Personal Understanding of Assessment and the Link to Assessment Practice: the Perspectives of Higher Education Staff

N. Reimann[a*] and I. Sadler[b]

[a]School of Education, Durham University, Durham, UK; [b]Faculty of Health and Life Sciences, York St John University, York, UK

---

[*] Corresponding author. E-mail: Nicola.reimann@durham.ac.uk

Personal Understanding of Assessment and the Link to Assessment Practice: the Perspectives of Higher Education Staff

Abstract

The study investigates how higher education staff understand assessment and the relationship between these understandings and their assessment practices. Nine individuals attended a workshop that guided them through the creation of a concept map about assessment, which was subsequently discussed in one-to-one semi-structured interviews. We found considerable variation in understanding of assessment both between and within participants, and this appeared to be a consequence of the varied contexts within which assessment operates. Some assessment practices were highly complex and at times closely entwined with teaching. In addition, individual's practices helped to illuminate variation in how underlying concepts (e.g. assessment for learning) were understood. The approach supported the construction of the participants' understanding of assessment and enabled the exploration of the interplay between thinking and reported practice, which were closely aligned. It also drew attention to the need to further develop methodologies which capture both the complexity of thinking about assessment and of real world assessment practices.

Introduction

Assessment in higher education is at the foreground of much contemporary research, policy and discussion in the sector, resulting from a general dissatisfaction with assessment and feedback both on the part of students and teachers (Nicol 2010). As a consequence there have been considerable attempts to reshape the assessment and feedback landscape both from a theoretical and practical perspective. This has led to a greater emphasis upon formative, assessment for learning (Sambell et al. 2013) and learning-oriented assessment (Carless 2015). However, in order to change and enhance current assessment practices, more needs to be known about the ways in which academics understand assessment and the relationship between their thinking and their assessment practices, but such research is still rare. This is the gap the present study aims to address.

Literature Review

Models of Assessment

Assessment is well theorised, whilst debates are still on-going. Since the seminal paper by Sadler (1989) on formative assessment in higher education there has been extensive writing about assessment and feedback in terms of how it can be most effectively understood. The overarching theme of much of this work is that assessment is not something that should simply be done to students, but something they are and should be actively involved in. Therefore, rather than seeing assessment as summative end-point testing and feedback as a grade or written comments provided to the student by the tutor, proponents of *assessment for learning* regard it as

an on-going, intrinsic component of instruction, involving teachers, learners and their peers in making evaluative judgements and taking action to close the gap between actual and desired performance (e.g. Wiliam 2011). Whilst there are variations in terminology (e.g. assessment for learning, formative assessment, learning-orientated assessment), they all tend to point to assessment as a constructivist process with the students as active, self-regulated learners. A contrasting model to assessment for learning is *assessment of learning* with a focus on measurement, certification and accountability (Gipps 2012). An additional concept has been proposed by Torrance (2007) who drew attention to instrumental interpretations of formative assessment. *Assessment as learning* involves a focus on assessment procedures and extensive coaching in order to meet explicitly stated requirements, to the extent that criteria compliance replaces learning. This has similarities with what Marshall and Drummond (2006) called 'letter' of assessment for learning, where teaching rigidly focuses upon assessment procedures and loses the 'spirit', i.e. the learning focus of assessment for learning.

A number of authors have critiqued the dichotomy between formative and summative assessment that has arisen in the literature and argue for their connectedness (e.g. Taras 2005; Lau 2015). Boud (2000) reminds us that assessment always does 'double duty' and cautions us to attend to all its purposes, even if they conflict with each other. The models outlined above, which mainly arise from conceptual discussions and theorisations of assessment, emphasise the complexity and tensions within assessment in higher education (Price et al. 2011). However, much less is known empirically about the ways in which this complexity plays out in the understanding

and practices of higher education teachers. Such insights have the potential to direct

efforts to enhance assessment practice, both by academic developers and managers.

Link between Understanding and Practice

Based on empirical studies of approaches to university teaching (e.g. Trigwell and

Prosser 1996), a unidirectional influence of thinking upon practice has been assumed

for many years. However, contextual and methodological limitations of these studies

have been pointed out (Kane et al. 2002), and results from interventions have been

equivocal (Ho et al. 2001). Several authors have questioned the direction of the

relationship and suggested that a change in practice may equally act as a catalyst for a

shift in attitude or conception (Devlin 2006; Eley 2006; Sadler 2012). Based on

empirical studies of lecturers' reflection on observed teaching situations, McAlpine et

al. (2006) capture the complex and non-linear relationship between abstract

conceptions and context-specific action in four 'zones' of teacher thinking, located on

a continuum of specificity. Although such ideas have been widely discussed in

relation to teaching, relatively little is known empirically whether and the way in

which they may apply to assessment.

Samuelowicz and Bain's (2002) research on assessment echoes some of the early

findings on the relationship between understandings of and approaches to teaching.

Their study showed that academics with a view of teaching as information

transmission also believed that assessment should test the retention of facts and

reported corresponding assessment practices. On the other hand, academics who

considered teaching as helping students to construct understanding focused on

integrating assessment with teaching and using feedback to improve understanding

and challenge misunderstandings. The latter understanding of assessment has strong resonances with the notion of assessment for learning. A number of interview-based studies have confirmed the existence of contrasting beliefs/conceptions of assessment. In Watkins, Dahlin and Ekholm's (2005) research, which focused on university teachers' conceptions of the backwash effect of assessment, one key dimension in which conceptions varied was the way in which the relationship between teaching and assessment was understood. A key distinction was between an 'external' relationship with assessment being regarded as end-point testing separate from teaching, and an 'internal' relationship where teaching and assessment are regarded as overlapping and interacting, which is akin to assessment for learning. Postareff et al. (2012) similarly found different conceptions of assessment and labelled these as 'reproductive' and 'transformational'. These involved assessment as measuring, on the one hand, whether students could correctly repeat information and, on the other hand, students' thinking processes. In their research a reproductive conception was highly consistent with 'more traditional' assessment practices such as pen and paper examinations, whilst a transformational conception was consistent with 'alternative practices' that support student learning rather than testing knowledge. Whilst their study showed a strong relationship between conceptions and reported practices, it also identified a few instances of incongruence between conception and practice. However, little is known about the exact nature of what was termed 'alternative practices' and details of the ways in which practices were affected by variations in conceptions.

Offerdahl and Tomanek's (2011) longitudinal study paints a more complex picture of the relationship between thinking and practice in relation to assessment. They report that thinking about assessment became more sophisticated through practical

experimentation with formative assessment strategies, but such revised thinking did not necessarily lead to changes in subsequent practice. Boyd and Bloxham's (2013) recent critique of the way in which the relationship between theory and practice is commonly represented is of specific relevance to assessment. Based on their research of university teachers learning to grade student work, they examine the metaphor of the 'gap' between theory and practice and argue that abstract knowledge on the one hand and practical, socially situated ways of working on the other are closely integrated. They propose a conceptualisation of professional learning as 'interplay' between vertical public knowledge, such as the knowledge encapsulated in assessment policy and theories of assessment, and horizontal practical wisdom, as enacted for instance in the holistic and instinctive actions involved in marking. Participation in assessment related activities and the creation and negotiation of artefacts such as marking grids involve both the codification of practical wisdom as well as the mediation of public knowledge.

It has also been suggested that conceptions of assessment may be context-dependent (Samuelowicz and Bain 2002; Watkins, Dahlin and Ekholm 2005; Postareff et al. 2012). This is in line with McAlpine et al. (2006) who stress the crucial role of context-based knowledge for teaching and Marton and Pong (2005) who found that individuals could hold more than one conception for the same phenomenon. They also discuss the range of labels that have been used to denote 'different ways of understanding' (Marton and Pong 2005, p335) and the slight variation in meaning implied by each label. This also applies to the assessment specific studies reviewed above. For instance, Samuelowicz and Bain's (2002) study focuses on 'orientations to assessment practice', Postareff et al. (2012) examine 'conceptions of assessment',

7

and Offerdahl and Tomanek (2011) talk about 'assessment thinking'. In the present study we refer to understanding of or thinking about assessment, deliberately using broad and generic terms.

The review of the literature has highlighted that in general, understandings of assessment show variation and this has some similarities with the variation found in understandings of teaching. In addition, certain understandings of assessment appear to be closely aligned to certain assessment practices. However, there is a degree of uncertainty about the precise nature of the relationship between understandings of assessment and assessment practices, particularly in terms of the direction of this relationship and contextual variation. Recently, there has been a line of research into assessment practices in context, using in-depth, fine-grained approaches with small numbers of participants (Carless 2015). The present study aims to contribute to this line of research by investigating the ways in which HE staff think about and practise assessment, and the relationship between understandings and practices.

Methodology, Design and Methods

Interviews are a relatively common method in studies of conceptions of teaching and assessment (e.g. Prosser, Trigwell and Taylor 1994; Samuelowicz and Bain 2002), most notably in phenomenographic research, but the limitations of such data have been foregrounded by several authors. The main issue is how representative language, in the form of interview accounts, is to an individual's everyday experiences (Säljö 1997) and there is some suggestion that this connection may not always be particularly strong. A study specific to teaching in higher education

exhibited significant disjunction between conceptions described during interview and claimed educational practices (Murray and MacDonald 1997). Argyris and Schön's (1974) differentiation between theory-in-use and espoused theory also provides support for a potential lack of relationship between experience and descriptions. Most of the assessment specific studies reviewed above were also based on interview data and it could be argued that their findings may be an artefact of the method. Supplementing interview data with data generated through alternative methods should therefore shed further light on the research questions. Thus a key issue in the design of this research was how to best gain insights into the way in which staff thought about assessment and their assessment practices. Concept mapping offered a potentially fruitful approach to alleviating some of the limitations of using interview accounts in isolation. A concept map consists of concept labels, which are arranged hierarchically linked by labeled lines that make the relationship between concepts explicit, thus creating propositions (Novak and Cañas 2008). The concept mapping approach has previously been used in research and teaching in all sectors of education, including higher education, for the purpose of making tacit and abstract knowledge visible and assessing conceptual development over time (Hay 2007; Hay, Kinchin and Lygo-Baker 2008). However, due to the active construction process that is involved in creating a map and the fact that the nature of the knowledge represented in the map changes through the process of constructing it, concept maps should not be considered to simply provide a 'window to the mind'. We have therefore experimented with dialogic concept mapping (Hay 2008), which allowed individuals to construct and clarify their understanding during the data collection process.

Staff development workshops exploring personal understanding of assessment and assessment practices and introducing participants to concept mapping were broadly advertised in two UK universities. Prior to the workshops participants were provided with information about the option to take part in the research project. The workshop guided participants through the development of their own concept map and facilitated a discussion of each other's personal understandings of assessment in higher education and assessment practices. Participants were asked to bring 'assessment artefacts' (i.e. self selected representations of their assessment practice) to the workshop and incorporate these into their concept map as examples for specific concepts, using labels of a different colour. Ethical approval for the study had been obtained at both institutions and out of approximately 20 workshop participants nine members of staff volunteered to participate. The broad subject areas represented in the sample included health, business, psychology and education and staff with a variety of roles and levels of experience. Seven out of the nine had been working in higher education for more than five years, and all had current experience of designing assessment and assessing student work. For research participants only, the workshop was followed by a one-to-one semi-structured interview. Interviewees were asked to talk through their map and prompted to explain labels and propositions. The second part of the interview focused on examples of their practices, asking them to discuss the artefact they had brought and practice labels on the map, and finally there were questions about the influences upon thinking and practice and their experience of the concept mapping activity. Participants were encouraged to amend their map both before and during the interview and were later-on provided with a photograph of their map. The interviews were approximately one hour in duration and subsequently transcribed. The integration of a variety of methods aimed to triangulate and gain

multiple perspectives on personal understanding to address the limitations of previous research, make participants' practices visible and enable a focus on the relationship between thinking and practice.

Data Analysis

The data were analysed in a broadly interpretivist, qualitative and iterative manner, immersing ourselves in the maps and artefacts, reading and re-reading the interview transcripts and considering each participant in turn. Using the models of assessment discussed above as an analytical framework, each type of data was initially examined in its own right. Both researchers provided independent accounts of the data first, which in a second step were discussed in order to reach agreement on shared interpretations. The maps were analysed in terms of overall structure (e.g. whether they contained features such as clearly consisting of two halves), key oppositions incorporated into the map (e.g. formative – summative, assessment for learning – assessment of learning), concepts located towards the top of the hierarchy and concepts with a high number of attached link lines. Where relevant, we have referred to these structural features in the findings section, and selected excerpts from maps have been included for illustration purposes. Traditionally concept maps have been analysed quantitatively through comparison with expert maps, awarding scores for concepts, link lines and propositions. Kinchin and Hay (2000) developed an alternative qualitative method of describing maps where they distinguished between 'spoke', 'chain' and 'net' structures as representing different levels of complexity. However, as has been discovered by other authors (e.g. Liu and Hinchey 1996), our maps varied considerably in complexity, structure and propositional content, and

some maps resembled hybrids between mind maps and concept maps (Davies 2011) due to link lines not always being labeled. Since the analysis needed to take account of the dialogic process and the relationship between the data, maps, artefacts and interviews were considered in close conjunction with each other. The interviews increasingly came to the fore as they illuminated the thinking behind the maps and shed further light on the personal understandings and practices encapsulated in maps and artefacts. Interviews were analysed thematically and particular attention was devoted to verbal explanations of specific elements of the maps and of concepts that were emphasised as important or discussed at considerable length. Taking account of all the data, meaning was condensed by capturing the essence of personal understanding and of practices in vignettes for all participants.

Findings

i.   Contrasting Ways of Understanding Assessment and Personal Variation

At a general level, two contrasting ways of understanding assessment were evident in the data. These two ways of understanding assessment resonated with the conceptual models of assessment discussed earlier. Hence, where appropriate, terms referred to in the literature review have been used to identify the relevant model. However, the analysis also illuminated the existence of both individual and contextual variation as a result of the situations within which assessment was practised.

For example, Charlotte (pseudonym) expressed an *assessment for learning* understanding and considered students' experiences and perspectives of assessment throughout. Her understanding appeared complex and multi-faceted; in the interview

she highlighted a range of tensions and described her understanding as provisional. In the map she distinguished between three 'purposes' of assessment: 'giving feedback', 'enabling quality control' and 'awarding degrees', and two different categories of 'modes': 'formative' versus 'summative', and 'sudden death' versus 'continuous'. These were located at the top of the hierarchy. In the interview, Charlotte highlighted the importance of the purposes of assessment and who assessment is for. She stressed the role of formative assessment and feedback and regarded good assessment as continual, i.e. 'happening perhaps little and often all the way through', and involving peers. She also considered what she called the 'human cost', namely the fact that assessment can cause considerable stress and anxiety amongst students, and the importance of maintaining student dignity. As a teacher, assessment also informed her as to whether her teaching had been successful.

In contrast, for Olaf, students' perspectives featured much less and his understanding was much more akin to the assessment of learning model of assessment. In the interview, Olaf focused predominantly on summative assessment and marking and his understanding of assessment seemed to be based on measurement. Within this model, reliability of measurement was seen as crucial. Therefore Olaf devoted considerable attention to discussing the clarity, explicitness and transparency of assessment requirements. More specifically he referred to the importance of standards and criteria for summative assessment, which related to an assessment rubric he had developed (the artefact he brought). This was described as allowing him to achieve consistency of marking and providing detailed feedback for summatively assessed work, which he mainly used to justify marks. However, assessment for learning was not absent from Olaf's thinking as his map included both assessment for learning and

assessment of learning concept labels which were both discussed in the interview. Olaf's map was divided into two parts, with assessment for learning on the left and on the right and several link lines between the two sides of the map. Assessment for learning and assessment of learning were both linked directly to the concept of 'organisational strategy'. This appeared to be an important concept as seven link lines were used to connect it with various other parts of the map.

Anne also exhibited a varied understanding, which related to different models of assessment and was influenced by what she described in the interview as the 'type of audience' being catered for. One audience was a student group undertaking a professionally accredited undergraduate programme of study, the other was a group of international students on a one-year direct entry programme. She explicitly referred to constructing her entire concept map on the basis of these audiences. Each had different assessment practice labels associated with them and these practice labels were related to quite different concept labels (see Figure 1), which created a map with two separate parts. On the side that related to the professionally accredited programme the assessment practice was a traditional, closed book exam and this linked to concepts such as 'Summative Assessment', 'Assessment of Learning' and 'High Stakes'. However, the side of the map that related to the direct entry programme had a practice label that referred to a specific approach to providing formative feedback. In response to difficulties the students faced in doing a project, Anne had designed a practice whereby the students submitted the first part of the project to receive feedback before they completed the second part. This practice was linked to the concepts of 'Assessment for Learning', 'Formative Assessment' and 'Low Stakes'. There were no link lines between either side of the map.

Figure 1 here

As illustrated above, the analysis showed that contrasting aspects of assessment were represented simultaneously, but that individual participants attributed different emphases and importance to them. Taken together the map and the interview tended to convey a particular way of thinking about assessment in which certain facets dominated, but it was difficult to categorise an individual's understanding as adhering exclusively to one specific model of assessment. This may be due to what Boud (2000) has termed the 'double duty' of assessment. Therefore it should be considered whether binary oppositions using categories such as 'transformational' and 'reproductive' (Postareff et al. 2012) do justice to the multiple purposes of assessment and the resulting complexity of understanding. In addition, the analysis suggests that variation in understanding could be explained by the different practices being used. Individuals did not seem to think about assessment in an abstract way, but in terms of particular situations or assessment related activities. This echoes the work of Marton and Pong (2005) and McAlpine et al. (2006).

ii. Alignment between Understanding and Practice

Virtually all participants engaged in practices that were closely aligned with their understandings. In fact, thinking and practice were often so intricately linked that they were difficult to separate. Philip, for instance, reported a formative practice that he termed 'practical self assessment workshops', three of which occurred over the course of one module he taught. The workshops were made up of a number of different stations, each with associated cases, prompts and questions, which increased

in complexity during the course of the module. Students worked in pairs and spent ten minutes at each station using a form (one of the artefacts brought along) to reflect on their own performance and the performance of their peer. The form required them to rate their confidence with the task, identify strengths and weaknesses and develop an action plan to address any weaknesses.

Philip's understanding expressed in the map and the interview was closely related to this practice. The map excerpt below shows that the practice label 'self-assessment workshop' was linked to a number of concepts commonly associated with assessment for learning (Figure 2).

Figure 2 here

Within the interview, Philip explicitly referred to the student perspective and expressed an understanding that closely aligned with assessment for learning:

> The assessment has to be for students. I feel it's got to be meaningful, encouraging, authentic, challenging, and I also feel quite strongly that it should be aligned, students should be clearly able to see where the learning outcomes and the assessment kind of fit together, (…) It's very important that things are clear and transparent for them. (…) Assessment isn't just assessment, it's assessment and feedback, because I think that the two elements are really important. (Philip)

This understanding was also reflected in Philip's description of the workshops in that they offered students non-threatening formative opportunities to scaffold their learning and encourage deep learning. His intention was for students to revisit and consolidate their learning from work they had done earlier in the module by providing structured feedback that could be used to further develop competence and confidence.

Such a finding confirms previous research, which has shown a close relationship and consistency between assessment practice and understandings of assessment (Samuelowicz and Bain 2002; Postareff et al. 2012). Our research participants were asked to identify where their practice sat in relation to their understanding on the map and discuss it as an example of a particular concept or several concepts. It could therefore be argued that methodologically this task already presupposed a relationship between understanding and practice. Conceptually, understanding and practice are equally difficult to separate. In the maps, the colour, not the content and level of concreteness, tended to distinguish concept labels from practice labels, whilst some labels could equally be regarded as denoting either understanding or practice. This is evident in the map excerpt in Figure 2.

iii.   *Complexity of Assessment Practice*

In the examples provided by some of the participants, the complexity of their assessment practice was striking. At times the descriptions were so detailed and intricate that it made them difficult to unpick and summarise concisely. Such complexity was particularly evident for those assessment practices that were heavily integrated with teaching, resulting in assessment and instruction being virtually inseparable. This has been conceptualised as a core characteristic of assessment for

learning (Wiliam 2011) as well as a key dimension in which conceptions have been found to vary within empirical research (Watkins, Dahlin and Ekholm 2005).

Esme was a lecturer in initial teacher education and trained students to become teachers in primary education. One of her intricately designed assessment practices was where the students had to work in small groups in order to produce a range of teaching activities, which they then enacted to the rest of the students on the course as if they were their pupils. During and after the delivery of the activity their peers used a sheet to provide anonymous feedback, which included two strengths and an area for improvement. This feedback modelled school teaching practice where teachers provide 'two stars and a wish' when marking pupils' work. At the end of the session the group who delivered the activity had a 30-minute group discussion facilitated by the tutor. Within the discussion they self-assessed against the criteria and considered feedback from their peers. The group was encouraged to create development points to support them in producing the activity for a subsequent week. Such a design provided an example of how formative and summative assessment, peer and tutor feedback, group and self-assessment were all integrated, with the intention to facilitate learning and improvement. In the interview Esme described the dialogic approach she used to facilitate this peer and self-review process in considerable detail. This approach contained many of the characteristics of the practical self-assessment workshops developed by Philip, already discussed above. Both represent highly complex practices in the 'spirit' of assessment for learning, which follow a longitudinal design and involve peer, tutor and self-assessment.

Other authors have reported understandings of assessment akin to assessment for learning in terms of their focus on students' thinking processes (Postareff et al. 2012), the integration of assessment and teaching (Watkins, Dahlin and Ekholm 2005) and the role of feedback as a formative tool for challenging and improving student understanding (Samuelowicz and Bain 2002). However, previous consideration and categorisation of the associated practices has tended to be rather broad and abstract. For example, Postareff et al. (2012) simply refer to 'alternative practices' and do not consider the finer grained details of such assessment practices. In contrast, the data generated by this study start to shed light on the considerable complexity and attention to detail involved in assessment for learning practices.

iv. *Construction of Understanding*

In some instances we saw evidence of the construction of individuals' understandings of assessment during the data collection process. This construction came about for individuals in two slightly different ways. The first way was that the dialogic concept mapping itself helped participants to refine, foreground and develop their understanding throughout the process. This was particularly explicit for one participant, Paula. Paula indicated that she had come to realise that reliability and validity were at the heart of her understanding of assessment and such a view seemed to be foregrounded as a consequence of creating a map and the dialogue about the map. The following interview extract illustrates this:

> …Something that links all the way through is that…the assessment must be
> reliable and valid and that then links back […]. When I look at it as the whole
> thing now it's probably the most important thing…because it relates to
> everything, it doesn't matter what you're measuring, it doesn't matter how

you're measuring it, it doesn't matter what it's defining, if what you are doing is not reliable and is not valid then the whole thing is shot. And it wasn't the first thing that I thought of by any stage, I was thinking of what you're measuring, how you're measuring it and what you're leading to. (Paula)

The second way in which individuals constructed their understanding of assessment was that the mapping and dialogic process encouraged them to think about their practice. In itself this appeared to evoke new ways of understanding assessment based upon the practices they were using. This was particularly apparent in Pierre, and the following extract gives insight into a formative activity used in order to develop the students' ability to analyse case studies and oral presentation skills.

They have to bring practical examples in and … every single lecture I do, we look at case studies. I don't actually say to them 'this is for your assessment', but I get them used to drawing out the information from case studies. So we looked at three different case studies, they all split into twos and threes and … they all came back and they talked about it and told the other group what it was about. (Pierre)

What was of most interest is that this practice had not previously been evident in his concept map or understanding of assessment. As comes through in the extract below, Pierre was actually (re)constructing his understanding in light of this practice as he spoke and tried to amend his map.

I don't know how to write it, I don't know the right word for it, but it's … whatever I'm gonna assess them on in their assessment I'm building that into

every single lecture. […]. I'm sort of not doing it openly 'cos I don't want them to say this is great for your assessment and it's not unconscious. So what is it, just sort of continuing hidden, that doesn't make sense but you know what I mean, hidden development of skills applied in assessment. [Writes on map]. So that would be something like, group work and presentations and analysis of case studies. (Pierre)

As he spoke, Pierre added some concept labels to his map, which provided an insight into the construction of his understanding of assessment. In the absence of a pre-existing concept, Pierre created the label 'continuing "hidden" development of skills'. This new concept was added to the map to form a new proposition regarding what assessment should be, which is depicted in Figure 3. Therefore the relationship between understanding and practice appeared to be iterative and multi-directional.

Figure 3 here

We have already drawn attention to the fact that the relationship between conceptions and practice or between professional knowledge and doing is contested. Some studies have confirmed that thinking drives approaches to teaching (Trigwell and Prosser 1996) and to assessment (Samuelowicz and Bain 2002), whilst others argue that practice helps shape conceptions, also with respect to teaching (Eley 2006) as well as assessment (Offerdahl and Tomanek 2011). The particular approach taken in the current research helps to illuminate the 'constructedness' of understandings of assessment, their intricate connection to practice and the multi-directionality of the relationship.

Conclusion

The findings above indicate that there is considerable variation in the teachers'
understandings of assessment and to a certain extent this confirms prior empirical and
theoretical work regarding contrasting understandings of assessment. However, what
the present study has unearthed is the complexity of the thinking and practice in
relation to assessment. We have shown how understandings are foregrounded,
adjusted, shaped or developed based on the experience of using certain practices and
even the concept mapping approach itself. Therefore a focus on attributing one single
conception to an individual may not do justice to the phenomenon under
investigation.

It is possible that the presence of multiple understandings may be specific to
assessment and sets it apart from other conceptions held by HE staff, as found for
example in research on conceptions of teaching. We would suggest that the neat
congruence between the conceptions of teaching and conceptions of assessment,
which is implicitly suggested elsewhere (Samuelowicz and Bain 2002; Postareff et al.
2012) may not necessarily be appropriate. One reason for this could be the strong
regulatory framework within which assessment for certification operates in HE. This
results in differing demands, and the many drivers which staff have to respond to
could be responsible for individuals exhibiting multiple and at points seemingly
contradictory understandings of assessment. Such an argument is related to the
'double duty' of assessment. Boud (2000, p160) argues that 'if we do not pay
attention to (…) [the] multiple purposes of assessment activities we are in danger of
inadvertently sabotaging one or more of them'. Hence applying a conceptual change

model as advocated for the development of teaching (e.g. Ho et al. 2001) to the development of assessment practice may not do justice to the double duty of assessment.

The study has also drawn attention to the value of drawing out the complexity of assessment practice. We would argue that it provided a more in depth insight into individuals' understandings of assessment and is important for supporting the enhancement of assessment in higher education. Firstly, the data on assessment practice more fully exposed the individuals' understanding of assessment by providing clearer meaning in relation to rather abstract concepts such as formative assessment or assessment for learning. In other words, an individual's understanding of assessment is not purely about knowing or espousing a particular concept but how it is enacted. This has clear parallels with McAlpine et al.'s (2006) model and Marshall and Drummond's (2006) work that through analysis of classroom practice observed assessment for learning activities that embodied the 'spirit' and those that simply conformed to the 'letter' of assessment for learning. Therefore without some insight into individuals' practices it is difficult to be sure that the meaning of a particular concept is the same for different individuals. Secondly, as proposed by Wiliam (2011), if we want to enhance assessment in higher education, we need to gain a better understanding of those practices where assessment and instruction are closely integrated. The analysis in the current study gives an indication of what is for different individuals in practice. What becomes clear from this analysis is the challenge of intricately weaving assessment for learning into courses in a longitudinal manner.

There are, however, obvious limitations with studies that try to gain fine-grained understandings of thinking and practice. As this study illustrates, key issues are convenience sampling, the low numbers of cases and that the data are highly specific to the context, limiting the level to which they are representative more widely. Despite this we would argue that such approaches are important for fully understanding assessment practice based on empirical evidence and also highlighting discrepancy between theory and practice. Carless's (2015) study using in-depth observations and interviews with five award winning teachers is recent example. Here the design and implementation of assessment tasks were found to be intricate and ingrained in the subject in a similar way to some of the practices in the current investigation, for example in Esme's case. However, the focus on the students' evaluative expertise and student engagement with feedback in the reported practices of our group of participants appeared to be far more varied both in terms of its presence and nature than those in Carless (2015). More studies of this type in a variety of context and with a range of participant groups are needed. It is likely that we will start to see commonalities across in-depth studies with small samples, which will allow for translation into different situations to become more credible and support the development of assessment that improves student learning more broadly.

Finally, the current study also makes a methodological contribution. The use of dialogic concept mapping and artefacts has provided an approach that helped research participants to construct and depict a complex, personal understanding of assessment. It could be argued that it has acted as a generative tool that enabled the exploration of the interplay between practical wisdom about assessment and public knowledge, as conceptualised by Boyd and Bloxham (2013). The approach used in this study has

obvious practical implications for academic development in terms of supporting staff

to discuss and develop their understandings of assessment and their assessment

practices, thus generating professional learning about assessment. In addition, to date

the methods used to investigate understanding of assessment in higher education are

limited in their breadth and there is the potential for over-simplification by purely

relying on participants being able to 'tell you' their understanding. Although only

one of many, concept mapping as an approach has offered a new perspective for this

line of research and as a supplement to interviews has helped us to capture some

complex conceptual constructs. It is also important to acknowledge that concept

mapping is not without limitations. Fundamentally it is still a form of self-report and

therefore, in a similar way to interviews, open to communication of espoused theories.

Also it is a relatively time-consuming approach, in terms of first guiding participants

through the process of producing a map and map construction itself. This highlights

the need to further develop methodologies and methods that allow us to research the

complexity of understandings of assessment and real world assessment practices.

References

Argyris, C., and Schön, D. A. 1974. *Theory in practice: increasing professional
effectiveness*. San Francisco, Jossey-Bass.

Boud, D. 2000. "Sustainable assessment: rethinking assessment for the learning
society". *Studies in Continuing Education*, 22(2): 151–167.

Boyd, P., and Bloxham, S. 2013. "A situative metaphor for teacher learning: the case of university tutors learning to grade student coursework". *British Educational Research Journal*, 40(2): 337-352.

Carless, D. 2015. "Exploring learning-oriented assessment processes". *Higher Education*, 69(6): 963-976.

Davies, M. 2011. "Concept Mapping, Mind Mapping, Argument Mapping: What are the Differences and Do They Matter?" *Higher Education*, 62(3): 279-301.

Devlin, M. 2006. "Challenging accepted wisdom about the place of conceptions of teaching in university teaching improvement". *International Journal of Teaching and Learning in Higher Education*, 18(2): 112-119.

Eley, M. 2006. "Teachers' conceptions of teaching, and the making of specific decisions in planning to teach". *Higher Education*, 51(2): 191-214.

Gipps, C.V. 2012. *Beyond testing. Towards a theory of educational assessment*. Classic edition. Abingdon and New York: Routledge. First published in 1994 by Falmer Press.

Hay, D, B. 2007. "Using concept maps to measure deep, surface and non-learning outcomes". *Studies in Higher Education*, 32(1): 39-57.

Hay, D. 2008. "Developing dialogic concept mapping as e-learning technology". *British Journal of Educational Technology*, 39(6): 1057-1060.

Hay, D, Kinchin, I., and Lygo-Baker, S. 2008. "Making learning visible: the role of concept mapping in higher education". *Studies in Higher Education*, 33(3): 295-311.

Ho, A. S. P., Watkins, D., and Kelly, M. 2001. "The conceptual change approach to improving teaching and learning". *Higher Education* 42 (2): 143-169.

Kane, R., Sandretto, S. and Heath, C. (2002). "Telling half the story: A critical review of research on the teaching beliefs and practices of university academics". *Review of Educational Research*, 72, 177-228.

Kinchin, I., and Hay, D. 2000. "How a qualitative approach to concept map analysis can be used to aid learning by illustrating patterns of conceptual development". *Educational Research Review*, 42(1): 43-57.

Lau, A.M.S. 2015. " 'Formative good, summative bad?' – A review of the dichotomy in assessment literature". *Journal of Further and Higher Education*. DOI: 10.1080/0309877X.2014.984600.

Liu, X., and Hinchey, M. 1996. "The internal consistency of a concept mapping scoring scheme and its effect on prediction validity". *International Journal of Science Education*, 18(8): 921 – 938.

Marshall, B., and Drummond, M.J. 2006. "How teachers engage with Assessment for Learning: lessons from the classroom". *Research Papers in Education*, 21(2): 133-149.

McAlpine, L., Weston, C., Timmermans, J., Berthiaume, D., and Fairbank-Roch, G. 2006. "Zones: reconceptualizing teacher thinking in relation to action" *Studies in Higher Education*, 31(5): 601-615.

Marton, F., and Pong, W.Y. 2005. "On the unit of description in phenomenography". *Higher Education Research and Development*, 24(4): 335-348.

Murray, K., and MacDonald, R. 1997. "The disjunction between lecturers' conceptions of teaching and their claimed educational practice". *Higher Education*, 33(3): 331-349.

Nicol, D. 2010. "From monologue to dialogue: Improving written feedback processes in mass higher education". *Assessment and Evaluation in Higher Education*, 35(5): 501 -517.

Novak, J. D., and Cañas, A. J. 2008. *The Theory Underlying Concept Maps and How to Construct and Use Them*. Technical Report, 2006-01 Rev 01-2008. Florida: IHMC CmapTools, Institute for Human and Machine Cognition. Available at: http://cmap.ihmc.us/Publications/ResearchPapers/TheoryUnderlyingConceptMaps.pdf [last accessed on 4 December, 2015].

Offerdahl, E.G, and Tomanek, D. 2011. "Changes in instructors' assessment thinking related to experimentation with new strategies". A*ssessment and Evaluation in Higher Education*, 36(7): 781-795.

Postareff, L., Virtanen, V., Katajavuori, N., and Lindblom-Ylänne, S. 2012. „Academics' conceptions of assessment and their assessment practices". *Studies in Educational Evaluation*, 38(3/4): 84-92.

Price, M., Carroll, J., O'Donovan, B., and Rust, C. 2011. "If I was going there I wouldn't start from here: a critical commentary on current assessment practice". *Assessment & Evaluation in Higher Education,* 36(4): 479-492.

Prosser, M., Trigwell, K., and Taylor, P. 1994. "A phenomenographic study of academics' conceptions of science learning and teaching". *Learning and Instruction*, 4(3): 217-232.

Sadler, I. 2012. "The influence of interactions with students for the development of new academics as teachers in higher education". *Higher Education*, 64(2): 147-160.

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

Säljö, R. 1997. "Talk as data and practice – a critical look at phenomenographic inquiry and the appeal to experience". *Higher Education Research and Development*, 16(2): 173-190.

Sambell, K., McDowell, L., and Montgomery, C. 2013. *Assessment for Learning in Higher Education*. Abingdon and New York: Routledge.

Samuelowicz, K., and Bain, J.D. 2002. "Identifying academics' orientations to assessment practice". *Higher Education*, 43(2): 173-201.

Taras, M. 2005. "Assessment – summative and formative – some theoretical reflections". *British Journal of Educational Studies*, 53(4): 466-478.

Torrance, H. 2007. "Assessment *as* learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning". *Assessment in Education*, 14(3): 281-294.

Trigwell, K. and Prosser, M. 1996. "Congruence between intention and strategy in university science teachers' approaches to teaching". *Higher Education*, 32(1): 77-87.

Watkins, D., Dahlin, B., and Ekholm, M. 2005. "Awareness of the backwash effect of assessment: A phenomenographic study of the views of Hong Kong and Swedish lecturers". *Instructional Science*, 33: 283-309.

Wiliam, D. 2011. "What is assessment for learning?" *Studies in Educational Evaluation*, 37: 3-14.

Figure captions

Figure 1. Concept map extract from Anne. The dashed line has been added to illustrate the two separate sides of the map (greyscale boxes indicate practice labels).

Figure 2. Concept map extract to illustrate the link between Philip's understanding and practice (greyscale boxes indicate practice labels).

Figure 3. Concept map extract as constructed by Pierre during the interview process (greyscale boxes indicate practice labels).