and make use of it. Putting aside any personal biases the teacher may have towards or against particular students, the teacher looks at the work through more or less objective eyes, appraising it as an external observer or consumer. Although the teacher may make assumptions about what the producer was intending, this is not known for certain. It does have an influence, however, on how the feedback is framed. Additional factors come into play. In particular, the appraisal may be made at a time and site remote from the student (depending on the type of student response), and a teacher may have such large classes that they barely know students except through their works. The teacher nevertheless accepts a considerable responsibility in trying to turn an assessment episode into a significant learning event.

Students face several interpretive challenges in trying to capitalise on feedback. The first relates to the work as a whole: students may focus partly on the works exactly as submitted or performed, and partly on what they had intended them to be. The learner's personal investment in the production then blurs the boundary between the two. The second challenge arises when the feedback's implications for action are dependent on student understanding of certain concepts or criteria used in the communication. Teachers become accustomed to using certain terms, and can easily presume that students know what they mean. A third type of challenge is experienced by the student who lacks the tacit knowledge necessary to identify the feature of their work to which some part of the feedback refers. For instance, a teacher may annotate a section with the comment "This does not follow logically from what goes before". If the student cannot see any problem with the logic, no action can be taken. On the other hand, to explain why the logic does not follow might require a paragraph of explanation, and the teacher either cannot afford the time to compose it, or does not see it as necessary.

In all three situations, the student cannot make critical connections between the feedback and the work. The feedback statements then fail as communications, and the telling is to no avail. Recasting the disclosure statements or elaborating them to provide finer detail is futile

while the primary referents remain fuzzy. Assuming that the interpretation problem is solved, the fourth challenge for learners is to assimilate the teacher's feedback into their existing knowledge bases so that it can be drawn on, as needed, in future constructive activity. As with all learning, such newly conceptualised knowledge needs to be consolidated before it decays if it is to have a positive influence on future works. Both interpretation and assimilation depend on verbalisation and reasoning, and hence familiarity with the concepts and vocabulary that are relevant to translating specifically evaluative discourse into the discourse of (future) production.

Clearly, the gap between the teacher's feedback and the student's appreciation of its practical import has to be reduced or closed. One option could be to teach students the key concepts as a separate activity, maybe outside the context of real assessment events. The alternative outlined below goes back to the aims of higher education, which include having students learn complex concepts and skills relevant to various disciplines, fields and professions. Expressed in basic terms, students can be said to have learned when they can: do or produce on demand something they could not do or produce before; accomplish this independently of particular others (such as a tutor or specific group of other students, but not necessarily in strict isolation from others); and deliver the product at an acceptable level of quality. Implicit in this is the requirement that students be able to recognize different levels of quality in works of the types they are expected to produce. This is necessary for them to be able to monitor the quality of their own work while it is still under development.

The key feature of the alternative approach is to provide students with significant appraisal experience as part of the pedagogical design (Sadler 1989). The aim is to induct them into both explicit and tacit knowledge sufficient to enable them to recognize or judge quality when they see it, and to a considerable degree explain it. To the extent that this is achievable, several additional benefits could ensue. These include the development of evaluative knowledge and skills of the types that are valued in careers after graduation; deeper student engagement with the content and structure of the academic program; and a downplaying of teacher-constructed feedback as a critical element for improved learning.

The teacher as assessor

The typical teacher's experience with making judgments is now outlined as a potentially useful model for thinking about how students can develop their personal competence in appraisal. Regardless of discipline or type of assessment task, teachers typically make hundreds of qualitative judgments routinely each year as a normal part of their academic responsibilities (Sadler 1998). They bring the accumulation of judgments about broadly similar responses in the past, and apply this knowledge to new student works. This exposes teachers to a wide variety of ways in which the students argue, describe, compare, evaluate,

create, analyse, synthesize and solve problems. Those appraisal experiences also give access to students' imaginations and strategies, providing vicarious experience about the challenges of production, and extending and enriching the teachers' own knowledge of potential moves which could be passed on to other students. Observe that the teacher's ideas of how a given work could have been done better results from the interaction between the teacher's existing evaluative and actual student works. Sometimes novel, these ideas may not have occurred to the teacher in the absence of particular student works. Also observe that constructive feedback does not need to depend on assumptions about what the student may have intended.

As teachers provide explanations for their judgments, they invariably make use of criteria, which are constitutive elements of evaluative discourses. Furthermore they routinely invoke whichever criteria are salient to a particular judgment, which means that they are sensitive to which are relevant to it, and which are not. Behind each criterion sits an enduring quality-related concept that students, too, can acquire. Again, analysing teacher experience provides a clue as to how. The teacher's exposure to a variety of student works gives rise to two distinct types of knowledge. The first is exhibited by the range of overall quality of the entire set of works, extended by the range of works the teacher has appraised in the past. This existential range of quality plays a significant role in the formation and maintenance of the teacher's abstract concept of quality itself. It is not unique to responses to any particular assessment task but carries across assessment events. The second type of knowledge is comparability, which is framed by each subset of student works that are judged to be of about the same quality despite their differences. A given level of quality has many potential 'expressions'. This is so much part of the normal experience of appraising multiple works that it is not regarded as remarkable or particularly taxing. These two types of knowledge – quality and comparability – can be represented as two dimensions or axes. These define a two-dimensional space within which each student work can be located at a unique position. In marking a batch of student works, the assessor progressively populates that space with judgments about real cases. From within this rich experiential assessment space, the teacher constructs feedback.

No matter how expertly and conscientiously constructed, it is difficult to imagine how feedback, regardless of its properties, could be expected to carry the burden of being the primary instrument for improvement. The solution proposed here is to provide learners with appraisal experience that is as close in scope and kind to that of the teacher as resources will allow. The main tool is, as might be expected, peer assessment (Boud et al. 2001), for which typologies and various approaches to its implementation are well documented in the literature. However, it is not peer assessment as routine activity or busyness, but purposeful peer assessment that is designed with a clear pedagogical intent. The nature of that intent makes up the focus of the next section.

Concept formation and three key concepts

Knowledge of relevant appraisal terms and concepts is necessary for students to be able to think, plan, develop and monitor high-quality works; to be confident in appraising entire complex works; and to engage in intelligent conversations about quality and its determinants. For many students, understanding the key concepts and their implications for practice are non-trivial; simply being told, even with multiple messages, is rarely effective. As one might expect, a substantial research literature on concept formation exists; it includes, in particular, the seminal work of Vygotsky (1986/1934). For the purposes of this article, three classes of concepts are identified as important in developing appraisal expertise, namely, response genre, quality, and criteria. These make up the explicit agenda to be addressed through structured peer assessment, and the order in which they are listed here is not accidental. A broad hierarchy is implied: the work must conform to the required response genre before its quality can be assessed; the quality is explained or justified by appeal to salient criteria. All three are crucial elements in assessing complex student works across many disciplines. Before looking at them individually, however, it is useful to look at concept formation as a general learning task.

A concept represents phenomena that share certain similarities known as common characteristics or critical attributes and that carry a distinctive label. Concepts are fundamental to human cognition and thinking, language development and verbal communication. They are generally distinguishable from one another within a given context or discourse, although overlap is not ruled out. They provide many of the building blocks for language, enabling effective and economical communication and thought, including reasoning with oneself. Obviously, broad agreement on the critical attributes is necessary for concepts to have currency across users and contexts.

Concepts that are core to academic disciplines are well understood by experts in the respective fields, and help to define the distinctive characteristics of disciplines. Not surprisingly, a great deal of education at all levels is devoted to teaching concepts, many of which are quite complex. The same approaches used in teaching discipline concepts can be applied to teaching assessment concepts. The formation of sophisticated concepts typically requires considerable investment of time, with multiple exposures to cases, inductive reasoning, verbalisation, discourse, and practical experience in making classificatory or evaluative decisions.

Because many concepts have their origins in concrete examples or cases, teaching them to students by referring to the same types of examples in effect walks learners back through the intellectual processes of concept invention. Concrete referents are not just useful and helpful to the learner; they are probably indispensable. Shared language and shared (concrete)

instantiations need to work together in concert (Sadler 1980). The most accessible concrete referents are authentic examples, either real cases or plausible constructions. The person explaining a concept typically switches back and forth effortlessly and unobtrusively between the general (which is the concept as an abstraction) and the specific (which is the case-in-point, example or illustration). Real examples help make the connections with, and illustrate, the verbal material or to serve as the context within which properties and patterns are identified and verbalisation about them takes place. Response genre, quality, and (individual) criteria are typically matters of degree rather than the dichotomous states which characterise many simple concepts that can be demonstrated by multiple positive and negative instances. Concepts based on matters of degree require exposure to wide ranges that exhibit properties to a more or less degree. In that way, examples help to identify and give body to a concept that would otherwise remain abstract. They can then be distinguished from other concepts with which they could potentially be confused. Practicing judgments about previously unseen concrete cases, specifically about works or performances of the same genre or type that learners are expected to produce themselves, providing explanations of those judgments, and debating those explanations, can narrow the distance between the abstract and the concrete.

Response genre as a concept. In higher education, the form and structure of what is expected from a student as an ideal response to an assessment task is generally set by the teacher. If students are asked to produce a critique of, say, a public policy, a literary work or a proposed technological solution to a problem, the responses are expected to qualify as critiques. A critique is a form of response that is distinct from an explanation – because it involves the student in evaluating something and expressing a considered judgment about it. The student has to stand back some distance from the content or issue, weigh up the relevant features of it, examine the evidence or grounds for the appraisal, and then commit the result of deliberations to paper. A critique, therefore, has a definite structure to it, which is why it constitutes a distinct genre or class of production. This does not imply that critiques from different producers will be the same or even much alike. However, they will have a sufficiently similar purpose and structure for them to be recognisable as falling within the genre of critique. An explanation conforms to a different set of characteristics, is arrived at through different cognitive processes, and results in a distinctive form of final product. Other response genres are solution, comparison, proof, demonstration, extrapolation, and scenario.

The term 'response genre' used in this section refers both to the type of response set out in the assessment task specifications within which the teacher expects the students to respond, and to the type of response actually submitted by a student. The reason for raising response genre as an issue is that the works many students submit lie outside the genre specified in the assessment task. Furthermore, students typically seem oblivious of this mismatch. A student

who submits a work that is a combination of synopsis and conjecture is not addressing an assessment task that specifies a critique. Its quality as a critique therefore cannot be appraised. Giving credit for any work that is not a critique potentially has three negative effects: the intent of the assessment task is subverted; the student's action is rewarded and reinforced without requiring a change in thinking and practice; and the student gains credit through avoiding a worthwhile educational outcome.

Admittedly, not all teachers intend assessment task specifications to be taken literally, and many use familiar forms of words out of habit. This is more likely to occur when academics, in devising assessment tasks, focus primarily on the content they wish to test, with only peripheral attention paid to what students should do with that content in order to demonstrate achievement of high-level cognitive objectives. This carries through to marking: work that shows comprehensive coverage may be rewarded highly, without respect to genre. 'Students know a lot about the topic, and that deserves due recognition.' This attitude reflects a lack of awareness among many academics of the significance of response genre as a concept, as does its frequent omission from lists of assessment criteria.

To construct a response within a given genre, the student has to design, select or construct the content, and organise that content so that the final production conforms to enough of the critical attributes to warrant the label. Understanding the characteristics of the different genres that are possibilities as assessment tasks is crucial. Furthermore, the concept of response genre is central to the aims of the assessment process. The purpose of learning is not simply for students to memorize quantities of material and regurgitate it on demand but to have students achieve a level of command over a body of knowledge, and to know how to process and organise it to achieve specified ends and so provide evidence of the attainment of key intellectual outcomes. In many fields, high-level academic learning can be demonstrated by the learner's ability to tackle diverse rather than routinised intellectual and practical problems. Assessing this requires that learners be able to assemble different selections of, and angles on, the substantive content in ways that manifest themselves in a variety of response genres. Explicitly using the term response genre with students can provide the pragmatic basis for signalling it as a pre-emptive requirement. Students generally adapt rapidly to this as a hard constraint once they understand the situation. Furthermore, students who appraise peer responses through the lens of response genre typically discover for themselves that many of those responses do not strictly address the set task. Conforming to the nominated response genre is seen not only as being required by the task specifications but also as defensible. Unless a work is constructed within the designated genre, there is no answer to the question of how well it addresses the assessment task as specified. After students master the idea of a few particular genres, they start to see how the issue of genre is important and become more open to appreciation of the characteristics of other genres.

Quality as a concept. This concept applies in disciplinary and professional contexts where complex learner productions are expected to be non-standardised. The concept of quality is then abstract and typically difficult for students to grasp. By 'quality' is meant the degree to which a work comes together as a whole to achieve its intended purpose. When complex phenomena are being evaluated, quality is often determined configurally rather than as the 'sum' of 'measures' of its components. Such holistic judgments may amount to more, or sometimes less, than would result from formal consideration of the various qualities taken separately. These qualities are, of course, usually called criteria (Sadler 2009b). In practice, quality is often easier to recognise when it presents itself than it is to define in the abstract, or account for fully in the particular. Not uncommonly, something significant is lost when attempts are made to express quality in propositional or declarative form, that is, in words, including rubrics and expansions of fixed criteria.

Determinations of quality often require all-things-considered holistic judgments in which multiple criteria are attended to simultaneously in a configural way. They see beyond superficial differences in the forms of individual works and into the deeper, subtler and more abstract aspects. This is why quality as a concept is treated here as an integrated entity in its own right rather than as somehow 'composed' using criteria. Students need to be exposed to, and gain experience in making judgments about, a variety of works of different quality, and so populate for themselves the two-dimensional appraisal space outlined above. Students' need planned rather than random exposure to exemplars and experience in making judgments about quality. They need to create verbalised rationales and accounts of how a work could have been done better. They need to engage in evaluative conversations with teachers and other students. Together, these provide the means for students to form essentially the same concept of quality as the teacher possesses, and in particular to understand what makes for high quality. Although providing these experiences for students may appear to add more layers to the task of teaching, it is possible to organise peer assessment as a direct approach to higher education pedagogy (Sadler 2009a).

Criteria as concepts. By a criterion is meant a property or characteristic that is useful in the context of quality and quality determinations. Some criteria (such as presentation and word length for a written piece) are straightforward with sharp boundaries. It is relatively easy to tell when something has been complied with or a rule kept. Other criteria, probably the majority, and including some with disarmingly simple labels, are considerably more abstract. This makes them problematic for students until they become competent users of them, which often requires some sophistication or fineness in judgment. These abstract criteria are concepts that do not have sharp boundaries, so they have to become known in the same ways

such concepts are formed by individuals and then shared. Eventually, criteria need to become a regular part of the student's appraisal vocabulary, enabling them to rehearse and to reason through in their own minds as they arrive at judgments, and later to explain and justify them to others. Coherence (in a piece of academic writing) is an example of an abstract criterion. How well do students understand this concept? Can they recognise low and high levels of it in particular works? Do they effectively recognise this property but use different terminology for it (such as 'hangs together')? Can they sense and work towards coherence in their own productions while construction is under way? The crucial test of whether students understand the (assessment) concept of coherence is not whether they are able to define it formally but whether they can make sound judgments about the coherence of their own works and those of others, and whether they use the term appropriately in explaining judgments about quality, in providing feedback for improvement, and in conversations with others.

The same could be said of evidence (for assertions), cogency or rigour (in an argument), artistry (in a musical performance), elegance (of a mathematical solution), integrity (in a clinical interview), efficiency (in a project design), and reliability (in patients' self reports). Students need to understand what these mean and imply for real appraisal decisions. They need to know when particular criteria are appropriate to employ in particular cases. They need to know when to invoke a non-standard, rarely used, criterion because, in a particular case, it is critically important. They need to know when to ignore an ordinarily indispensable criterion because some super-ordinate criterion makes it irrelevant in a particular appraisal. The point being made here is that a deep knowledge of criteria and how to use them properly does not come about through feedback as the primary instructional strategy. Telling can inform and edify only when the meanings and implications of the terms and the structure of the communication are understood by the student as message recipient.

The sleeper issue with respect to criteria as concepts is that the same terms often mean different things to different teachers, and these differences remain largely unexplored. Space does not permit a detailed analysis of this, except to note it and make a few brief comments. Although academics teaching different courses might agree on a fixed list of criteria, vigorous debate can follow an attempt to formalise the interpretations of those same criteria to enable consistent use by all teachers and assessors. Furthermore, research studies into how teachers can best convey their expectations (including the meanings of criteria) tend to focus specifically on qualitative assessment within a course rather than the issue of consistency of interpretation across courses, and this is not seen as problematic. Two recent examples are the studies by Andrade (2007) and Rust et al. (2003) in the **USA** and UK respectively, where the major interest was in how professors make their expectations known to students. Consistency in the interpretation of key terms may not pose a problem for individual higher education academics, but it does for students, who are faced with interpretations that can vary

significantly across teachers or courses. Finally, it might be thought that students need to be familiarised with standards as well as the criteria. The reason for not raising this aspect specifically is that standards can be thought of in relation to contiguous segments of the quality continuum. The more fundamental issue here is for students to grasp the constitutive nature of quality itself.

Tacit knowledge. A learning environment that attends to developing students' conceptual understandings of response genre, quality and criteria will, as a consequence, extend the students' tacit knowledge (Polanyi 1962). It cannot be taught any other way. Making numerous and purposeful peer assessments is the crucible within which these three concepts can be seen to react and interact. This activity gives rise not just to an appraisal judgment but also to a body of unseen, unarticulated and often unheralded know-how of the intricate relationships between the appraisal elements and how they are applied. Competent practitioners and assessors constantly draw on their reservoirs of tacit knowledge; it is the very essence of a great deal of professional expertise. Higher education institutions have a responsibility to induct students not only into the mechanics of appraisal, but also into a deep appreciation of how complex qualitative judgments can be made with integrity.

Conclusion

It may be assumed that the teacher's responsibility for making best use of assessment to improve student learning is fully discharges by providing extensive verbal communications relating to assessment task specifications (and possibly fixed assessment criteria); information about the quality of an appraised work; and advice about how future responses to similar assessment tasks could be improved. These are commonly referred to as feedforward and feedback respectively. However, empirical evidence shows that while these communications may be appreciated by students, they often lead to little if any improvement in subsequent submissions. Because high quality feedback is laborious to compose, and many students seem to take little interest in and benefit from it, teachers can be forgiven for feeling disconcerted or discouraged. That the situation turns out as it does can be traced back to a fundamental flaw in the assumption: feedforward and feedback are essentially limited to telling, and the conditions for learning through being told are rarely satisfied adequately.

In the analysis presented in this article, it has been argued that the problem lies less with the quality of feedback than with the fundamental assumption that telling, even detailed telling, is an effective approach to complex learning. Because feedback is commonly expressed in verbal form, learning from being told is flawed as a general strategy because the conditions for the statements to make intimate connection with the student work (with a view

to future work) are frequently not satisfied. Assuming that low student dispositions to learn is not the reason for their failure to capitalise on learning opportunities, the issue is how to create a different learning environment that works effectively. A proposed alternative to the usual sequence of [task – response – appraisal – feedback] is to make intensive use of purposeful peer assessment as a pedagogical strategy, not just for assessment but for substantive aspects of the course as well. Students need to grasp three groups of concepts in particular – response genre, quality, and criteria – if interactions between teachers and learners are to be formatively effective, and capability in complex appraisal is to be developed. These assessment concepts must be understood not as abstractions but as core concepts that are internalised, operationalised and applied to concrete productions. Unless this occurs, the key assessment concepts are likely to remain submerged and invisible. To the extent that the alternative outlined in this article is put into action, the significance of feedback-as-telling is reduced. If the process were to be entirely successful, the need for feedback would be obviated altogether.

References

Andrade, H., and Y. Du. 2007. Student responses to criteria-referenced self-assessment. *Assessment and Evaluation in Higher Education* 32: 159-181.

Ausubel, D. 1963. *The psychology of meaningful verbal learning*. New York: Grune & Stratton.

Black, P. and D. Wiliam. 1998. Assessment and classroom learning, *Assessment in Education* 5: 7-74.

Bjorkman, M. 1972. Feedforward and feedback as determiners of knowledge and policy: Notes on a neglected issue. *Scandinavian Journal of Psychology* 13: 52-158.

Boud, D., R. Cohen, and J. Sampson. 2001. *Peer learning in higher education: Learning from & with each other.* London: Routledge

Carroll, J.B. 1968. On learning from being told. *Educational Psychologist* 5: 4-10.

Crooks, T.J. 1988. The impact of classroom evaluation practices on students. *Review of Educational Research* 58: 438-481.

Polanyi, M. 1962. *Personal knowledge: Towards a post-critical philosophy*. London: Routledge & Kegan Paul.

Higgins, R., P. Hartley, and A. Skelton. 2002. The conscientious consumer: Reconsidering the role of assessment feedback in student learning. *Studies in Higher Education* 27: 53-64.

Rust, C., M. Price, and B. O'Donovan. 2003. Improving students' learning by developing their understanding of assessment criteria and processes. *Assessment & Evaluation in Higher Education* 28: 147-164.

Sadler, D.R. 1980. Conveying the findings of evaluative inquiry. *Educational Evaluation and Policy Analysis* 2(2): 53-57.

-------. 1989. Formative assessment and the design of instructional systems. *Instructional Science* 18: 119-144.

-------. 1998. Formative assessment: Revisiting the territory. *Assessment in Education: Principles, Policy & Practice* 5: 77-84.

-------. 2009a. Transforming holistic assessment and grading into a vehicle for complex learning. In *Assessment, learning and judgement in higher education,* ed. G. Joughin, 49-64. Dordrecht: Springer.

-------. 2009b. Indeterminacy in the use of preset criteria for assessment and grading in higher education. *Assessment and Evaluation in Higher Education* 34: 159-179.

Vygotsky, L.S. 1986. *Thought and language*. Revised and edited by Kozulin, A. Cambridge, Mass: MIT Press. (Originally published in Russian as *Myshlenie I rech*, (1934).