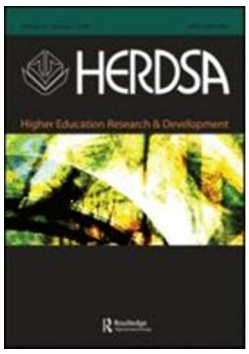

Variation in the development of teachers' understandings of assessment and their assessment practices in higher education.

Journal:	*Higher Education Research & Development*
Manuscript ID	CHER-2016-0491.R1
Manuscript Type:	Article
Keywords:	Academic development, Concept mapping, Assessment-for-learning, Conceptions

Variation in the development of teachers' understandings of assessment and their assessment practices in higher education

Abstract

This paper reports a study into the development of staff understanding of assessment and assessment practice. Eight teachers from two universities constructed an initial concept map about assessment that was discussed in a one-to-one semi-structured interview. A year later a new map was created and the interview focused on change in thinking and practice. Multiple models of assessment were evident in the participants' understandings at the same time and change was characterised by subtle evolution in thinking. Development in practice was more significant and often associated with the foregrounding of assessment-for-learning. Vignettes are used to illustrate the variation in nature and scale of development. Interplay between this development of practice and understanding was multidirectional and external context played an important role. The approach offers detailed insight into the relationship between assessment thinking and practice and demonstrates that both research and academic development need to go beyond conventional approaches to conceptualising the development of academics and take account of the finer grained complexities of assessment thinking and practices.

Keywords: conceptions, assessment for learning, concept mapping, academic development

Introduction

The theme of assessment and feedback has recently received considerable attention in the higher education (HE) literature. Due to persistent student dissatisfaction there is pressure on institutions and the sector as a whole to enhance assessment practice. In parallel several authors have urged us to reconceptualise assessment and feedback and change assessment practices in order to foreground learning as a key aspect of the process (e.g. Boud and Molloy 2013; Sambell et al. 2013). Although studies examining HE teaching have highlighted the importance of academics' conceptions for their practices (e.g. Trigwell and Prosser 1996), there is relatively little corresponding empirical research into how HE teachers understand assessment and their associated assessment practices. Existing work (Postareff et al 2013; Samuelowicz and Bain 2002) identifies broad categories of understanding, however, very little is known about the ways in which understandings of assessment and practices develop over time. Formal academic development tends to be aimed at early career academics, but does not necessarily focus on assessment, while academics' beliefs about assessment acquired through experience are likely to consolidate over time and be less amenable to change. Based on the significant implications of assessment for learning and certification, further understanding of the development of teacher thinking is a key aspect for enhancing assessment practice. The current study provides insights into such development.

Literature review

In the conceptions of teaching literature (e.g. Trigwell and Prosser 1996) it is inferred that development involves coming to understand teaching as more learning-orientated and student-centred, however there is limited empirical data on which to base this

assumption. Only a handful of studies have monitored teacher change over time in relation to their conceptions and approaches to teaching (Martin and Ramsden 1992; McKenzie 2002) and the influences upon their development (Sadler 2012). Generally, these studies suggest a change towards more student-centred practices, but this appears highly varied and context dependent. There is also conflicting evidence for the impact of academic development activities for such change (Stes et al. 2010), and it has been argued that academics predominantly develop as a result of everyday non-formal workplace learning (Knight et al. 2006). Another point of debate in this literature is the extent to which conceptions guide practice or practice stimulates conceptual change (Devlin 2006; Sadler 2012). Eley (2006) has questioned the existence of a strong directional influence of conceptions on practice. Based on research with school teachers, Guskey (2002) argues that development is predominantly an experiential process where teachers need to experience that their changes in classroom practices improve student learning outcomes before changing their beliefs and attitudes. However, other research with school teachers found that changes in the cognition of teachers were much more common than changes in actual teaching behaviour, with 'friction' being a key influence for change in practice (Bakkenes et al. 2010). This resonates with Pickering's (2006) finding that change in novice lecturers was influenced by a 'disturbance of beliefs' (p.325). Hence it is clear that research on teacher development is characterised by conflicting evidence, however, it would not seem unreasonable to assume that some of the findings might be relevant to thinking and behavioural change in relation to assessment.

Current thinking about assessment tends to adhere to constructivist conceptualisations and emphasises the importance of formative assessment (Sadler 1989) and student self-regulation (Nicol and Macfarlane-Dick 2006). Key aspects of

such a model, usually referred to as *assessment for learning* (AfL) (Wiliam 2011; Sambell et al. 2013) or *learning-oriented assessment* (Carless 2015), are that assessment should be an integral component of instruction, located within collaborative learning environments that engage students as active participants in the assessment and feedback process, foster meaningful, authentic engagement with the discipline, and support students in the development of evaluative expertise. This contrasts with conventional models that cast students as passive recipients and focus on end-point testing, psychometric measurement, certification and accountability (Gipps 2012), also referred to as *assessment of learning* (AoL). However, it is important to be aware that teachers have understood and used the notion of AfL in quite different ways. Marshall and Drummond (2006) distinguished between classroom practices in the 'spirit' of AfL, which embrace the underlying principle of developing student autonomy, and in the 'letter' where only techniques associated with AfL are in place. Torrance (2007) found interpretations of AfL that involved a narrow focus on criteria and instrumentally fulfilling requirements, to the extent that the assessment became the learning experience *per se.* Boud (2000) reminds us that due to its conflicting purposes, assessment always has 'double duty'. This has been further extended by authors such as Lau (2015) who argues that the literature has unintentionally created a harmful dichotomy between formative and summative assessment and that there is a need to re-connect the two as they are integral to one another. Similarly, Carless (2015) indicates that well-designed, learning-orientated summative assessment provides opportunity for the development of associated formative assessment strategies (e.g. peer learning and self-evaluation).

There is limited empirical evidence for the extent to which teachers' understandings relate to these models of assessment. Postareff et al. (2013) have found a hierarchy of conceptions of assessment ranging from 'reproductive' to

'transformational' and these tended to correspond with 'traditional' or 'alternative' assessment practices, respectively. However, few details are provided about what exactly traditional and alternative practices consist of and the way in which they may be influenced by variations in conceptions. The categories in Postareff et al. (2013) align with Samuelowicz and Bain's (2002) earlier work on academics' orientations to assessment practice and both studies suggest congruence between academics' understanding of assessment and their conception of teaching. However, there are obvious limitations of research that relies solely on interviews and other forms of self-report, as these are liable to elicit espoused theories rather than actual practices. This was exposed in assessment-based research with school teachers where observations provided evidence of disjuncture between beliefs and practice (Dixon, Hawe, and Parr 2011) and has also been seen in research specific to teaching in HE (Murray and MacDonald 1997). Therefore a key issue in the design of research in this area is how to best gain insights into staff thinking and their practices in assessment.

Only a small body of literature has considered the development of conceptions and practices in relation to assessment to date and it is quite diverse in terms of its focus, approach, contexts within which it has been investigated and whether it involved a formal intervention. Some of the findings from these studies echo the work specific to the development of conceptions of teaching, whilst others go beyond it. Within a school context, Dixon and Haigh (2009) identified changes in understanding formative assessment to varying degrees and related this to how teachers viewed their role and the role of the students in relation to learning and assessment. Offerdahl and Tomanek (2011) discovered that experimentation with alternative assessment strategies produced some development towards more sophisticated thinking about assessment, but surprisingly this did not result in revision of future practices. Finally, the importance of

departmental communities and cultures for influencing the development of assessment thinking and practice in relation to marking a complex assessment task was a key finding in a study by Jawitz (2009). Such an idea regarding informal development would appear to be important, particularly in HE where initial and continuous professional development tend to be mixed and relatively unstructured. This is supported by Poskitt (2014) who suggests that as assessment is subject to multiple dynamic political, cultural and educational influences, informal everyday workplace learning is likely to play a significant role in shaping teachers' thinking and practice. Similarly, a review of 100 studies found that the impact of pre-service teacher education on assessment is either non-existent or weak (Xu and Brown 2016). Based on this the authors called for more evidence about the role of 'on-the-job' learning for developing understanding of assessment and suggest that greater contextual understanding may prove useful for educating teachers about assessment.

In summary, the literature on HE teachers' understanding of assessment and their associated practices is somewhat limited and mixed. Little is known about the ways in which they understand the different conceptual models that can be found in the assessment literature, teachers' development over time, and the relationship between understanding and practices within such development. This paper seeks to contribute to this emerging body of literature through reporting a fine-grained study which aims to understand the way in which HE teachers' understandings of assessment and their assessment practices develop over time, using dialogic concept mapping.

Method

Two rounds of data collection, one year apart, were undertaken with eight members of staff from two UK universities. The broad subject areas represented in the sample included health, business, psychology and education. There were a variety of levels of

experience, however the majority (6 out of the 8) had been teaching in HE for more than five years. Round 1 (reported in more detail by Authors 2016) consisted of 2-hour staff development workshops broadly advertised in the universities, in which participants were introduced to concept mapping and then each constructed a map of their personal understanding of assessment. Concept maps were used with the aim of providing a visual representation of understanding, making tacit and abstract knowledge visible and helping to assess conceptual change (Hay, Kinchin and Lygo-Baker 2008; Kinchin and Hay 2000). Participants were also asked to bring an 'assessment artefact' (i.e. self-selected representations of their assessment practice), which, alongside other examples of their assessment practice, was added to their concept map, using different coloured labels. Both the concept mapping and use of artefacts aimed to address the limitation of interviews by incorporating concrete examples of practices, which were also explicitly linked to concepts, as a way of generating richer and triangulated data. Workshop participants were informed of the opportunity to volunteer for the research component, and following the workshop research participants attended a one-to-one semi-structured interview that used their map and artefact(s) to discuss their personal understanding of assessment and reported assessment practices. This applied the principles of dialogic concept mapping (Hay, 2008) by incorporating an iterative process where participants had several opportunities to construct, discuss and review maps.

Round 2, which took place a year later, aimed to specifically identify change. Participants were invited to a second interview and given the same concept labels they had used in their round 1 map as a basis to create a new map; however it was emphasised that they could discard any unwanted labels used previously and add new ones. Interviews were conducted immediately following the production of the second map where participants were asked to first explain the reasoning behind their new map

and then comment on any differences or similarities they observed between the two maps presented side by side. This was followed by a discussion of their perceptions of the development of their understanding of assessment and their reported assessment practice, including consideration of a round 2 artefact.

Data analysis

Since we sought to understand whether and in what way understanding and reported practices had changed, analysis focused upon comparing round 1 and round 2 data for each individual. This was done in a broadly interpretivist, qualitative manner using thematic analysis and the phases identified by Braun and Clarke (2006), working both inductively and deductively. Initially this involved immersion in maps, interviews and artefacts and considering them in close conjunction with each other. As found in previous concept mapping research (e.g. Liu and Hinchey 1996), analysis of the maps was not straightforward as they varied considerably in complexity, structure and propositional content. Link lines were not always labelled and some maps therefore resembled hybrids between mind maps and concept maps (Davies 2011). However, the interviews helped to expand and shed further light on changes to personal understandings and reported practices encapsulated in maps and artefacts. Attention was paid to sections of the maps and interviews including new practices or highlighting changes to existing practice. The nature of development in understanding and practice emerged both implicitly (i.e. development identified by the researchers through comparison of round 1 and round 2 data, such as additional concept labels incorporated in round 2) and explicitly (i.e. development described or identified by the participants themselves in the round 2 interview).

First, the two researchers produced independent accounts of the data in which changes in thinking, changes in reported practice and the reasons for such changes were

identified for each individual. These were discussed to agree themes that then were checked and refined by comparison between participants, data sources (maps, interviews and artefacts) and across the two rounds of data collection. When comparing individual participants and changes between round 1 and 2, conceptions and practices associated with AfL emerged as significant for understanding difference and change over time. In the next step AfL was therefore applied as a theoretical lens to identify characteristics of AfL evident in the data, which allowed to capture variation between participants and in development. A final step in the analysis was to focus on the development of each individual participant in contrast to others. These resulted in vignettes which were developed to illustrate the uniqueness of individual development in greater detail and are reported below for selected participants.

Findings

Changes in understanding

Individuals' understandings of assessment were complex. In both rounds, all participants' maps and interviews demonstrated an awareness of both AoL and AfL. These contrasting models of assessment were often present simultaneously, albeit with differing emphases. There was limited evidence of significant development in understanding between round 1 and round 2. Maps in both rounds were extremely similar in terms of content and structure, and interviews contained explicit statements to the effect that there had been no change. Where change was evident, it tended to take the form of a subtle evolution or slight re-shaping of participants' original thinking. This involved the foregrounding, confirming, adjustment or re-alignment of an existing understanding, as the following example of one participant illustrates.

Philip: increased awareness of collaboration and dialogue

In round 2, Philip introduced and emphasised, both within the map and the interview, the concept of 'collaboration'. This was to foreground the importance and nature of dialogues for the way in which assessment supports learning. The following extract from round 2 makes explicit that this subtle change in thinking involved a specific concept coming to the fore, resulting in the development of a clarified understanding rather than a significant qualitative shift.

> The only (concept label) I added in was collaborative, I think I had dialogues but I think I'm perhaps appreciating more the importance to the students of dialogic feedback and them being able to talk through. But also the fact that they're working together on tasks that relate to their assessment, then the collaborative work that they do they are developing trust with each other and getting quite a lot of feedback […] to where they need to be for the future. Although I was using that in practice I perhaps wasn't so aware of the importance of the sort of collaborative development and the dialogic element and I'd incorporated this year I think possibly more. I haven't really incorporated more activities but what I have tried to do is identify them. (Philip, round 2).

Changes in practice

Whilst understanding of assessment tended to remain relatively stable, changes to practices were more common. Development of practice was brought about by a variety of changes in the external context, including, for example, the reduction of classroom contact time, institutional policy initiatives such as on-line marking, external examiner comments, and the availability of new technologies. However, development of practice was not evident for all individuals and there was considerable variation in the nature and scale of such development, with some making adjustments to existing practices and others introducing completely new strategies. These changes either involved the adoption of new practices that embodied some characteristics of AfL, or the evolution of existing AfL practices by deepening and enhancing them further. Interestingly, even

in cases where AfL practices had already been in place, considerable development was still evident.

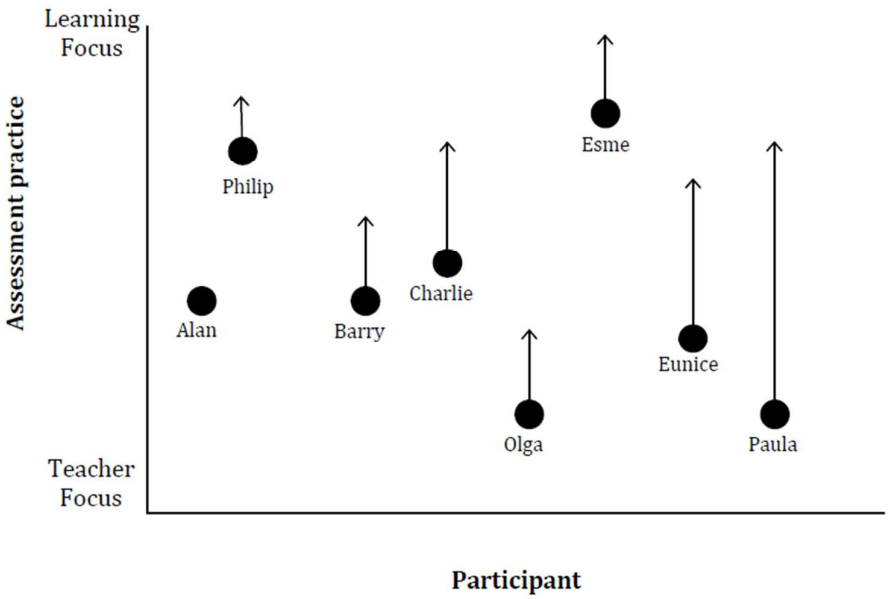

Figure 1. Schematic representation of the direction and scale of the development in assessment practices for all participants.

Figure 1 aims to represent the practice related changes visually for all research participants. The emphasis here is upon the extent to which the reported assessment practices became more focused upon student learning over time and further incorporated AfL principles and procedures. What needs to be made clear is that the figure should not be regarded as an attempt to quantify the qualitative differences that were found as AfL cannot be considered an absolute. However, the figure attempts to illustrate the variation between individuals. As the majority of teachers were already employing learning focused assessment practices at the time of round 1, considerable shifts in practice were uncommon. For some, there was no or little evidence of change in practice (e.g. Philip), whilst others developed considerably (e.g. Paula). Also of

interest was the further development of individuals already engaged in AfL practices (e.g. Esme). Details are discussed in the vignettes provided below.

Vignettes of changes in practice

The following vignettes illustrate patterns in the development of assessment practices by using three participants as examples. This aims to provide fine-grained insight and illuminate contrasting ways in which practices changed in relation to AfL. In addition, attention has been given to the interplay between understanding and practice throughout each individual's development.

Esme: enhancing formative assessment dialogue through technology

Esme's vignette illustrates the way in which key characteristics of AfL are present in her thinking as well as her practice and how they have been developed further. Esme trains students to become teachers in primary education in the UK. She uses very complex, student focused assessment practices that are highly integrated into teaching. Such understanding and practices were clearly apparent in round 1. However, in round 2 there has been a subtle development in her thinking, from feedback being the central concept in round 1, to emphasising the integral role of dialogue and interaction with students to form shared judgments that lead to student improvement (Figure 2). There are clear parallels here with the concept of AfL as being located within collaborative learning environments.

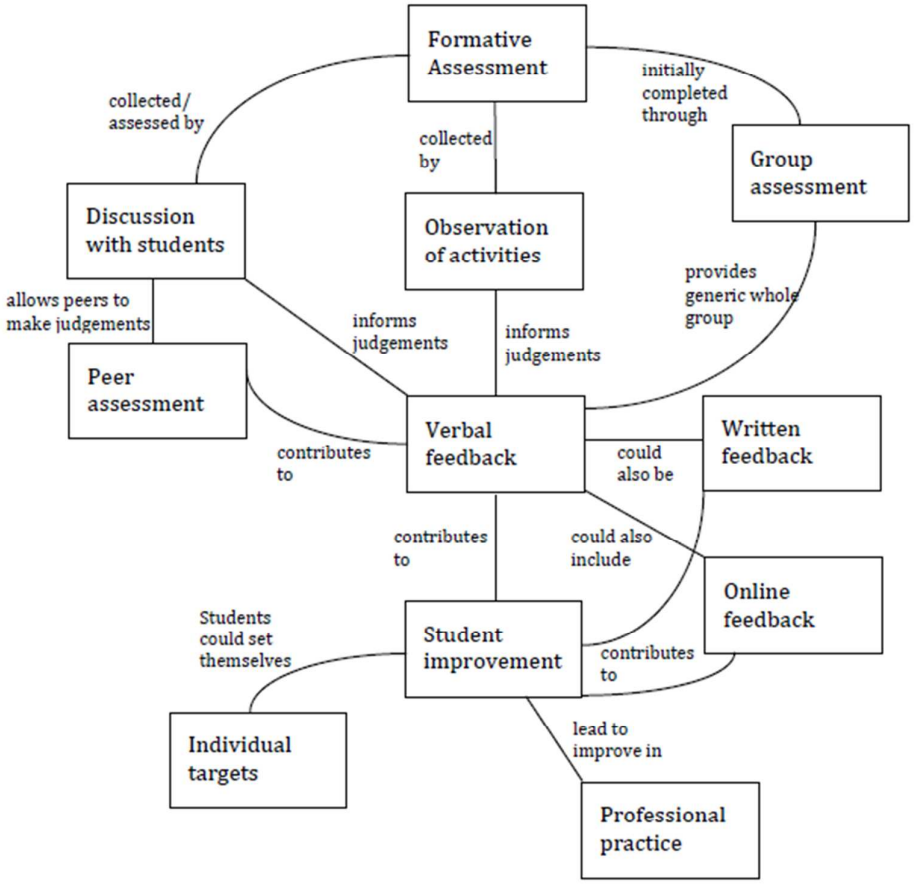

Figure 2. Map excerpt to illustrate Esme's understanding of AfL (Round 2).

The interview illustrates that Esme has thought more deeply about the circumstances that lead to student improvement. She has come to the conclusion that effective formative assessment is more than feedback as it should involve making judgments jointly with students and therefore needs to be verbal, dialogic and involve interaction:

> Formative assessment is most important, because formative assessment is on-going, it's that discussion, it's that dialogic face, (…) it's forming judgments together and I think if you do that, students have a better understanding and that leads to better improvements. (Esme, round 2)

Esme described a specific formative assessment practice, which changed between round 1 and 2 due to the integration of technology. The practice consists of small groups of students producing teaching activities and materials presented to their peers and on which the peers provide anonymous feedback. Following the activity, the group have a 30-minute group discussion facilitated by the tutor, during which they consider the peer feedback and self-assess against criteria. In round 1 peer feedback was paper-based. In round 2 technology was incorporated to further facilitate the peer feedback and self-assessment process: the group was now videoed when presenting the activity so that they could watch themselves afterwards, and rather than providing comments on paper, their peers took photographs of the materials using a tablet computer and annotated them electronically.

> They can add in text boxes (…) so it's almost like a dialogue pointing to things on the display rather than just saying 'this was good'. (…) They might say, 'we really like the way you have made the display interactive,…', or things like, (…) 'the writing is too small here, the children will find it difficult' so it can be critical as well as positive. (…) The students who did the display then get to look at it and (…) they almost like want to talk to the people who have given the judgment, because they want to discuss it, they want to say 'this is why we did it, how do you think....' (…). It's a lot better using the iPads because beforehand it was written down on a piece of paper and it just said ' I really like the interactive activity of this', but there was nothing identified because the display wasn't there anymore. (…) What we find is that it is actually improving their feedback to other students. (…) or feedback when they're marking work with the children. (Esme, round 2)

This example makes clear that such use of technology is in line with Esme's slightly evolved understanding of assessment. Through the use of technology she has further integrated assessment with instruction, engaged her students even more actively in the assessment process and provided them with additional opportunities to develop their

evaluative expertise. By doing so she has deepened several key features of AfL, which were already present in her round 1 practice.

Eunice: Developing a marking code and using it in formative peer assessment

Eunice's vignette demonstrates a qualitative shift towards AfL practices, described retrospectively by Eunice as having taken place before the first round of data collection. In both interviews, Eunice refers to the development of a marking code that was included as a practice label in both concept maps:

> I kept thinking: I keep writing the same comments! So what I did is I developed a set of codes. So now instead of writing 'don't reference like this, reference like this', (…) I just put a code on it, so it could be 'xrefst' and then I give the students a sheet: it tells you what the code means and it gives them an example. (Eunice, round 1)

Eunice stresses that she initially developed the marking code for her own benefit. However, introducing this new practice seemed to have stimulated a change in her thinking:

> Having done that for my own purposes I then thought (…) 'they don't need me, the students could use this to mark their own work and each other's'. (Eunice, round 1)

This practice of constructing and using marking codes appeared to influence Eunice's thinking, which in turn stimulated her to develop an approach that is more akin to AfL. She continues to describe the 'peer assessment exercise' (included in her maps as another practice label), in which students use the marking codes to mark each other's essays, each piece of work being marked by two peers. Eunice then took this further by providing examples of good and bad essay in the virtual learning environment (VLE).

The interview suggests that these are explained in class prior to the peer assessment exercise being carried out:

> I gave them examples of essays that I'd marked so they know what I would be doing. (…) showing them typical mistakes, even things like why it's best not to use too many quotes. (Eunice, round 1)

The peer assessment exercise and the exemplars generated another change in Eunice's assessment practice that took place between round 1 and round 2. In the round 2 interview, Eunice describes how she used grade descriptors and exemplars for each band. Once again, this was initially for her benefit as a marker of summative work, but then she moved to using them with the students in formative tasks. This time this change was initiated by what Eunice calls the 'wagging finger by the moderator':

> Originally I used this for my own purposes, so I would mark assignments and I'd think ok where would I then put the student they've done that, ok it's in that band and I did get pulled up a few times by moderators saying 'why don't you make it explicit to students by referring to your marking criteria 'cos they have it' and after a while of getting the wagging finger I thought well why don't I? Because in actual fact that's more explicit guidance. (Eunice, round 2)

In Eunice's vignette the relationship between her thinking about assessment and her practices is complex and multidirectional, as an initial change in practice appears to have led to a change in thinking, which in turn has triggered additional practice developments. There is evidence of a qualitative shift from assessment practices that were adopted for the benefit of the teacher towards practices that are more in line with AfL principles. Eunice has integrated assessment into instruction and the peer assessment exercise has the potential to develop her students' evaluative expertise. However, Eunice's practice did not engage with the 'spirit' of AfL in the same way as Esme's. For instance, much of Eunice's thinking still revolves around marking

summative work. The exemplars provided in the VLE do not appear to be actively used by students in the classroom, and the peer assessment exercise appears to focus on surface features such a referencing rather than promoting deeper, more authentic engagement with the discipline.

Paula: implementing formative peer review

Paula's vignette was chosen since it exemplifies a significant development in assessment practice, which did *not* seem to be accompanied by a corresponding change in understanding. Paula's thinking, which remained virtually unchanged between round 1 and 2, focused on assessment as a valid and reliable measurement tool whose main purpose is to differentiate between different levels of understanding and performance, to indicate success or failure and result in the award of a mark or qualification. Feedback was understood as comments on an end-point assessment that would be used to help with the summative assessment of a subsequent module. However, there was no consideration for the ways in which students would engage with such feedback, either before or following submission of their work.

In round 2, Paula described the introduction of a formative peer review activity she had not used previously. This new practice stands out as it seems incongruent with her understanding of assessment. The following extract is Paula's description of her development in which she confirms the change in her practice whilst highlighting the relative stability in her understanding:

> Things do evolve over time and new people come into the department with new ideas and you start to think slightly differently about how you should assess and what you should assess. We haven't made much in the way of wholesale changes over the last year but I think that you do evolve. I think my understanding of assessment and what we're trying to achieve through assessment is more or less the same. I think it's perhaps the understanding of how you achieve it that perhaps

> changes a little bit more because you think you're actually doing something, then somebody comes along and says 'yeah but that's x and y' and you go 'oh I hadn't really thought of it'. (Paula, round 2).

The formative peer review activity was introduced in a first-year group developing laboratory report writing skills. Rather than getting the students to submit five reports throughout the module that were all graded, the module was changed so that just one report was graded at the end, with the other pieces being formative. After laboratory sessions students had to submit a write-up of the session to the VLE and then in groups of four discussed and reviewed the work of the others in the group and provided peer feedback via the VLE. Two key factors appeared to have prompted this change. The first was due to the volume of marking and issues in terms of consistency of multiple markers. The second factor was Paula taking on a new module and re-designing the activity alongside a colleague. The idea of input from colleagues as an influence on Paula's practice is also hinted at more broadly in the extract above. Although there did not appear to be any 'wholesale' changes in understanding, some of Paula's comments suggest insights gained from the new practices: 'they (the laboratory write-ups) were probably as good as when they were getting the summative feedback'; 'the actual reading of three other people's pieces of work about the same topic was really valuable to them, to actually see how other people had done it'. The following extract also suggests that experiencing the practice has started to challenge some of Paula's pre-existing assumptions:

> I was very pleased, it could have been the case that because it was formative, because it was peer assessed, they might have just thought… We did have some concerns that in the first week that everybody would turn up and sort of; 'have you done that?' 'no'. But I think because the first semester, first year you really tell them to do this and they actually do it. If we tried to get some of the second years to do something like that you'd have no chance.

Like in Eunice's vignette, there is evidence of a qualitative shift towards AfL practices since the peer review activity incorporates student collaboration and aims to develop their evaluative expertise of a genre that is core to their discipline. Similar to the findings of Jawitz (2009), there are hints from Paula about the influence of peers and the community of practice. Such an influence might start to explain the curious mismatch between Paula's understanding and practice. The quote above suggests that the incongruence between understanding and practice may have resulted in some level of 'friction' (Bakkenes et al. 2010) or 'disturbance of beliefs' (Pickering 2006), which may have the potential to become an important catalyst for further development.

Discussion and Conclusions

This study offers insights into the ways in which HE staff thinking and assessment practice develops over time. Although some qualitative shifts were found, change tended to be small scale, particularly in relation to understanding. The vignettes provide evidence for change towards practices that were increasingly in line with AfL principles, however, the extent to which this was embraced and implemented varied considerably. In the vignettes of Eunice and Paula it can be argued that although there was a significant shift towards a focus on learning, not all aspects of AfL were attended to. It is also notable that some of the AfL practices outlined within the vignettes started off as practices originally implemented to enhance AoL from an assessor focused perspective. For example, the use of marking codes by Eunice and the reduction in number of summative submissions for Paula were seen to evolve into practices that more closely resembled AfL. Such transitions and transformations in practice seemed to be stimulated by colleague and student engagement with them. This could suggest that there is a possibility that adopting procedures in the 'letter' of AfL may have the potential, in the long term, to facilitate practice in the 'spirit' of AfL. Overall, it is

important to be aware that the development of assessment practice is likely to be staged and protracted and that change in understanding is particularly hard to achieve. This has implications for the work of policy makers and academic developers who need to allow time for enhancements to become embedded and evolve.

The idea that small-scale change in practice may be significant for an individual's understanding has been argued elsewhere in relation to the conceptions and approaches to teaching (Devlin 2006; Eley 2006). More specifically, Sadler (2012) found that practices that encouraged instances of interactions with students, regardless of the initial intention for the practice, acted as a key influence on teacher development. The current study indicates that how assessment thinking and practice interact with one another, in the process of development, varies considerably between individuals. In some instances a change in thinking stimulated a change in practice (e.g. Esme), whilst in others instances a change in practice stimulated a change in thinking (e.g. Eunice). Hence the relationship between thinking and practicing appears to be two-directional. This emphasises and extends the complexity that Offerdahl and Tomanek's (2011) study revealed and highlights the need for further research which captures the relationship between thinking and practice in development over time.

The congruence between the individuals' understanding of assessment and their practice was an interesting point of variation in the participants' development. While some of the changes in assessment practices were in line with individuals' thinking (e.g. Esme), others were not (e.g. Paula). Paula's case indicates that a purely conceptual change approach to educational development maybe an oversimplification and that the relationship between thinking and practising may be more complex than assumed, not only in relation to teaching but also to assessment. Paula's incongruence between a focus on AoL in thinking and the introduction of an AfL practice contains elements of

'friction' or 'disturbance of beliefs' as reported by Bakkenes et al. (2010) and Pickering (2006). The conscious processing of this incongruence might ultimately lead to a change in thinking, and there are some indications, as highlighted by the interview excerpts, that the newly adopted practices have started to challenge Paula's assumptions. This suggests that reflection on such frictions (e.g. instigated within the context of an in-service course) might be a possible avenue towards achieving change. Such an intervention could also form the basis of future research.

The finding that HE teachers who already have an AfL focused conception of assessment can still develop their assessment practice in significant ways, further highlights the limitations of the conventional approaches and frameworks for conceptualising and researching the development of academics in relation to assessment. Our study suggests that research needs to go beyond categorisations such as student versus teacher focused (Trigwell and Prosser 1996) or reproductive versus transformational (Postareff et al. 2013) and instead recognize, make explicit and investigate the finer grained complexities of assessment thinking and practices and their development over time. This also raises questions about methodologies and methods used in HE assessment research, and interview-based research in particular. The present study complemented interviews with concept maps and artefacts; Carless (2015) has recently demonstrated the potential of interviews complemented with observational data for advancing our understanding of assessment practices. This suggests that more of this kind of research is needed.

The relatively small level of change observed in the participants' understanding and practice may be related to the specific nature of the sample, the methods employed and the period of investigation. Although to varying degrees, the participants seemed to hold a relatively broad awareness of assessment from the outset, which for most

participants included both AfL and AoL perspectives. Therefore it could be argued that our data provide insights into the next stage of development in terms of understanding assessment. It would be reasonable to suggest that this is likely to be longer-term, less drastic in nature and about getting to grips with the nuances and all dimensions of AfL. A limitation of the study was the absence of data regarding the participants' academic development activities, either in the months prior to the initial data collection or in the year between the two data collection rounds. Future research should consider longer periods of time and focus on a wide range of staff including both early career and experienced academics.

References

Bakkenes, I., Vermunt, J.D. and Wubbels, T. 2010. "Teacher learning in the context of educational innovation: Learning activities and learning outcomes of experienced teachers." *Learning and Instruction* 20: 533-548.

Boud, D., and Molloy, E. 2013. "Rethinking models of feedback for learning: the challenge of design." *Assessment & Evaluation in Higher Education* 38 (6): 698-712.

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

Braun, V. and Clarke, V. 2006. "Using thematic analysis in psychology." *Qualitative Research in Psychology* 3: 77-101.

Carless, D. 2015. "Exploring learning-oriented assessment processes". *Higher Education.* doi: 10.1007/s10734-014-9816-z.

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: 112-119.

Dixon, H. and Haigh, M. 2009. "Changing mathematics teachers' conceptions of assessment and feedback." *Teacher Development* 13 (2): 173-186.

Dixon, H., Hawe, E., and Parr J. 2011. "Enacting assessment for learning: the beliefs practice nexus" *Assessment in Education: Principles, Policy and Practice* 18 (4): 365-379.

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

Guskey, T.R. (2002). "Professional Development and Teacher Change". *Teachers and Teaching* 8 (3): 381-391.

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: 295-311.

Jawitz, J. 2009. "Learning in the academic workplace: the harmonization of the collective and the individual habitus." *Studies in Higher Education* 34: 601-614.

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.

Knight, P., Tait, J., and Yorke, M. (2006). "The professional learning of teachers in higher education." Studies in Higher Education, 31(3): 319-339.

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.

McKenzie, J. 2002. "Variation and relevance structures for university teachers' learning: Bringing about change in ways of experiencing teaching, research and development." *Higher Education Research and Development* 25: 234-241.

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.

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

Martin, E. and Ramsden, P. 1992. "An expanding awareness: how lecturers change their understanding of teaching." *Research and Development in Higher Education* 15: 148-155.

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.J., and Macfarlane-Dick, D. 2006. "Formative assessment and self-regulated learning: A model and seven principles of good feedback practice." *Studies in Higher Education* 31 (2): 199-218.

Offerdahl, E. and Tomanek, D. 2011. "Changes in instructors' assessment thinking related to experimentation with new strategies." *Assessment and Evaluation in Higher Education* 36 (7): 501 -517.

Pickering, A..M. 2006. "Learning about university teaching: reflections on a research study investigating influences for change." *Teaching in Higher Education* 11 (3): 319-335.

Poskitt, J. 2014. "Transforming professional learning and practice in assessment for learning." *The Curriculum Journal* 25 (4): 542-566.

Postareff, L., Virtanen, V., Katajavuori, N. and Lindblom-Ylanne, S. 2013. "Academics conceptions of assessment and their assessment practices." *Studies in Educational Evaluation* 38: 84-92.

Authors. 2016. "Personal Understanding of Assessment and the Link to Assessment Practice: the Perspectives of Higher Education Staff." *Assessment and Evaluation in Higher Education.* doi: 10.1080/02602938.2016.1184225.

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

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.

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.

Stes, A., Min-Leliveld, M., Gijbels, D., and Van Petegem, P. (2010). "The impact of instructional development in higher education: The state-of-the-art of the research." *Educational Research Review,* 5(1), 25-49

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: Principles, Policy and Practice* 14 (3): 281-294.

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

Trigwell, K., Prosser, M. and Waterhouse F. 1999. "Relations between teachers' approaches to teaching and students' approaches to learning." *Higher Education* 37: 57-70.

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

Xu, Y. and Brown, G.T.L. 2016. "Teacher assessment literacy in practice: A reconceptualization." Teaching and Teacher Education 58: 149-162.

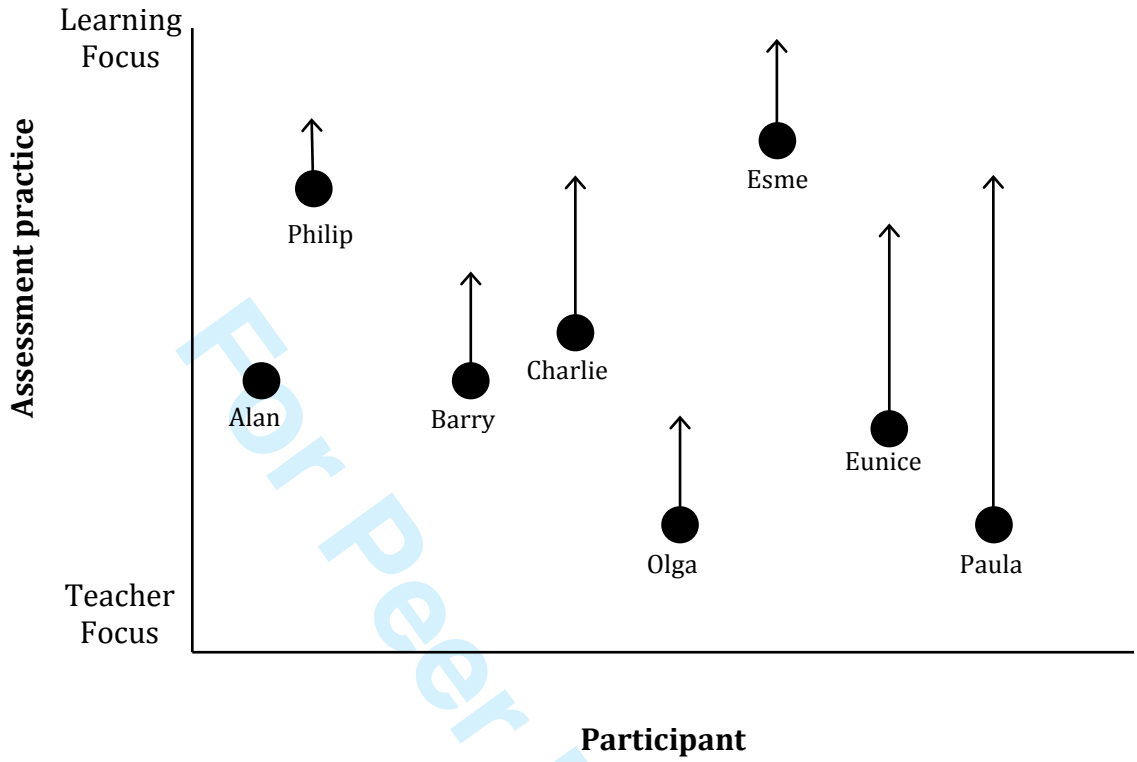

Figure 1. Development of assessment practices towards a focus on student learning.

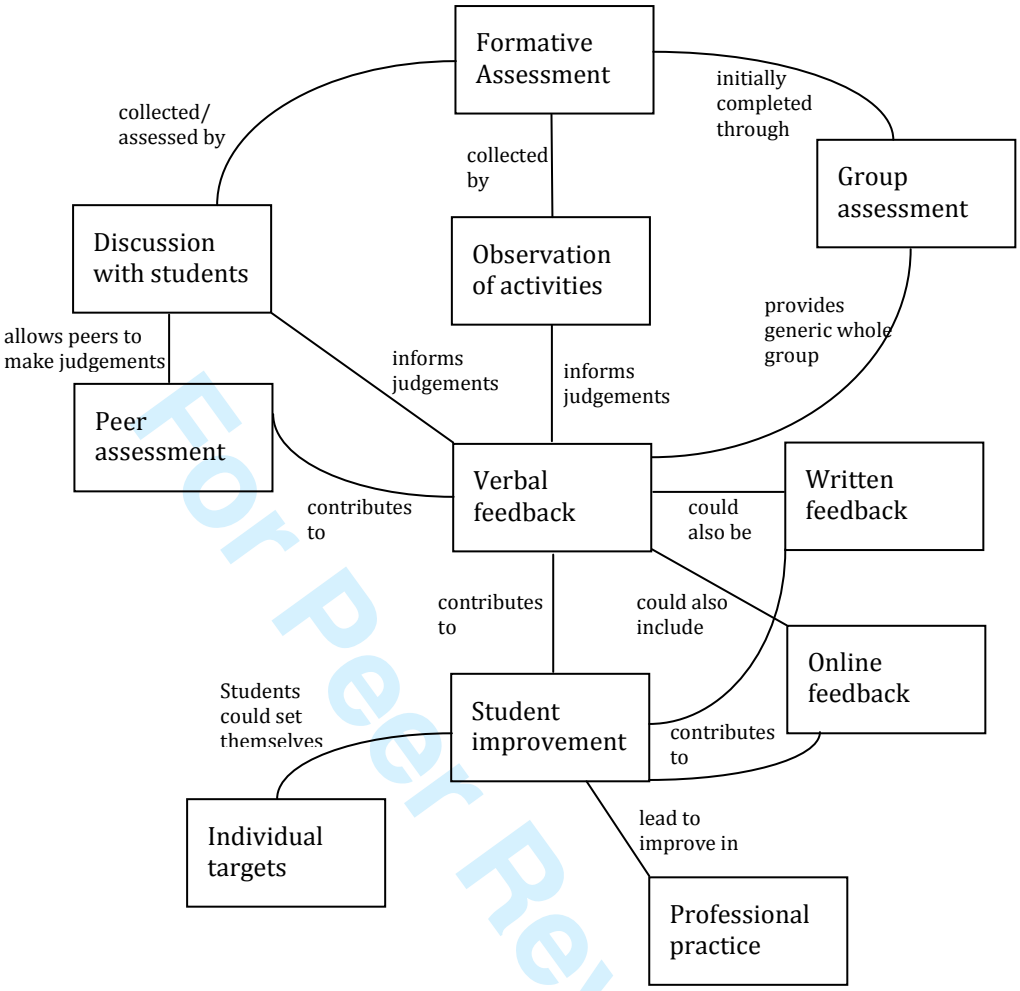

Figure 2. Map excerpt to illustrate Esme's understanding of AfL (Round 2)