Assessment in Education
Vol. 14, No. 2, July 2007, pp. 000–000

Running heads
E. Hargreaves
The validity of collaborative assessment

The validity of collaborative assessment for learning

Eleanore Hargreaves*
University of London, UK

This article explores the features relating to the validity of assessment for learning, in particular the features of a *collaborative* assessment for learning, because of the learning benefits associated with collaborative learning. The article indicates what some of the learning benefits of highly valid collaborative assessment for learning might be, assuming that a valid assessment for learning *actually* promotes learning. It explores the idea that, for an assessment for learning to be valid, its learning outcomes must be socially appropriate for learners of the twenty-first century. The article illustrates some of these conceptual points, using descriptions of three collaborative assessments for learning currently being practised. Two of the illustrations are taken from the UK and one from the Eastern Caribbean.

Introduction

The validity of tests is a long-debated issue, but the validity of assessments for learning (or formative assessments) is a newer area of dialogue. The purpose of a summative test is summarizing learning, while the purpose of an assessment for learning is promoting learning: therefore it is important to *distinguish* between what characterizes a highly valid summative test and what characterizes a highly valid assessment for learning. This article first explores the features that contribute to the validity of assessment for learning, in particular the features of *collaborative* assessment for learning, because of the learning benefits associated with collaborative learning. Subsequently, the article indicates what some of the learning benefits of valid collaborative assessment for learning might be, assuming that an assessment for learning with high validity *actually* promotes learning (Black & Wiliam, 1998; Stobart, 2006). It explores the idea that an assessment will have more validity, if its learning outcomes are socially appropriate for learners of the twenty-first century. The article illustrates some of these conceptual points, using descriptions of three collaborative assessments for learning currently being practised. Two of the illustrations are taken from the UK and one from the Eastern Caribbean.

* School of Curriculum, Pedagogy and Assessment, Institute of Education, University of London, 20 Bedford Way, London WC1H 0AL, UK. Email: e.hargreaves@ioe.ac.uk

Assessment for learning

The phrase 'assessment for learning' (or 'formative assessment') denotes assessment used specifically to *enhance* learning processes or performances, rather than just *measure* them. Although any assessment will *affect* the student's learning in direct or less direct ways (see for example, Torrance & Pryor, 1998) the idea of assessment for learning is that the primary intention and the result of the assessment is enhanced learning. An extensive array of practices have become associated with 'doing' assessment for learning, such as teachers sharing assessment criteria with pupils, teachers encouraging pupils to self-assess and both teachers and pupils focusing on how we learn (ARG, 2002). In England, assessment for learning has also been included in government policy directives at primary and lower secondary levels, but these have invited teachers to use assessment rather to raise standards of performance than to enhance learning (for example, Miliband, 2003).

The meaning of validity in the context of collaborative assessment for learning

Consequential validity

In relation to summative assessment, a test is no longer seen to possess validity as an inherent property (Messick, 1989). Rather, its consequential validity is stressed and an assessment has high validity if the '… consequences of test interpretation and use are not only supportive of intended purposes but also consistent with other social values' (Gipps, 1994, p. 62). In terms of formative assessment, then, *the validity of an assessment for learning depends on how far the interpretation and use of the assessment actually leads to further learning.* This means that simply designing an assessment which is intended to promote learning does not make it valid. According to this view, to be highly valid, an assessment must actually engender a high degree of the desired learning too. In addition, the learning that ensues must be *consistent with other social values.* Who decides what the social priorities are at any one time is, of course, a problematic issue but this article tries to provoke thought on the issue.

The above comments suggest that the validity of an assessment for learning necessitates learning actually happening; but this outcome depends on how the assessor defines 'learning', which in turn depends on particular social values. Theoretically, then, if an assessment actually promotes learning, but that learning is out of keeping with social priorities, the assessment lacks validity. Therefore, to judge the validity of an assessment for learning, the assessor (and the assessed?) needs to know that the learning it promotes is valuable. A dialogue about what sort of learning we value needs to precede our evaluation of whether or not an assessment for learning is valid.

How do we know, however, that any further learning—processes or outcomes—has resulted from an assessment? Much valuable learning is neither predictable nor observable and we are all familiar with the experience of making meaning a long time after an event. One learner will respond differently from another: what is meaningful to one student may pass another by. Perrenoud (1998) reminded us that, 'In between what the pupil does and what passes through his or her mind, the mediations are complex. And what happens in the mind does not necessarily affect learning' (p. 89). So on all these counts, we may not know sometimes what further learning has resulted from our assessments for learning, and in this sense, we cannot fully judge the extent to which the

assessment is valid. But these are not good reasons to avoid promoting valid assessments for learning, using the best evidence we can to protect their validity. In the three illustrations given below, no attempt is made to quantify or even prove that valuable learning takes place, and yet the descriptions help to illustrate what we might be aiming at in a highly valid assessment for learning.

Fitness for purpose and the impact of assessment on learning

Stobart (2006) reminds us that validity in assessment is crucially about making an appropriate articulation between the assessment and the construct it samples. This, he tells us, 'is essentially about fitness-for-purpose' (p. 134). In terms of formative assessment or assessment for learning, where the emphasis is on consequences, the question is whether the *form* of the assessment is promoting the kind of learning we want our students to engage in. Different forms of assessment encourage different forms of learning, whether the assessment is formative or summative. If learners are engaging in collaborative learning on a regular basis, but the assessment is not collaborative, then the assessment is unlikely to sample its construct effectively.

In addition, assessment has an impact on the learning that happens in the classroom. If our assessment is for learning, then there is persuasive reason to make sure the backwash effects of any assessment s are positive; that is, that their form and content encourage the knowledge, skills, understanding and attitudes that are valued (Frederickson & Collins, 1989). In the case of collaborative learning, *the form and content of the assessment needs both to reflect and to encourage collaborative learning processes and outcomes.*

The assessment activities themselves can be a prime arena for generating useful learning through the assessment. If collaborative assessment activities themselves give rise to meaningful learning processes during the assessment activities being used, they will have some validity as assessments for learning, regardless of their longer term outcomes. In this context, the further learning generated may take the form of knowledge construction in an experience of participation, rather than the form of acquisition of specific skills afterwards (Sfard, 1998).

Other characteristics relating to the validity of collaborative assessment for learning

I am arguing, then, that the extent to which a collaborative assessment for learning is valid, depends on the extent to which it actually promotes valuable learning. Crooks' (1996, 2001) articles about the validity of formative assessment do not have collaborative learning as their focus; but they are useful in their suggestion that the validity of formative assessments can be *threatened* if the learning conditions are not supportive. In the case of collaborative assessment, some essential conditions that contribute to the validity of assessment for learning include: a classroom focus on learning; students' and teachers' familiarity with collaborative learning situations; their familiarity with the purposes and processes of the assessment; the appropriateness of the assessment task; and the meaningfulness of 'feedback' constructed. In traditional assessments, feedback is often considered to be what the teacher does to steer students towards the right track in their learning journey (Black & Wiliam, 1998). In a collaborative situation, feedback is more likely to be in the form of challenges posed by peers or other adults, self-regulatory

activity on the part of the learner her/himself (Zimmerman, 1989; Perrenoud, 1998) or 'productive helping' (Webb *et al.*, 2002) among members of a group.

Serafini (2001) equated the word 'assessment ' with either 'measurement', 'procedure' or 'inquiry'. A collaborative assessment for learning is unlikely to take the form of a measurement resulting in levelling or grading (although it might include some elements of measurement or be accompanied by measurements). This is because measuring depends on a view of knowledge in which knowledge is believed to '… exist separately from the learner, and students work to acquire it, not construct it' (Serafini, 2001, p. 385). This view of knowledge is contrary to a collaborative view in which students and teachers participate in knowledge-constructing dialogue. *A collaborative assessment for learning is likely to be an inquiry*, in which 'Knowledge is believed to be constructed by the individual within the social contexts of the learning event' and '… multiple interpretations are encouraged, and each learner transacts with different texts and the world to create meanings (Rosenblatt, 1979)' (Serafini, 2001, p. 387).

The value of learning associated with collaboration

I have argued above that a valid assessment for learning promotes valuable learning. I have suggested that validity is likely to be highest when the form of the assessment fits the purpose, and when the conditions are appropriate. I have hinted that the valuable learning may be associated with socially constructed knowledge; and that socially constructed knowledge is one feature of collaborative learning and assessment. In the next section, I examine the processes by which knowledge can be socially constructed in the classroom and why such learning could be considered 'valuable'.

The processes of collaborative learning

Knowledge can be socially constructed through collaborative learning processes. Collaborative learning is learning in which everyone has a more or less equal role in constructing the designated knowledge. All members of the group, whether a whole class group or a small one, have equal value although their contributions are various. Collaborative learners complement and build on each others' views to construct shared knowledge. For example, if members of a group have each read a different text relating to a topic (for example, the topic 'collaborative learning') through sharing the meanings they have each made from a different text, as a group they build up a rich picture (of 'collaborative learning'). This is socially constructed knowledge as the learning only partially consists of acquiring information that is already fixed.

Collaborative learners are likely to voice emergent or intuitive ideas and so have a chance to try them out in public (Crook, 1994). When there is a disagreement, this promotes reflection and a critical exploration of a topic within the group; and at the same time helps each individual firm up her/his own stance. Collaborative learning is therefore characterized (in Watkins' words) '… by the interplay of private and public reflection' (Watkins, 2004, p. 5). When a group is successfully collaborating, the agent of inquiry is not the individual, but the knowledge-constructing group: students take responsibility for knowing what needs to be known and for insuring that others know what needs to be known.

Other terms for this process of constructing knowledge socially include 'peer learning' and 'cooperative learning' (Boud *et al.*, 1999). Collaborative learning should

not be confused with peer teaching in which one student takes the role of teacher in relation to another student. And collaborative learning does not happen just by putting students in small groups. Although the teacher has a less intrusive role in students' learning, her/his importance in directing collaborative learning is no less than in directing traditional classrooms. The teacher's challenge is to encourage students to collaborate, rather than to compete or to work alongside each other. This does not happen overnight and can be a hugely challenging talk. Especially in classrooms where competition and hierarchical relationships are entrenched, it demands the teacher's and learners' skill and persistence (see Mercer *et al.*, 2004).

The learning outcomes associated with collaborative learning

The knowledge constructed through collaborative learning can be wide-ranging, but certain learning outcomes are particularly associated with collaborative learning (and its assessment). Boud *et al.* (1999) have noted that collaborative learning strategies are so successful at improving *subject-knowledge* that often the strategies are employed even when the longer term agenda is about individualistic achievement, rather than about cultivating collaborative skills and outlooks. In addition, collaborative learning can produce impressive learning gains without extra staffing in the classroom. Certainly, as Brown and Campione (1994) have shown, collaborative learners seem to retain subject matter better, make more flexible use of information and apply it more widely. Therefore collaborative learners are more likely to succeed even in traditional versions of 'learning'.

However, collaborative learners also show a more positive attitude to their subject matter and approach it in a more active, involved, exploratory and in-depth way, as well as being more aware of how they learn: all of which lead both to better lifelong learners as well as to better performers in formal examinations (Watkins *et al.*, 2001). Student-teacher relationships improve through the use of collaborative learning, and as they do, learners' motivation and feelings of autonomy and competence increase, which in turn also lead to further learning processes (Battistich *et al.*, 1995). A collaborative assessment based on collaborative learning may well contribute to the learner's disposition to carry on learning and her/his improved strategies for doing so.

Whatever the subject matter being studied, collaborative learning develops in learners higher-level thinking skills, including the ability to reflect and think critically. If learning is defined as 'meaning-making' rather than 'acquiring information', reflection and critical thinking are important tools, especially in this era where information is easily accessed but less easily made sense of.

Other skills promoted by collaborative learning include oral communication skills, argumentation skills, empathy, social interaction skills and leadership skills. Students' relationships among themselves improve as they take responsibility for one another in collaboration, overriding racial and other differences. Learning can be seen in this context as knowledge of participating and contributing socially.

While all the above may not be traditional learning goals, their usefulness as learning outcomes is underlined by a new emphasis today on generic attributes or capabilities, because of the increasingly complex situations people meet inside and outside the workplace and hence a need to embrace 'lifelong learning' (Boud *et al.*, 1999). Indeed, some argue that all '… human learning is necessarily and fundamentally social' (Watkins, 2005, p. 14). When we talk about assessment for learning promoting

'learning', while we might mean *retention and use of subject-specific information or skills* (as assessed through written examinations), an additional notion might be valuable, that of *meaning-making, and constructing knowledge of participating and contributing socially and of learning in diverse contexts.*

Is this the kind of learning we consider valuable? If so, any assessment that actually promotes meaning-making, and knowledge of participating and contributing socially and of learning in diverse contexts has some validity as an assessment for learning. If we consider only individually-measurable outcomes to be valuable learning, then any assessment for learning that promotes those has high validity.

During a university teaching session about assessment for learning, I asked a group of 21 teachers and examiners what they would like pupils to have learnt as a result of their schooling, and their responses included the following. They wanted pupils to have learnt to be: empathetic, self-aware, socially competent, resilient, creative, confident, responsible, reflective, self-evaluative, critical, inquisitive, lifelong learners, enthusiastic about learning, good citizens, independent, self-reliant, happy, curious, inventive, excited, risk-taking, tolerant, kind, respectful, caring, good team players, streetwise, spontaneous, well qualified, physically active and healthy. They wanted pupils to have the skills of: applying knowledge, learning to learn, problem solving, self-management, communicating fluently and making sound moral judgements. Finally, they wanted pupils to know how to make a positive social contribution. Although this survey was carried out in an MA classroom, it was not a teaching module relating particularly to collaborative learning. And yet many of the skills mentioned by participants as being valuable learning outcomes are those particularly promoted by collaborative learning and its assessment for learning (perhaps not the being physically active and healthy outcomes —and becoming streetwise through collaborative learning might be a little ambitious!)

In summary, I am suggesting that a valuable version of learning is that of meaning-making, and constructing knowledge of participating and contributing socially and learning in diverse contexts. I am proposing that this version of learning is supported by the social need today for people who are confident, self-directed and continuing learners, have a sensitive awareness of others, be able to interact well with others and be flexible in their application of information. I am questioning whether a version of learning as acquiring and using information and skills still has the social currency it had before the information 'revolution' in which information is readily available but wise application of it still depends on choices made by social beings. To evaluate the validity of assessment for learning, we assessors need to be clear which versions we value.

Descriptions of collaborative assessments for learning in practice, to illustrate some points about validity

Collaborative assessment for learning in a London primary school

Paul teaches in the Year 6 class (10- to 11-year-olds) of Coleridge Primary School, situated in London, UK. Collaboratively with two other colleagues, Paul developed a form of collaborative assessment to help his pupils learn deeply about specific concepts, and also to help them learn to become confident learners, learn about communicating and learn to cooperate (Campbell *et al.*, 2005).

Paul used his design of assessment for learning in his teaching of discrete and continuous data and how the data should be represented. When he began, his pupils were

already familiar with learning collaboratively. On this occasion, the children spent time in groups of three to come up with ideas about *What information we need in order to make a graph useful*. Paul then compiled a plenary flipchart list of his pupils' suggestions. When the whole class was satisfied with the list of criteria about what makes a graph useful, each child individually created a story around what a given graph meant, making sure to include on their own graph all the information mentioned as necessary on the class list. Having worked alone on this task, the group of three reconvened to think again of any information that should have been included on the class list. When all members of the group agreed, the new suggestions were submitted for inclusion on the revised class list. After much discussion, their final list read simply as:

- A full title ('A graph showing the distance travelled by … during …etc.')
- Labelled axes (with units clearly stated and scale demarcated).

Now the pupils moved into new groups of three. Each of the three pupils in a group shared with the other two children their own graph and the story they had created around it. When all three had done this, the group had to divide up 60 marks among the three of them, to reflect how well each child had taken on board the necessary pieces of information, on their graph. Although this process sounds anti-collaborative, Paul reported that the activity generated a very rich dialogue about the most important features of the graph, and necessitated a large degree of cooperation and communication among group members as they fed back their reasons to each other. The teacher, Paul, never knew how they had allocated marks, nor what their stories were. He was busy listening in on the children's assessment dialogues taking place.

Finally, Paul facilitated a whole-class inquiry about what pupils concluded about representing data on graphs, and also how they had experienced learning in groups. Their comments suggested that, at least for some children, the assessment had led to further mathematical learning, including making sense of mathematical information. For example, Jai wrote: *'I found that once I'd chosen the scale for my axes the story started telling itself'*. Another pupil, Stanley, told Paul: *'I didn't realise the same line could tell lots of different stories'*. However, Carol remarked that: *'Knowing what made a good story helped me get my ideas right'*. And Andros said, *'Looking at other people's graphs and talking about how many marks they should get helped me see what was good and not so good about my graph'*.

Discussion

In terms of the above ideas relating to validity, as far as we can tell from observing the children, this assessment seems to have had a high degree of validity in terms of its fitness for purpose. The collaborative form of the actual task and the teacher's role as guide not assessor, was consistent with the teacher's valued view of learning as: developing mathematical skills; constructing knowledge of participating and contributing socially; and learning about learning. In addition the learning conditions were supportive since the children were accustomed to collaborative learning and dialogue. Although in this design of assessment for learning, the pupils actually measured their achievement and each others', the purpose of the measurement was *not* for the teacher to have a snapshot of each child's achievement, but to help the pupils think deeply and inquire about a particular (mathematical) concept and to learn about their own learning and negotiating

with others. It was during the measurement exercise that the pupils were able to give each other the most useful feedback.

Collaborative assessment for learning in St. Kitts, West Indies

The year 2005–2006 heralded the start of an innovative and collaborative form of assessment for learning on the island of St. Kitts in the Eastern Caribbean, West Indies. Traditionally, teachers and learners of the island's 24 primary schools concentrate their energies towards success in the primary leaving examination, the Test of Standards. However, parents, educators and employers recognize that the educational tunnel vision that results from a focus on the examination may deny pupils the rounded learning experiences that prepare them to be socially responsible people, or flexible and adaptable employees.

This was why the Department for Social Studies within the Curriculum Division of the Ministry embarked on the ambitious venture in which collaborative learning and collaborative assessment for learning play a major role. The aim of the venture is to promote the subject area of citizenship, and to promote learning-awareness and learning responsibility in students, as well as social awareness and social responsibility. The programme instigators agree that if the learning outcomes they are hoping for include learning about learning and social learning, then to increase their validity, assessments for learning should include collaborative aspects. In order to raise the profile of this non-examination focused learning and assessment, the Department sought funding from UNESCO and also invited the Prime Minister, Hon. Dr. Denzil Douglas, to support the scheme.

The programme covers two school terms. To participate in the project, all children in the top grade of primary school (aged 11–12 years old) engage in classroom learning about 'responsible citizenship'. After some such sessions, they decide on a project that they would like to carry out within the community (outside school) in order to improve an aspect of the community and so benefit the wider community and themselves. In doing this, they work collaboratively as a whole class (with numbers of up to 25 pupils). Intermittently, in the classroom their teacher orchestrates a process of collaborative assessment for learning in which the whole group focuses on learning about learning, making the following inquiries: Is everyone completing her/his task? Are we working as a team? How can we help each other out? What are we pleased about? What are we not pleased about? What changes can we make to how we are working? (Jacobsen *et al.*, 2005). The pupils give themselves and each other feedback about the processes as they collaboratively self-regulate.

The next stage in the programme is for the pupils to carry out the project they have designed, in the community. For example, one group of pupils decided to make a mural to enhance the environment and display important images. Another group started a project with people in prison. A third and fourth group started working with people who have alcohol addiction and drug addiction respectively. Community officers from the Department of Community affairs support groups of classes with their projects. The children are encouraged to *engage* with the people in the local community in their activities in order that the community can learn from them as well as help them.

On completion of the project (for example, when the mural is finished) the children write up and submit a collaboratively compiled report of the project, as one

product of the programme. In compiling this, they self-assess their learning processes and the outcomes. The other product is an assessment in the form of a presentation of the project to the people who now benefit from it: the pupils choose to describe verbally or visually what they have done; and they explain their assessment of how the project has helped their learning as a group. During this assessment process, the community representatives ask the children questions about their actions and their learning as feedback to promote further self-reflection and assessment. The presentation is collaborative, as is the response to questions (despite the large group size of up to 25). The aim is that learning happens during the presentation assessment itself; that the children are learning as they are engaging with community personnel and as they are reflecting critically on what they have done and how they have learnt.

Discussion

Probably the greatest threat to the validity of the assessments described above is the traditional learning conditions that the participating children are used to in all other areas of their school experience. The unusual nature of the citizenship assessments and the preparation towards them might be an obstacle to their learning. Clearly, the programme instigators had a clear idea about the broad and social kinds of learning that this project would promote, and in that sense, the assessments are fit for purpose. However, socially in St. Kitts, rote learning for the Test of Standards is still valued by some in the twenty-first century, lessening the validity of the assessments as 'consistent with other social values'. The fact that these are inquiry-based assessments, led collaboratively by pupils, in this case might *reduce* their learning benefits unless they are being very skilfully facilitated: and the classroom teachers in this project may not share the same learning aims as the programme instigators and so may not offer appropriate support. The large group sizes add to the difficulty of this skilful facilitation.

At present it is impossible to judge whether these assessment inquiries have actually promoted any valued learning, since the presentations are yet to happen at the time of writing, and no feedback has been reported by the children to their teachers. However, in this case of collaborative assessment for learning, there are several reasons to suspect that validity may be lessened, at least initially.

Collaborative assessment for learning in a University Masters module

Tutors and students participating in the Masters Degree module about assessment for learning at the Institute of Education, London University, UK, resolved to turn an obligatory summative assessment into a learning event (Hargreaves, 2005). All concerned aimed to promote valuable learning through valid collaborative assessment for learning, despite being obliged to assess learners summatively as well. The valuable learning we aimed to promote included constructing meaningful knowledge about assessment for learning itself; learning to learn effectively through critical reflection; and constructing knowledge of participating and communicating as part of an assessment for learning community.

This module about assessment for learning was the same module mentioned earlier, in which participants were invited to prioritize school learning outcomes. Its participants came from a range of countries including the UK, Rwanda, Cyprus, USA and the Caribbean. In this module, students became familiar with learning collaboratively

from the start of the module. We aimed to establish a learning community in this classroom by encouraging students to interact with different combinations of other students and get to know each other. Most of the learning was through group or pair activity, in response to some teacher stimulus. One student, Katherine, described how these early activities helped her when it came to the collaborative assessment that followed. She wrote that the collaborative assessment '*... was not a task I had undertaken before and my overall feelings were positive towards group learning, which mainly stemmed from the other group activities we had completed in earlier sessions*' [Reflections, Feb 2006. Each student wrote and submitted some reflections on the experience of the module, following its conclusion: these and like quotations are taken from these]

After five weekly sessions, we invited each student to write on a large post-it a very brief description of how they imagined a 'valid assessment for learning'. They stuck their post-its on a long sheet of paper along the wall and then grouped them according to their views about valid assessment for learning. The groupings formed the basis for six Assessment Task writing groups. The selection process was not perfect (one group had to disperse and join others), but it did mean most people were in groups with a shared interest and no one was excluded for social reasons. One student, Patricia, recognized the difficulty of being forced to be collaborative in class: '*The social elements of the module were both threatening and immensely rewarding, but un-learning and re-learning was necessary* [Reflections, Feb 2006].

Students are normally assessed summatively, in the university, by writing 5000 words on a set topic related to each module they attend. In this module, in these interest groups, students had to write collaboratively 2,500 words to exemplify a valid assessment for learning, and to explain what made their assessment examples valid (and later wrote an individual piece too). It was this activity that gave rise to the idea for this article and the examples described by two of the groups are used as illustrations above. The group had three weeks to produce the assignment which would also be published and circulated for other interested people to read. It would provide half the basis for the final grade for the module. The assessment criteria for the assignments had been agreed by the community.

During each class session, the six small groups stopped their activity to reconvene as a whole community, to assess the collaborative processes, feed back to each other and make plans for improved collaboration. The rest of the time, they focused on their task. Different groups approached the task in different ways: some did everything collaboratively all the way through; others divided up the task and kept coming back together to coordinate. Most people used 'phone and email extensively and some groups met, outside of class time. Paul (whose own classroom assessments are represented above) described eloquently how the assignment-writing process was one of constant and instant feedback, which became itself a process of collaborative knowledge construction:

> Ideas are proffered, group members assess them in open dialogue and in doing so the person proffering receives immediate feedback that moves their thinking forward (they learn). More than this, in a cascade of similar learning events, new knowledge is constructed by the group… [Reflections, Feb 2006]

Another student, Sharon (whose community assessments are also described above), referred to one group moment when she '*wanted to scream'*. However, she went on to say that this conflict did enrich her thinking, but it depended on her good initial relationship with peers:

> It was in this instance that the social chat we had before [starting to work on the assessment task] helped to provide the direction for our activities. The challenge of each others' thoughts brought the most richness to my learning [Reflections, Feb 2006]

During the penultimate session of the module, the six small groups paired up into three pairs. Each pair swapped assignments, so that everyone in Group A read Group B's assignment and vice versa. Group A then discussed Group B's assignment in relation to the issues. Group A and Group B then had a feedback dialogue together about each other's assignment in relation to the issues. The goal was to enhance each other's learning, not to nit-pick about the format of the writing. For this reason, neither group could change their text after the dialogue. What they could do, was collaboratively to self-assess against the assessment criteria as a group, taking on board what they had learned from the feedback dialogue. The assignments and self-assessments were then handed in to the tutors and both were considered when the final grade was decided. While the process of peer feedback was uncomfortable in some ways, for example, because participants had some fear of being judged and of making criticisms of peers, Katherine was later able to reflect that the process increased her confidence in learning:

> The fact that the assessment represented 'us' and not an individual was also less threatening and we were more able to share our work. This collaboration built up respect and allowed all members of the community a learning voice [Reflections, Feb 2006]

Discussion

In this case it was clear that some learners had constructed new knowledge about the subject, assessment for learning. Through the assessment for learning inquiry processes, both during the writing of an assignment and in assessing each others', they had become more aware of learning about their own learning and about being socially active and responsible, all of which were learning priorities for this module. The learners' subsequent written *Reflections,* as well as the group assignments, suggested the same, that the assessments had promoted valuable learning, indicating a good level of validity for these assessments. Perhaps those students who could already reflect critically on their own learning, who already self-regulated and who already found social participation and communication easier would have benefited most in that they were used to maximizing their learning through these processes. However, it seems that classroom conditions were anyway conducive to learning.

Because the module focused on learning as well as assessment, many students had thought about what kind of learning they valued and shared the learning goals of the tutors, even though the assessment model went against the university norm. Since those goals were about acquiring subject knowledge as well as participating as members of a learning community and developing learning strategies, the assessments fitted their learning purposes in that they were collaborative, inquiry rather than measurement-oriented (although within a measurement framework), and primarily took the form of

group self-initiated assessment although they were also assessed by the tutors. In this context, feedback was particularly timely and relevant, as given by peers during productive helping as they constructed their assignments together. This would be another factor to improve the assessment's validity, as we are in no doubt that at least some students actually did learn through these processes.

Concluding comments

Above are three illustrations of collaborative assessments for learning in practice and some comments about the extent of their validity. My argument has been that such assessments are likely to have a high degree of validity if the following conditions are met:

- the assessment for learning actually leads to further learning of a kind that is consistent with other social values
- the form and content of the assessment for learning reflects and encourages valuable learning
- the assessment is an inquiry rather than a measurement
- classroom conditions are conducive to valuable learning.

However, this list is only useful when we have evaluated the learning that we, as a society, want to prioritize. While we could assume that our schools' curriculum embodies prevalent social values, such a notion could usefully be challenged. I have argued in this paper that learning is particularly valuable if it includes learners making their own meanings in a particular area of knowledge, constructing knowledge of participating and communicating socially and reflecting critically on their learning in diverse contexts; as well as retaining, using and applying information appropriately. These are some of the learning outcomes and processes associated with collaborative learning, and therefore with valid collaborative assessment for learning in that context.

Despite their desirable outcomes, collaborative learning and collaborative assessment for learning are not common in classrooms across the world. A long list of factors mitigates against teachers taking the risk of exploring the benefits in their own classrooms. Perhaps top of the list is the perceived pressure of externally imposed summative assessments and the individual, national and international competition they encourage: despite the evidence that exam results improve where students learn collaboratively. Other perceived reasons to avoid collaborative learning and assessment for learning include resourcing; this is despite the fact that the main resource needed is familiarization with collaborative processes, because other resources can actually be reduced by collaboration. However, cultural scripts or habits (Stiegler & Hiebert, 1998) that triumph individualism or social hierarchy, plus teachers' and students' traditional expectations are probably the hardest to challenge.

Acknowledgments

I would like to thank Paul Campbell and Sharon Rattan for their special help with this article, and all the participants on the 2005 Assessment for Learning Module who informed and inspired it.

Notes on contributor

Eleanore Hargreaves is lecturer in Assessment for Learning, Teaching and Learning in Classrooms and Building Learning Communities, at London University's Institute of Education, UK. She has worked in Test Development at the National Foundation for Educational Research, carried out assessment projects in Egypt, India, South Africa and Pakistan and led Action Research enterprises into Assessment for Learning in local authorities across the UK.

References

Assessment Reform Group (2002) *Assessment for learning: 10 principles*. Available online at: http://www.assessment-reform-group.org.uk (accessed 01/10/2005).

Battistich, V., Solomon, D., Watson, M. & Schaps, E. (1995) Schools as communities, poverty level of student populations, and students' attitudes, motives and performance: a multilevel analysis, *American Educational Research Journal*, 32(3), 627–658.

Black, P. & Wiliam, D. (1998) *Inside the black box* (London, King's College).

Boud, D., Cohen, R. & Sampson, J. (1999) Peer learning and assessment, *Assessment and Evaluation in Higher Education* 24(4), 413–426.

Brown, A. & Campione, J. (1994) Guided discovery in a community of learners in: K. McGuilly (Ed) *Classroom lessons* (Cambridge MA, MIR Press), 153–186.

Campbell, P., Grey, J. & Caldwell, S. (2005) Talking and thinking together in assessment for learning, in: E. Hargreaves (Ed) *Examples of effective/valid assessment for learning and explanations of what makes it effective/valid* (London University Institute of Education, unpublished document), 1-7.

Crook, C. (1994) *Computers and the collaborative experience of learning* (London, Routledge).

Crooks, T. (2001) The validity of formative assessments. Draft paper presented at *BERA 27th Annual Conference*, University of Leeds, dates? September.

Crooks, T., Kane, M. & Cohen, A. (1996) Threats to the valid use of assessments, *Assessment in Education*, 3(3), 265–285.

Frederikson, J. & Collins, A. (1989) A systems approach to educational testing, *Educational Researcher*, 18(9), 27–32.

Gipps, C. (1994) *Beyond testing: towards a theory of educational assessment* (Lewes, Falmer Press).

Hargreaves, E. (2005) Assessment for learning? Thinking outside the [black] box, *Cambridge Journal of Education*, 35(2), 213–224.

Jacobsen, H., McFarlane, M. & Rattan, S. (2005) Metalearning and reflection in assessment for learning, in: E. Hargreaves (Ed) *Examples of effective/valid assessment for learning and explanations of what makes it effective/valid* (London University Institute of Education, unpublished document), 7-15.

Mercer, N., Dawes, L., Wegerif, R. & Sams, C. (2004) Reasoning as a scientist: ways of helping children to use language to learn science, *British Educational Journal*, 30(3), 359–377.

Messick, S. (1989) Validity, in: R. Linn (Ed) *Educational measurement* (3rd edn.) (New York, American Council on Education and Macmillan), 13–103.

Miliband, D. (2003) Every schoolchild is special, *The Observer*, Sunday 1 June. Available online at: http://observer.guardian.co.uk/comment/story/0,6903,968001,00.html (accessed on 5/3/2003).

Perrenoud, P. (1998) From formative evaluation to a controlled regulation of learning processes. Towards a wider conceptual field, *Assessment in Education*, 5(1), 85–102.

Rosenblatt, L. (1979) *The reader, the text, the poem* (Carbondale, IL, Southern Illinois University Press).

Serafini, F. (2001) Three paradigms of assessment: measurement, procedure, and inquiry, *The Reading Teacher*, 54(4), 384–393.

Sfard, A. (1998) On two metaphors for learning and the dangers of choosing just one, *Educational Researcher*, 27(2), 4–13.

Stiegler, J. & Hiebert, J. (1998) 'Teaching is a cultural activity', *American Educator*, 22(4), 4–11.

Stobart, G. (2006) The validity of formative assessment, in: J. Gardner (Ed) *Assessment and learning* (London, Sage), 133–146.

Torrance, H. & Pryor, J. (1998) *Investigating formative assessment* (Buckingham, OUP).

Watkins, C. with Carnell, E., Lodge, C., Wagner, P. & Whalley, C. (2001) Learning about learning enhances performance, *NSIN Research Matters*, 13, 000–000.

Watkins, C. (2004) Classrooms as learning communities, *NSIN Research Matters*, 24, 000–000.

Watkins, C. (2005) *Classrooms as learning communities* (London, Routledge).

Webb, N., Farivar, S. & Mastergeorge, A. (2002) Productive helping in cooperative groups, *Theory into Practice*, 41(1), 40–48.

Zimmerman, B. (1989) Models of self-regulated learning and academic achievement, in: B. Zimmerman & D. Schunk (Eds.) *Self-regulated learning and academic achievement* (New York, Springer-Verlag), 1–24.