Beyond model answers: learners' perceptions of self-assessment materials in e-learning applications

Karen Handley[a]* and Benita Cox[b]

[a]*Oxford Brookes University, UK;* [b]*Imperial College London, UK*

The importance of feedback as an aid to self-assessment is widely acknowledged. A common form of feedback that is used widely in e-learning is the use of model answers. However, model answers are deficient in many respects. In particular, the notion of a 'model' answer implies the existence of a single correct answer applicable across multiple contexts with no scope for permissible variation. This reductive assumption is rarely the case with complex problems that are supposed to test students' higher-order learning. Nevertheless, the challenge remains of how to support students as they assess their own performance using model answers and other forms of non-verificational 'feedback'. To explore this challenge, the research investigated a management development e-learning application and investigated the effectiveness of model answers that followed problem-based questions. The research was exploratory, using semi-structured interviews with 29 adult learners employed in a global organisation. Given interviewees' generally negative perceptions of the model-answers, they were asked to describe their ideal form of self-assessment materials, and to evaluate nine alternative designs. The results suggest that, as support for higher-order learning, self-assessment materials that merely present an idealised model answer are inadequate. As alternatives, learners preferred materials that helped them understand what behaviours to avoid (and not just 'do'), how to think through the problem (i.e. critical thinking skills), and the key issues that provide a framework for thinking. These findings have broader relevance within higher education, particularly in postgraduate programmes for business students where the importance of prior business experience is emphasised and the profile of students is similar to that of the participants in this research.

Introduction

Formative feedback is a vital part of learners' efforts to apply and practice the principles that they learn: as Laurillard (1993, p. 61) has said, 'action without feedback is completely unproductive for a learner'. Feedback is usually interpreted as material that assesses, verifies and comments on the learner's response to assignment questions. In

---

*Business School, Oxford Brookes University, Wheatley, Oxford, OX33 1HX, UK. Email: khandley@ brookes.ac.uk

some educational settings, however, the possibilities for giving verificational feedback, followed by adaptive comments on errors or areas for improvement, is limited. This is the case both in self-paced e-learning (traditionally known as computer-aided instruction [CAI]) and also occasionally in large-scale lectures where lecturers may resort to rhetorical question-and-answer sessions addressed to the entire audience.

In e-learning, the desire to provide verification has tended to produce one of two approaches: either a proliferation of multiple choice questions (MCQs) amenable to computer-aided assessment; or the transfer of 'complex' forms of assessment to electronic discussion forums where academic staff can respond to individual postings with the benefit of their expertise and understanding. Increasingly, these approaches are combined as newer pedagogic technologies become integrated with the old, and approaches are blended to balance the advantages of one with the disadvantages of others. Furthermore, recent technologies such as blogging and wikis have expanded the repertoire of communication tools to promote student reflection and knowledge integration (for example, Downes, 2004; Fountain, 2005). Nevertheless, there continues to be a reliance on MCQ quizzes (despite pedagogic limitations) and on staff-moderated discussion forums (despite resource constraints).

With respect to MCQs and similar types of binary exercises, Schank and Cleary (1995) have derided the over-use of this method of assessment because it risks 'dumbing down' students' learning experience. The reliability and validity of MCQs have also been questioned (Davies, 2002; Conole & Warburton, 2005). Validity is a particular problem when assessing students' higher-order learning that has traditionally relied on essays or short-answer questions (Rowntree, 1987). However, the limitations of artificial intelligence technologies such as natural language processing means that computational assessment of free-text continues to be problematic (Gardner, 1987).

The second approach, which we have broadly categorised as the use of discussion forums, is not without its own problems. The principal one concerns resourcing: staff (and indeed students themselves; Moore, 2002) often feel they do not have the time to respond carefully and individually to each posting because of other demands on their time. This means that students may be left without feedback—a 'black hole experience' in cyberspace (Suler, 1997).

The twin problems of limited staff resources to moderate discussion forums, and the deficiencies of MCQs and artificial intelligence technologies, continue to create challenges for staff engaged in designing materials for students. A conventional response to these challenges has been to offer students 'model answers' to enable them to self-assess their own work (for example, Laurillard, 1998) rather than give adaptive feedback that provides an external assessment. The proposed benefits of self-assessment are that students are encouraged and empowered to develop their own critical faculties rather than rely on the authoritative judgement of external sources such as academic staff (Boud, 1995). However, model answers are deficient in many ways, particular in their presumption of a single correct answer regardless of context, aims, personal characteristics, and so on. This reductive assumption is rarely the case with complex problems designed to test students' higher-order learning.

Nevertheless the challenge remains of how to develop students if the conventional form of verificational, adaptive feedback cannot be computer generated.

To address this challenge, the research presented in this article investigates the effectiveness (and deficiencies) of model-answer feedback, and identifies alternative materials that can facilitate students' self-assessment. The research was exploratory, using semi-structured interviews with 29 adult learners who used a seven-hour CAI application to help develop their project management skills. The application included tutorials and a number of activities, which combined open questions with model answers as 'feedback'. The model answers did not provide verification in the form of '*correct*' or '*incorrect*'. Dissatisfaction with model answers quickly became apparent during early interviews, and so the research re-focused its efforts on two questions: what made feedback 'useful' (whether verificational or not), and what alternatives to model-answer feedback could be provided.

Literature review

The importance of formative feedback is widely acknowledged as a critical input to the process of learning (Butler & Winne, 1995; Askew, 2000; Laurillard, 2002; Black & Wiliam, 2003; Taras, 2003; Roos & Hamilton, 2005). However, educational theory differs in its articulation of the function of adaptive feedback and self-assessment materials such as model answers. In her review of contemporary feedback paradigms, Askew (2000, pp. 3–15) distinguishes between three paradigms: receptive-transmission, constructivist, and co-constructivist. In the receptive-transmission paradigm, the 'expert' gives corrective or explanatory information to students to help them learn 'the truth'. In the constructivist paradigm, experts engage in an 'expanded discourse' to help students gain new understandings, but cannot dictate what those understandings will be since students are influenced by a number of sources. In the co-constructivist paradigm, the dichotomy of expert and novice dissipates and there is recognition that the 'teacher' also learns from the 'student' through dialogue and participation in shared practices (Lave & Wenger, 1991). In an older literature on 'programmed instruction' inspired by behaviourists such as Skinner (1958) and Thorndike (1932), feedback was viewed as a reinforcement mechanism: either rewarding desirable behaviour, or 'punishing' or withdrawing rewards if the student made errors. Emphasis was given to behaviour that could be observed and possibly shaped through a schedule of feedback interventions.

The cognitive science revolution of the 1970s and early 1980s firmly swept aside the behaviourist paradigm and refocused attention on the mind and its purported parallels with computational and symbolic processes. Inspired by these new ideas, the e-learning literature on intelligent tutoring systems and CAI became firmly grounded in the receptive-transmission feedback theories (for example, Wenger, 1987). As a consequence, intelligent tutoring systems and CAI designers tended—and to some extent still do—to espouse the MCQ and its binary equivalents because they could be used to 'diagnose' and then 'correct' students' faulty knowledge. This assumption is embedded in seminal feedback models such as that of Kulhavy and Stock (1989) and

Bangert-Drowns *et al.* (1991). The problem with 'model' answers as a form of self-assessment is that they, too, imply a receptive-transmission pedagogy that is considerably out of favour in current academic debates.

In the context of self-paced learning, a more appropriate way to conceptualise e-learning feedback and self-assessment materials may therefore be one that acknowledges the role of self-regulated learning, and that then investigates the type of self-assessment material that facilitate that process. In this regard, a classic model of self-regulated learning that can be applied to e-learning is that of Butler and Winne (1995). They argue that although learners tend to seek out external sources of feedback (e.g. tutor comments), the most effective students also generate internal feedback by monitoring their performance against self-generated or given criteria. It is in this connection that self-assessment becomes particularly important.

The pedagogic benefits of self-assessment have been widely discussed. Brown and Dove (1991, p. 59), for example, cite the advantages of student motivation, exchange of ideas, transferable personal skills and the development of a community of learning. However, self-assessment comes with its own cautions and challenges; for example, it can be demanding and time-consuming, and students may lack the capability to understand and apply assessment criteria (Brown & Dove, 1991, p. 61). Self-assessment 'in its more rigorous form' involves learners not only in evaluating their work, but also in identifying and specifying the criteria and standards which should be applied to it (Race, 1991, p. 7). Here we see the developmental potential of self-assessment methods, which is 'more transformational, elusive and confronting to conventional teaching than it is normally expedient to recognise' (Boud, 1995, p. 1). This is because the process of identifying criteria—especially if these are debated openly with tutors—acknowledges the socio-constructivist process of scaffolding student learning to broadly fit the norms of the educational community to which they belong (see also Vygotsky, 1978).

The experience of many students using stand-alone CAI, WebCT quizzes, and so on (albeit as one among many elements of learning and teaching) is that the predominant forms of activity are the closed question plus MCQ, or the open question plus model answer. However, given the deficiencies of MCQs and model answers, which have been discussed above, the need for alternative ways to facilitate student self-assessment is particularly important.

Research study aims and objectives

In view of the relative lack of prior empirical research on alternatives to model answers following students' short-answer responses, this research aimed to investigate learners' perspectives on what constitutes effective non-verificational e-feedback. To address this aim, three research objectives were identified:

1. Obtain learners' perceptions of examples of model answer e-feedback taken from an e-learning application they have recently completed.

2.   Solicit their views on what constitutes 'useful' e-feedback.
3.   Solicit their perceptions and evaluations of alternatives to model-answer feedback.

Research methods

The research adopted an exploratory approach relying mainly on semi-structured interviews, asking learners about specific examples of feedback as well as hypothetical examples. The research participants were selected from a global consulting organisation, 'ProfCo'. ProfCo was selected because it produced high-quality, innovative CD-ROM applications for its employees.

It was not feasible (nor an intention) to make statistical generalisations covering the whole population in ProfCo, because interviews were in depth and were with a smaller ($n = 29$) number of people than would be the case in large-scale hypothetico-deductive research. Nevertheless, the risk of systematic bias was reduced by adopting a time-based strategy (Sapsford, 1999) whereby all learners based in the UK and US offices who had completed *Zentoria* within three weeks prior to requesting an interview were selected for interview. The 29 interviewees were then stratified according to their level of expertise, because prior research (Chi *et al.*, 1988) indicates that expertise is an important mediator of learner expectations and therefore also of their evaluations of educational materials. Level of expertise was attributed to learners using the proxy of their managerial 'grade', where one represented 'new entrant' and seven represented 'partner'. New entrants were typically in their early 20s while consultants were often in the early 30s. Thus the profile of younger research participants resembles that of postgraduate Masters-level programmes (e.g. M.Sc. in Management), while the profile of consultants resembles that of Executive Masters and short-course programmes, which are a growing area of higher education.

A mixed methodology was used, and methods were selected for their appropriateness to each of the three objectives. Given the small sample size, no statistical generalisations have been made; nevertheless, we present some numerical data for informational purposes, recognising that the data are indicative only.

Obtaining learners' perceptions of specific examples of model answer e-feedback

The first objective was to ascertain learners' perceptions of 'real' model-answer feedback, based on their prior experience of an authentic e-learning application. The application selected for this purpose was *Zentoria*, which was designed specifically for the participating organisation with the aim of developing learners' project management skills. The storyline running through *Zentoria* was of a project team of management consultants helping their client to develop an environmentally friendly car. The application was created using Macromedia Director and Adobe Photoshop by a team of information technology developers and educational designers. The seven-hour course could be taken at the learner's own pace, and was structured around two elements: *tutorials*, which provided a grounding in project management principles; and *activities* (MCQs and open questions), which portrayed problems typical of project

management. The application included a 'Scratchpad'—a notebook facility in which learners could type their short-answer responses to open questions. On finishing their responses, learners were presented with model answers that varied in several ways.

The 29 interviewees were asked about their experiences of completing three open-question activities. These activities (scenario > task > type of 'feedback') are summarised in Table 1. Semi-structured interviews were conducted to elicit interviewees' views on the activities, and particularly on the model-answer e-feedback. Interviews were analysed using 'meaning interpretation' and 'meaning categorisation' approaches (Kvale, 1996), as well as open-coding, constant comparative methods, deviant case analysis, tabulations and a modified form of inter-rater cross-checks. Analysis was facilitated through the use of Microsoft Excel, and NVivo.[1]

The qualitative data analysis package NVivo was used to collate and then analyse comments from interviewees. Analysis was facilitated by developing a 'coding template' to categorise interviewees' comments according to theme. Some codes were identified from the theoretical literature (e.g. the importance of 'student engagement'), but most were developed inductively by reading and re-reading the transcripts, or by listening to the audio-tapes. Selected extracts were discussed with two colleagues engaged in similar research in order to clarify and improve consistency of interpretations. This strategy provided a balance between seeking full inter-rate reliability (which is difficult if not undesirable in qualitative research because of the 'tyranny of the lowest common denominator'; Kvale, 1996, p. 181), and reliance on one person's subjective interpretations.

Microsoft Excel was used to analyse numerical data. This was done in the final piece of research where learners were asked to comment on nine self-assessment designs. Comments were transformed into a rating scale of −2 to +2 according to the strength and direction (favourable or not towards the design) of those comments (Boyatzis, 1998).

Soliciting learners' views on what constitutes 'useful' e-feedback

As will be discussed later, many learners expressed dissatisfaction with the model answer e-feedback in *Zentoria*. From the perspective of this research, a different

Table 1.   Summary of activities used to ascertain learners' perceptions of model answers in *Zentoria*

Activity reference	Scenario	Task	Feedback
ZEN1	Unauthorised extension to project scope	Open question	Model answer
ZEN2	Loss of confidence by project director in the calibre of the project team	Open question, accompanied by a finished example of a similar open question	Model answer
ZEN3	Preparation for the first meeting of a new project team	Open question	Presentation of some of the issues to consider

interviewing strategy was required if alternative feedback designs were to be identified. Therefore the interview schedule was supplemented with new questions, asking interviewees what made feedback 'useful', and what they perceived was generally 'missing' from the *Zentoria* feedback. These qualitative data were analysed using the methods described above.

Soliciting learners' evaluations of alternative non-verificational e-feedback designs

The third research objective was developed once it became apparent that learners were dissatisfied with model answers. In addition to asking for their general opinions about 'useful' feedback, the remaining interviewees were presented with further material in the form of descriptions of nine hypothetical feedback designs that could follow short-answer open questions (Table 2). The designs were presented to interviewees for evaluation and comment. Qualitative responses were transformed into a rating scale of −2 to +2 as described above. Where no comments were made, a zero ranking was recorded.

Results

Learners' perceptions of specific examples of model answer e-feedback

When asked about the model-answer feedback, many learners recalled their disappointment about the lack of assessment and verification. Most, however, also recognised that automated verificational feedback on free-text answers was an unrealistic expectation. Over and above these generally shared sentiments, learners varied somewhat in the way they responded to the model answers: as a new piece of instruction, as reinforcement for existing knowledge, or as a prompt to 'tune' existing mental schemata.

Table 2.   Nine hypothetical non-verificational feedback designs evaluated by interviewees

Short-hand name	Description given to interviewees
MA	Model answer
KD&D	Key dos and don'ts
Exmples	Several real-life examples, each describing a project situation, the intervention and the outcome
BP	Description of the formal, best practice intervention
Issues	Explanation of the underlying issues behind the problem
Exp. thnk proc	Description of how an expert would think through the problem
Consequences	Video showing the consequences of several problem interventions
Exp. P&C	Video showing an expert comparing the pros and cons of several problem interventions
Peer debate	Video showing peers in an informal setting debating the merits of several problem interventions

e-Feedback as a new piece of instruction.   Many learners stopped typing free-text answers when they realised that the e-feedback in *Zentoria* was non-verificational. Instead, they read the e-feedback as a new piece of instructional text.

e-Feedback as reinforcement.   For some interviewees, non-verificational feedback was useful in re-enforcing their views. For example:

> … it helped to crystallise some ideas that weren't fully formed in my mind. (Z8)

e-Feedback as a prompt for reflection and self-assessment.   For a few interviewees, and mainly those already experienced in project management, the model answers acted as a prompt for further reflection and self-assessment. By thinking through the solutions and explanations given in the model answer, learners felt that they 'tuned' and improved their existing knowledge schemata (Rumelhart & Norman, 1978). The feedback helped them to achieve a sense of closure to the questions raised in their minds by doing the task, while at the same time validating and modifying their current ways of thinking about the type of problem represented.

However, only a few learners experienced this sense of reflection and closure. An important differentiating factor seemed to be their tolerance of the ambiguity of the feedback. Some interviewees were tolerant of feedback that did not verify their own answers. These learners seemed satisfied with feedback that offered 'food for thought'—new ideas, perspectives and possibilities that allowed them to assess the feasibility and applicability of their own 'solution' to the problem task and that added richness to their existing ways of thinking about the problem. Although the model answers gave no 'feed back' in the true cybernetic sense, they did promote self-generated learning.

Learners' views about what constitutes useful feedback

Having asked learners about three specific instances of model-answer feedback, they were asked two supplementary questions and invited to talk about their generic expectations of e-learning feedback following MCQs or short-answer questions. The first question was 'what makes feedback useful?' This was asked in recognition of the vital importance of understanding student *expectations* of e-learning feedback; indeed, Alavi and Gallupe (2003) argue that the management of student expectations is a key principle of technology-mediated learning. In response to the question, interviewees commented mainly on the instructional function of feedback—what they wanted it to 'do'. The responses were analysed and grouped into six categories, which are presented in Table 3 using the terminology of instructional design, enriched by inter-viewees' own comments. The prevalence of each category of comment is given in the final column.

Many interviewees cited several functions of e-feedback, and thus the total for comments exceeds the number of interviewees. Comments indicate that e-feedback

Table 3.   Illustrations of interviewees' expectations of 'useful' feedback

Instructional function of feedback	Illustration of interviewees' expectations	Cited by interviewees (*n*)
Verify	'… confirms what you in your own mind mentally have come up with' (Z6)	34
Reinforce	'[Feedback] is very useful for confirming your ideas, and helping you -, you know if you think of something on your own and then it gets confirmed, it helps you remember that point' (Z7)	26
Correct errors, or state correct behaviours	'… straight answers … like "do this in this situation", for example. In the real world of course it may not turn out that way but maybe [it'd be nice to have] a classic case like "this is what you would do". That'd be nice so that we don't have to figure it out on our own' (Z17)	21
Inform	'… contains all the points. It's more of a checklist' (Z18)	10
Explain	'… if it's possible to get the "why" in there. If you're not right, are you on the right track? or … or "gee that's a common mistake, but …"' (Z20)	31
Promote reflection	'… not rehash the things that are basic, but really to give a more stretching analysis' (Z19)	14

was more frequently perceived as useful for verifying and reinforcing existing knowledge or explaining concepts, rather than facilitating more radical knowledge (re)construction by offering alternative insights, reflections and perspectives. However, a different picture emerged when interviewees were asked what was 'missing' from the model-answer feedback they received in *Zentoria*, and what they would have gained from a (human) tutor. A wide range of responses were given, including 'what-if scenarios', 'behaviours to avoid', 'sense of realism', 'alternative perspectives' and 'opportunity for debate'.

An analysis of results that address the first two objectives of the research study suggests that, overall, learners held two key expectations of e-feedback, which at first sight seem contradictory: firstly, the expectation for more verificational feedback (implying a 'right' answer); and secondly, the expectation for more discursive and relativistic feedback. The latter expectation was indicated by comments recommending a move away from a theoretically 'model' answer, towards feedback that exemplified different ways of applying theory to practice. Interviewees' desire for alternative perspectives and 'what-ifs' suggested an expectation for less dualistic, causal-explanatory feedback. By comparison, the primary expectation—for verificational feedback—seemed somewhat of a paradox. This apparent paradox becomes understandable if one considers the mediating variable of 'experience'; for example, that less experienced students seemed most likely to want to know the behaviours they should *avoid*.

Learners' evaluations of alternative non-verificational feedback designs

To probe the apparent contradictions in the results, 26 of the 29 interviewees were asked to comment on nine hypothetical non-verificational e-feedback designs that were developed following analysis of early interviews. Interviewees were asked how effective the designs would be as a follow-on from learners' short-answer responses to open questions. The nine designs are described in Table 2. Interviewees' extrapolated ratings are summarised in Figure 1, which show that, overall, the highest average rating was given to materials giving 'several real-life examples'. The lowest average rating was given to 'model answer'. Given the exploratory nature of this part of the research, numerical analysis does not generate statistical generalisations; nevertheless, we provide some quantitative indicators where this may help the reader gauge the range of responses (Table 4). We also present a summary of interviewees' comments for each design.

'*Real-life examples*' had the highest overall ranking of the nine designs, with learners preferring 'real-life examples' over feedback, which was 'too textbook-like' (Z22). Learners wanted 'authenticity' (see also Leung, 2003). However, some interviewees (e.g. Z26) tempered their positive comments with qualifying remarks.

> With the real-life examples, they're kind of a fairy tale and it's oftentimes hard to give enough background information … to present a complex issue in the time available. (Z26)

An '*expert's process of thinking through the problem*' was ranked highly by most interviewees, but was ranked only sixth by grade-two junior consultants. This result is not surprising if one considers that an emphasis on process and on reasoning skills is
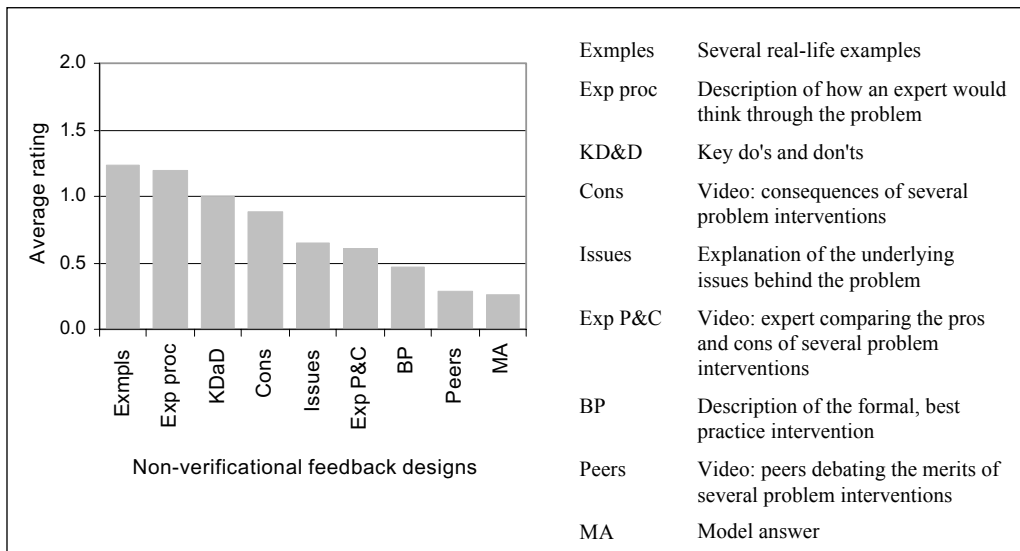

Exmples	Several real-life examples
Exp proc	Description of how an expert would think through the problem
KD&D	Key do's and don'ts
Cons	Video: consequences of several problem interventions
Issues	Explanation of the underlying issues behind the problem
Exp P&C	Video: expert comparing the pros and cons of several problem interventions
BP	Description of the formal, best practice intervention
Peers	Video: peers debating the merits of several problem interventions
MA	Model answer

Figure 1.   Averaged and ranked interviewee ratings for hypothetical non-verificational feedback designs (n = 26)

Table 4.   Description, average ratings and overall ranking of nine non-verificational feedback designs

Reference	Description	Average rating between −2 and +2 (total sample $n = 26$)	Overall ranking
Exmpls	Several real-life examples, each describing a project situation, the intervention and the outcome	1.24	First
Exp proc	Description of how an expert would think through the problem	1.19	Second
KD&D	Key dos and don'ts	1.00	Third
Cons	Video showing the consequences of several possible interventions	0.88	Fourth
Issues	Explanation of the underlying issues behind the problem	0.65	Fifth
Exp P&C	Video showing an expert comparing the pros and cons of several possible interventions	0.62	Sixth
BP	Description of the formal, best practice intervention	0.46	Seventh
Peers	Video showing peers in an informal setting debating the merits of several possible interventions	0.29	Eighth
MA	Model answer	0.26	Ninth

embedded in the ethos of many management consultancy firms. Many firms subscribe to what Schein (1987) referred to as 'process consulting': they apply standard methodologies to evaluate the effectiveness of processes in client operations, and generally provide solutions that impact at the level of those client processes. Findings indicate that experts were admired because they were credited with understanding these processes. However, inexperienced interviewees, perhaps because they had not yet been socialised into the organisation and had not yet adopted a processual approach to problems, ranked the 'expert-thinking process' feedback only sixth.

'*Key dos and don'ts*' were perceived as useful by experienced as well as inexperienced learners. An advantage of this type of feedback is that it explicitly warns of behaviours to avoid.

> The key dos and don'ts would be very good because then you *know* what is wrong, and those types of things stay with you. (Z23, emphasis in original speech)

'*Consequences*' feedback was highly valued by the more experienced interviewees (grade five, ranked third; grade four, ranked third) and modestly valued by the less experienced (grade three, ranked fifth; grade two, ranked fourth). This type of feedback seemed to fulfil two main functions. The first was to promote social learning (Bandura, 1986) through vicarious experience. There was also recognition that, unless well done, consequence feedback could be 'cheesy' (Z7) (if using video) or

'something to laugh about' (Z15). Several interviewees suggested that the value of consequence feedback would be enhanced if supplemented by comments from an expert.

'*Issues*' underlying the given problem, provided by an expert, was ranked fifth over-all. For inexperienced learners it was ranked last, for example, because they saw 'issues' as being too dependent on the circumstances of the specific problem. Interviewees with more knowledge and experience were more accommodating, and saw issue-based feedback as a useful prompt for reflection and for understanding the rationale behind decisions.

It is possible that interviewees who were positive about 'issues' feedback recognised that issues tend to be repeated across multiple problems. They become a way of indexing experiences of problems and their solutions, and of recalling them when analogous problems arise. Issues-based thinking is one example of a more structured, 'critical' method of thinking that students develop during their time at university.

'*Experts discussing pros and cons*' was ranked low by most interviewees. This was surprising if compared with the other 'expert' feedback, which related to the process of solving problems. This discrepancy was possibly because although experts were credited with using effective *processes* for thinking through problems, the culmination of their thinking process—the *solution*—was not credited with being necessarily correct.

'*Best practice*' scored a low average rating, for reasons similar to those given for the model answer. For some interviewees, however, best practice feedback was welcomed precisely because it established the organisation's 'policy' against which actual practice could be benchmarked.

For some interviewees, '*peer debate*' was perceived as useful because peers would advocate new ideas and perspectives that could be reflected on. However, several interviewees suggested that peer-debate feedback would create confusion unless 'controlled or even scripted to be sure to get the appropriate points across'. This suggests that interviewees were uncomfortable having to deal with divergent opinions.

'*Model-answer feedback*' was ranked the lowest of the nine designs. Interviewees implied that model answers were inappropriate for complex questions or topics: 'Getting the perfect answer is not always great, because it's not a perfect world' (Z16).

Discussion

Analysis shows that this group of learners were generally unsatisfied with model answers as a response to their short-answer responses, and that only the more reflective learners used them for self-assessment. As alternatives to model answers, interviewees gave a number of suggestions about the sort of responses they would prefer in the context of a stand-alone element of a course; for example, learning what to 'avoid' and not just what to 'do'; learning about *process* (e.g. how to think through a problem) and not just *action* (e.g. what to do in a specific situation); and learning by hearing of other people's experiences and not just from theory. While the importance of these situated aspects are recognised in the educational literature, the principles are rarely applied to the design of e-feedback in CAI applications.

What this group of learners seemed to desire at the end of each exercise was a sense of 'closure'. A definitive single, correct answer was not necessarily required (see also Kruglanski, 1989) nor deemed appropriate if learners recognised the complex and situated nature of the problem. In those circumstances, a programmed response that addressed the task in terms of relevant issues, alternative perspectives, or the *process* of resolving the problem were sometimes considered more valuable because they prompted reflection leading to a deeper understanding of the problem. However, the number of learners who adopted this perspective was relatively few. Those who *did* tended to have sufficient expertise and self-confidence to rely on their own judgements about whether their response was 'right'. They would then reflect on what else they could learn from the model-answer feedback. Inexperienced learners, on the other hand, wanted first and foremost to know whether they were 'right', because they did not have the judgemental maturity with which to make a self-assessment. Here we see parallels with the work of Miller and Parlett (1974) drawing on the research by Perry (1970) on the intellectual development of students. These authors identified student transitions from a 'cue-deaf' (dualist) orientation in which they believed there to be a body of 'right' knowledge to be accumulated in order to pass examinations, through 'cue-consciousness' (multiplicitous/relativist) and leading (for some) to a 'cue-seeking' (committed) orientation. *Zentoria* students who were cue-seeking recognised the relativity of knowledge and of the tasks they were given to solve, but were also secure in their own values and self-knowledge, which meant they sometimes disagreed with a solution imposed in a model answer. Indeed, one could argue that the advanced students were not only questioning the validity of the model answers, but were also questioning the appropriateness of the questions themselves. In other words, their suggestions implied a rejection of questions of the type 'what would you do?' and their replacement by more nuanced questions such as 'how would you think through the options?' or 'what are the key issues?'. Going forward, this indicates a possible progression from involving students in self-assessment via criteria development to their involvement via question development. This is an important area for further research.

Bearing in mind the range of student experiences, there may be some benefit in developing a multi-layered exercise that seeks to promote reflection in those students who are ready for it, while also providing a verificational element for those needing the security of benchmark answers. The proposed exercise is depicted in Figure 2 and includes a divergent-reflective cycle as well as a convergent-verificational cycle, leading, ideally, to the students' sense of closure.

Another alterative would be to exploit the newer technologies that facilitate conversation between students as they collaboratively debate the assignment questions and devise a repertoire of solutions for them. One possibility is the use of wiki technologies that enable the co-production and co-editing of a single text (Fountain, 2005). The limitation, however, is that wikis may inhibit inexperienced or shy students who are struggling to develop their own voice and authority (Barton, 2004). An alternative, therefore, would be the use of blogging technologies to encourage individuals to engage in conversations with each other while maintaining their sense of identity (for
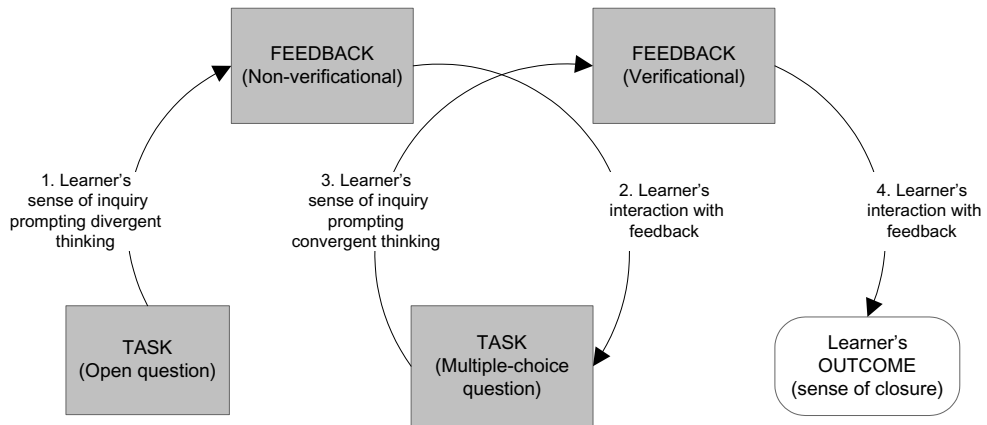
Figure 2.   Proposed structure for divergent:convergent activity

example, Downes, 2004). These are important areas for future development and research.

Conclusion

This research has investigated learners' perceptions of model-answer e-feedback, including their evaluation of alternatives. Alternatives to the conventional model-answer format need to be developed if e-learning is to move away from an over-reliance on MCQs and model answers as the standard method of assessment, particularly where students are expected to meet higher-order learning objectives.

Using semi-structured interviews, the research found that learners are generally dissatisfied with model-answer e-feedback for a variety of reasons, which vary according to their level of experience and prior knowledge in the subject taught. For complex topics requiring judgement and experience, many of the learners interviewed felt that model answers were inappropriate because they failed to indicate the nuances of the given problem. Instead, alternative non-verificational feedback designs were considered more effective. These included 'examples from real-life', and 'issues' to help students structure their thinking of the problem.

The findings of this research have broader relevance within higher education, particularly in postgraduate programmes for business students where the importance of prior business experience is emphasised (Mintzberg, 2004), and where the profile of students is similar to that of the participants in this research. These students need to be supported in their efforts to build on existing knowledge and experiences, and to reconsider what they perceive to be the relevant conceptual and practical issues. The topic of project management, which was the basis of the application used for this research, is an ideal vehicle for achieving this because it encourages an integrative approach to business studies that combines theory and practice.

Note

1. Nvivo is a qualitative data analysis package owned by QSR International (http://www.qsrinternational.com/).

References

Alavi, M. & Gallupe, R. (2003) Using information technology in learning: case studies in business and management education programs, *Academy of Management Learning and Education,* 2(2), 139–153.

Askew, S. (Ed.) (2000) *Feedback for learning* (London, Routledge).

Bandura, A. (1986) *Social foundations of thought and action: a social cognitive theory* (Englewood Cliffs, NJ, Prentice-Hall).

Bangert-Drowns, R. L., Kulik, C. L., Kulik, J. A. & Morgan, M. (1991) The instructional effect of feedback in test-like events, *Review of Educational Research,* 61(2), 213–237.

Barton, M. (2004) *The future of rational–critical debate in online public spheres.* Available online at: http://collegewriting.us/barton/Shared%20Documents/RevisedArticle.doc (accessed 17 July 2006)

Black, P. & Wiliam, D. (2003) 'In praise of educational research': formative assessment, *British Educational Research Journal,* 29(5), 623–637

Boud, D. (1995) *Enhancing learning through self assessment* (London, Kogan Page).

Boyatzis, R. E. (1998) , *Transforming qualitative data: thematic analysis and code development* (Thousand Oaks, CA, Sage Publications).

Brown, S. & Dove, P. (1991) Opening mouths to change feet: some views on self and peer assessment, in: S. Brown & P. Dove (Eds) *Self and peer assessment, Standing Conference on Educational Development,* Birmingham, SCED Paper 63.

Butler, L. B. & Winne, P. H. (1995) Feedback and self-regulated learning: a theoretical Synthesis, *Review of Educational Research,* 65(3), 245–281.

Chi, M. T. H., Glaser, R. & Farr, M. J. (Eds) (1988) *The nature of expertise* (Hillsdale, NJ, Lawrence Erlbaum Associates).

Conole, G. & Warburton, B. (2005) A review of computer-assisted assessment, *ALT-J Research in Learning Technology,* 13(1), 17–31.

Davies, P. (2002) There's no confidence in multiple-choice testing, in: *Proceedings of the 6th International CAA Conference,* Loughborough University, 4–5 July, pp. 117–130. Available online at: http://www.caaconference.com/pastConferences/2002/proceedings/davies_p1.pdf (accessed 11 August 2006).

Downes, S. (2004) Educational blogging, *EDUCAUSE Review,* 39(5), 14–26.

Fountain, R. (2005) *Wiki pedagogy.* Available online at: http://www.profetic.org:16080/dossiers/rubrique.php3?id_rubrique=110 (accessed 20 July 2006).

Gardner, H. (1987) Epilogue, in: *The mind's new science: a history of the cognitive revolution* (New York, BasicBooks) (Original text published 1985).

Kruglanski, A. W. (1989) *Lay epistemics and human knowledge: cognitive and motivational bases* (New York, Plenum Press).

Kulhavy, R. W. & Stock, W. A. (1989) Feedback in written instruction—the place of response certitude, *Educational Psychology Review,* 1(4), 279–308.

Kvale, S. (1996) *Interviews: an introduction to qualitative research interviewing* (Thousand Oaks, CA, Sage Publications).

Laurillard, D. (1993) *Rethinking university teaching: a framework for the effective use of educational technology* (1st edn) (New York, Routledge).

Laurillard, D. (1998) Multimedia and the learner's experience of narrative, *Computers & Education,* 31, 229–242.

Laurillard, D. (2002) *Rethinking university teaching: a framework for the effective use of educational technology* (2nd edn) (New York, Routledge).

Lave, J. & Wenger, E. (1991) *Situated learning: legitimate peripheral participation* (Cambridge, Cambridge University Press).

Leung, A. C. K. (2003) Contextual issues in the construction of computer-based learning programs, *Journal of Computer Assisted Learning,* 19, 501–516.

Miller, C. M. L. & Parlett, M. (1974) *Up to the mark: a study of the examination game,* Monograph 21 (Guildford, Society for Research in Higher Education).

Mintzberg, H. (2004) *Managers not MBAs: a hard look at the soft practice of managing and management development* (London, FT/Prentice Hall).

Moore, M. G. (2002) What does research say about the learners using computer-mediated communication in distance learning?, *The American Journal of Distance Education,* 16(2), 61–64.

Perry, W. G. (1970) *Forms of intellectual and ethical development in the college years* (New York, Holt, Rinehart & Winston).

Race, P. (1991) Learning through assessment, in: S. Brown & P. Dove (Eds) *Self and peer assessment, Standing Conference on Educational Development,* Birmingham, SCED Paper 63.

Roos, B. & Hamilton, D. (2005) Formative assessment: a cybernetic viewpoint, *Assessment in Education: Principles, Policy and Practice,* 12(1), 7–20.

Rowntree, D. (1987) *Assessing students: how shall we know them?* (2nd edn) (London, Kogan Page).

Rumelhart, D. & Norman, D. (1978) Accretion, tuning and restructuring: three modes of learning, in: J. W. Cotton & R. L. Klatzky (Eds) *Semantic factors in cognition* (Hillsdale, NJ, Lawrence Erlbaum Associates).

Sapsford, R. (1999) *Survey research* (London, Sage Publications).

Schank, R. C. & Cleary, C. (1995) *Engines for education* (Hillsdale, NJ, Lawrence Erlbaum Associates).

Schein, E. H. (1987) *Process consultation: lessons for managers and consultants* (Reading, MA, Addison-Wesley).

Skinner, B. F. (1958) Teaching machines, *Science,* 128, 969–977.

Suler, J. R. (1997) *The black hole of cyberspace.* Available online at: www.rider.edu/suler/psycyber/psycyber.html (accessed 28 July 2006).

Taras, M. (2003) To feedback or not to feedback in student self-assessment, *Assessment and Evaluation in Higher Education,* 28(5), 549–565.

Thorndike, E. L. (1932) , *The fundamentals of learning* (New York, Teachers College Press).

Vygotsky, L. S. (1978) *Mind in society: the development of higher mental processes* (Cambridge, MA, Harvard University Press).

Wenger, E. (1987) *Artificial intelligence and tutoring system* (Los Altos, CA, Morgan Kaufmann).