*Studies in Higher Education*
*Vol. 32, No. 1, February 2007, pp. 39–57*

# Using concept maps to measure deep, surface and non-learning outcomes

David B. Hay*
*King's College, London, UK*

This article reports the use of concept mapping to reveal patterns of student learning (or non-learning) in the course of master's level teaching for research methods. The work was done with a group of 12 postgraduate students, and the concept maps of four individuals produced *before* and *after* a single teaching intervention are shown in detail. The data are presented as case studies that document the incidence of deep learning, surface learning and non-learning. These are terms that are widely used in the educational research literature, but most evidence for these learning approaches comes from students' conceptions of learning, not from empirical measures of changes in knowledge structure. Here precise criteria for defining change in terms of deep, surface and non-learning are developed, and concept mapping is used for assessment of learning *quality* using these criteria. The results show that deep, surface and non-learning are tangible measures of learning that can be observed directly as a consequence of concept mapping. Concept mapping has considerable utility for tracking change in the course of learning, and has the capacity to distinguish between changes that are meaningful, and those that are not. This is discussed in the wider context of learning, and teaching and research.

## Introduction

A better understanding of the ways in which people learn is a clear goal of educational research. That people learn differently is intuitive, and that people conceive of learning in different ways is phenomenologically demonstrable (see Säljö, 1975; Marton & Säljö, 1976, 1984; Marton, 1986; Entwistle, 1990; Entwistle *et al.*, 1991, 2001). Furthermore, there is now a large body of empirical research data to suggest that individual approaches to learning can have considerable impact on the outcomes they achieve (e.g. Entwistle & Tait, 1994), and that the assessment of learning 'style' may be a useful means of identifying students at risk as a consequence of their approach (e.g. Entwistle, 1990; Entwistle & Tait, 1994). Nevertheless, the literature that reports the evaluation of learning style assessment is mixed and often overtly critical

*King's Institute of Learning and Teaching, King's College, London, James Clerk Maxwell Building, 57 Waterloo Road, London SE1 8WA, UK. Email: david2.hay@kcl.ac.uk

ISSN 0307-5079 (print)/ISSN 1470-174X (online)/07/010039–19

DOI: 10.1080/03075070601099432

(see, for example, Garner, 2000; Coffield *et al.*, 2004a, b). This is largely because the field lacks a broad theoretical underpinning that is explanatory for the discipline as a whole (Sternberg, 2001; Coffield *et al.*, 2004a), and tangible evidence for change that can be defined in terms of recognisable 'styles' is lacking.

This is not to say that research on learning style has not done much to advance the debates about learning difference and educational practice. In the UK, for example, the distinction between deep and surface learning (see Säljö, 1975; Marton & Säljö, 1976) has had considerable impact on theory development and teaching practice, and has informed some of the most important texts for learning and teaching in higher education (e.g. Ramsden, 1992; Nicholls, 2002). The diagnostic approach of Entwistle (1990) is grounded in the deep versus surface approach, but it is criticised by Coffield *et al.* (2004a) because its predictive power has rarely been tested, and it may imply an overt tendency towards labelling individuals in ways that are further barriers to learning. Nevertheless, they also find much to praise in the approach and suggest that active engagement in debates about learning among teachers and students is likely to have considerable benefit for both. This is similar to the views of Garner (2000), who reviews the Kolb Learning Style Inventory (Kolb, 1999) and concludes that, though instrumental assessment of styles may be flawed, significant benefits can accrue as a consequence of ensuing discussion of different learning processes and approaches.

Despite these debates, it is probably that learning is such a complex and intractable process that has made its study difficult and contentious. A different (but complementary) contribution to the field is the work of Jarvis (e.g. Jarvis, 1992, 1998; Jarvis *et al.*, 1998). Jarvis (1992) provides a model of the learning process that is derived from hundreds of individual descriptions of learning. It has a number of similarities with Kolb's Learning Cycle (Kolb & Fry, 1975) because it suggests that all learning is rooted in experience and that it comprises steps of reflection, evaluation and experiment. It is, however, notably different in three important ways. First, it suggests that learning is always an individual process: that personal change is indicative of learning, and that non-learning is reinforcement or the absence of change. Second, it includes memorisation as one of the possible stages in the learning process, but suggests that memorisation on its own leads to non-learning if other steps in the learning processes are ignored. Third, it documents a variety of routes through the steps of learning that are not necessarily sequential (as they are in Kolb & Fry [1975]), but can be flexible, iterative and recursive. Thus, Jarvis's model (Jarvis, 1992) accommodates the descriptions of learning that are indicative of individuals with different learning approaches (or styles). Importantly, it also integrates non-learning as a potential outcome of any experience (including formal education), in ways that are defined either because the situation is ignored (presumption, non-consideration or rejection [Jarvis *et al.*, 1998]) or because the learning is trivial and does not lead to change.

Work that has been done to define and distinguish between meaningful and rote learning (e.g. Novak & Symington, 1982; Novak, 1998) complements the approaches of Jarvis. Like Jarvis (1992), the definition of meaningful learning (Novak, 1998) is

based on evidence of change in individual knowledge and understanding. Where Jarvis (1992), however, suggests only that learning can be measured by documenting change (and that the absence of change is non-learning), Novak (1998) suggests that change can occur in qualitatively different ways: ways that are indicative of meaningful learning and ways indicative of rote learning. This is largely synonymous with the distinction between deep and surface approaches (and also memorisation as a route to non-learning [Jarvis, 1992]), but here it is better defined. Novak (1998, p. 53) provides a model of meaningful learning that is amenable to measurement. It is shown in Figure 1. Briefly, meaningful learning is defined by three traits:

1. that the learner has prior knowledge that is relevant to the new learning to be done;
2. that what is to be learnt is presented in ways that have meaning;
3. that the learner must choose to learn meaningfully.

These criteria can be used to detect change that is meaningful. First, individuals must be able to show knowledge (and understanding) prior to the learning event. Second,

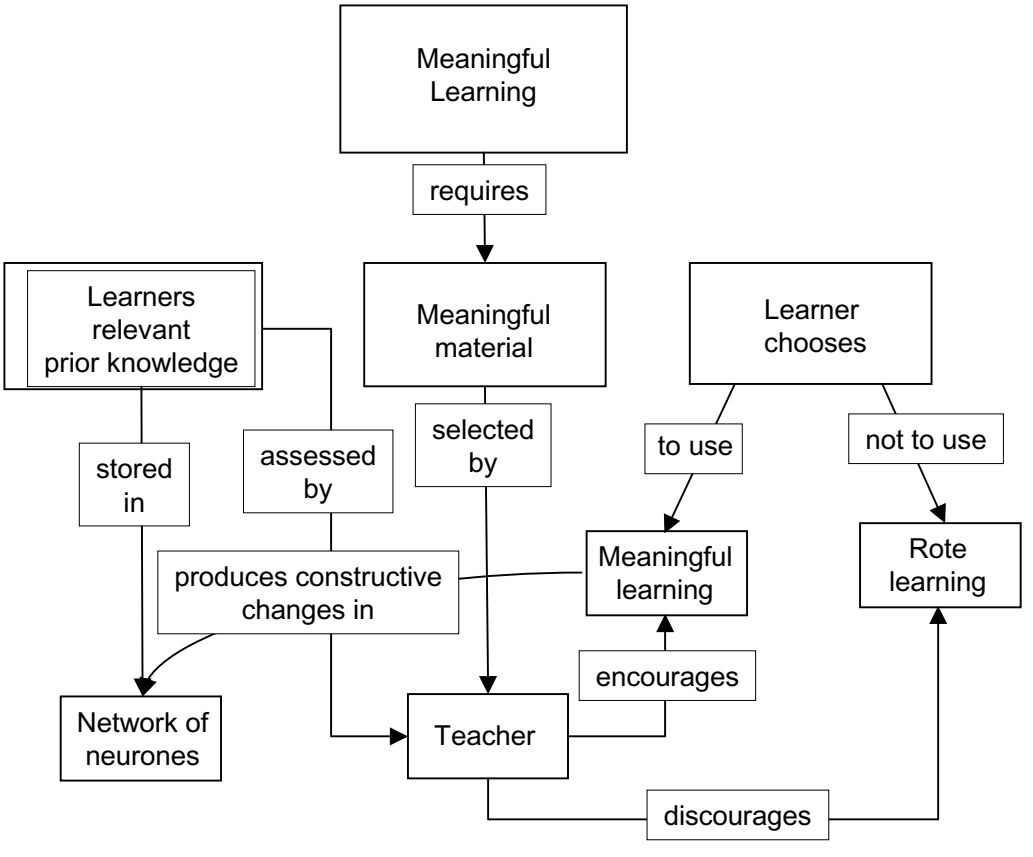


Figure 1. The definition of meaningful learning (adapted from Novak, 1998)

what is subsequently learnt must have meaning in the context of what was known before. Third, there must be evidence of the active integration of the new with the old on behalf of the learner.

Concept mapping (Novak, 1998) has powerful utility for the demonstration of learning in these terms. It can be used to display individual knowledge structures for comparison at different stages of the learning process (Novak & Symington, 1982; Novak, 1998), and to distinguish between expert and novice knowledge structures (Novak & Symington, 1982; Kinchin, 2001). In a departure from the original approach, however, Kinchin *et al.* (2000) have shown how the qualitative analysis of concept maps can reveal typologies of gross knowledge structures indicative of different patterns of knowledge and understanding. This approach does not attempt to measure change in quantitative ways (through change in concept richness or map linkage, for example), as has been done by many (e.g. Novak, 1998; Hoz & Gonik, 2001; Cassata *et al.*, 2004), nor does it attempt to validate change in terms of 'expert notions of correctness' or predetermined learning outcomes. Instead, it shares an epistemological approach with Jarvis, and suggests that 'meaningful change' is any that gives greater potential for the exposition of individual interpretations of a subject. The approach is a significant departure from Novak's original definition of meaningful learning (Novak, 1998), and it implies a focus on individual student approaches to learning rather than change as a process of development towards goals specified by teachers. The framework (used here) for assessing learning in qualitative terms is developed in the context of deep versus surface learning (rather than meaningful versus rote learning). This is justified because the deep and surface learning distinction is more appropriate for description of students' approaches, rather than the analysis of teaching and learning as processes. Finally, the work done here attempts to integrate the non-learning definition of Jarvis (Jarvis, 1992) with the deep and surface learning approach. This is valuable because non-learning is commonly neglected as a distinct outcome of potential educational experiences, but it has characteristics that are clearly distinct from surface learning.

## Methods

This research was done in the course of a postgraduate teaching programme in Research Methods. The author was required to take two three-hour theory classes on 'interview techniques', one week apart. The opportunity was used to conduct the research presented here and the students were encouraged to produce concept maps of the topic *before* and *after* the class and set-reading (to be done in the intervening time). Before the subject teaching began, the students were given an introduction to the concept mapping method. These were shown how to make maps and taught the principles of the methodology as defined by Novak (1998). In particular, they were shown how concepts can be arranged graphically and in a hierarchical manner (more inclusive concepts at the top and detail and example at the bottom) to make visual representations of their knowledge and understanding. They were taught to link all the concepts that they could (with directional arrows indicating the way in

which a link was to be read), and to label these links with statements that explained the interaction. It was made clear that links without statements would be dismissed in any subsequent analysis (since connections between concepts that are not explained lack meaning), and the importance of developing maximum linkage was emphasised (as the best exposition of rich understanding). The students were also encouraged to take time to reflect and reorganise their maps as often as possible, so that the final products would be the most parsimonious exposition they could make. Then they were asked to make the first (naive) map of pre-existing or prior knowledge. For this they were allocated three-quarters of an hour. The maps were collected by the researcher/teacher and the subject-based class commenced.

The lesson consisted of seminar-style presentation and facilitated discussion. The student group comprised adults aged from 26 to 55, and was a mix of UK and overseas learners. There was no one in the group who did not have some previous experience of educational research. At the end of the class references were given for the reading that was to be done before the next meeting. The class the following week was then begun with a brief reminder of the concept mapping method, and a further 45 minute period for a repeat of the concept mapping exercise. At the second meeting, the students' first maps were not returned until the concept mapping exercise was complete. In both the *before* and *after* mapping exercises the participants were simply asked to make their maps to explain their understanding of the use of interviewing techniques in educational research.

In the analysis, some simple criteria were developed a priori to distinguish deep, surface and non-learning. The definition of deep learning was largely derived from Novak's definition of meaningful learning (Novak, 1998), but all references to the teacher role in learning were removed. Thus, the following criteria were used to identify deep learning:

1. The second map must show both newly learnt concepts (that were not included in the first) and original (prior) conceptions.
2. The second map must show that the new knowledge has be linked to the prior knowledge in ways that are meaningful (i.e. that the linking statements are valid and explanatory and provide evidence of meaning in the mind of the map author).
3. The overall knowledge structure of the second map is a significant improvement on the first (i.e. that is shows better organisation, higher linkage and richer exposition of meaning).

Surface learning was defined by failure to fulfil these criteria in cases where there was still evidence of significant introduction of change (replacement of old concepts with new ones or the simple addition of new conceptions). Thus, the criteria for surface learning were:

1. The second map must show significant numbers of newly introduced concepts (ones that were not evident in the first), but these are not integrated with prior knowledge by linkage to concepts that are persistent from the first to the second map.

2. The second map will contain new concepts, but the conceptual linkage of the map as a whole will not be increased as a result.
3. The second map will not constitute a significant improvement on the first, either in terms of structural richness (linkage) or explanatory power (meaning).

Definition of the non-learning typology was derived from Jarvis (1993), and comprised evidence of consistency from the first map to the second (i.e. the lack of conceptual change). Thus, the criteria for non-learning were:

1. Persistence of prior knowledge from the first map to the second.
2. The lack of evidence of significant reorganisation of conceptual structures from one map to the next.
3. The absence of newly introduced concepts in the second map.
4. The absence of newly developed links in the second map.
5. The absence of newly developed expositions of meaning among previously existing linking statements.

Each pair of *before* and *after* maps were then tested against these criteria and typified. This was done by analysis of gross structures (Kinchin *et al.*, 2000) and the validity of propositions by 'expert' judgment. Concepts linked with arrows but lacking explanatory statements were treated as if they were not linked at all, and linking statements that could not be understood by the author were similarly dismissed as meaningless.

## Results

The data comprised a total of 24 concept maps, two each for each participant; one *before* and one *after* the teaching intervention and reading. Here only eight maps are shown in detail; just the *before* and *after* maps from four participants, and the names under which these maps are presented are pseudonyms. These were chosen from the total data set to best illustrate readily identifiable patterns of learning. They are presented under three headings—Deep Learning, Surface Learning and Non-Learning—even though several included evidence of learning corresponding to more than one definition.

### *Deep learning*

Of the four case studies presented here, Asha provided the clearest example of deep learning. Asha's first map (Figure 2) used eight concepts to explain the topic. There were three levels of hierarchy and all concepts were linked by explanatory statements. The concepts of the interview 'process' and interview 'style' had an organisational role, and subsumed detail that was explained under these headings. However, there was no cross-linking anywhere in the map. The concept of the 'environment' was loosely bolted to the rest of the map by a link to 'method', but this was not adequately explained or justified by the linking statement.

Overall, Asha's first map is a trivial exposition of the topic and it gives only a tenuous description of the interview method. It suggests that Asha had a broad prior

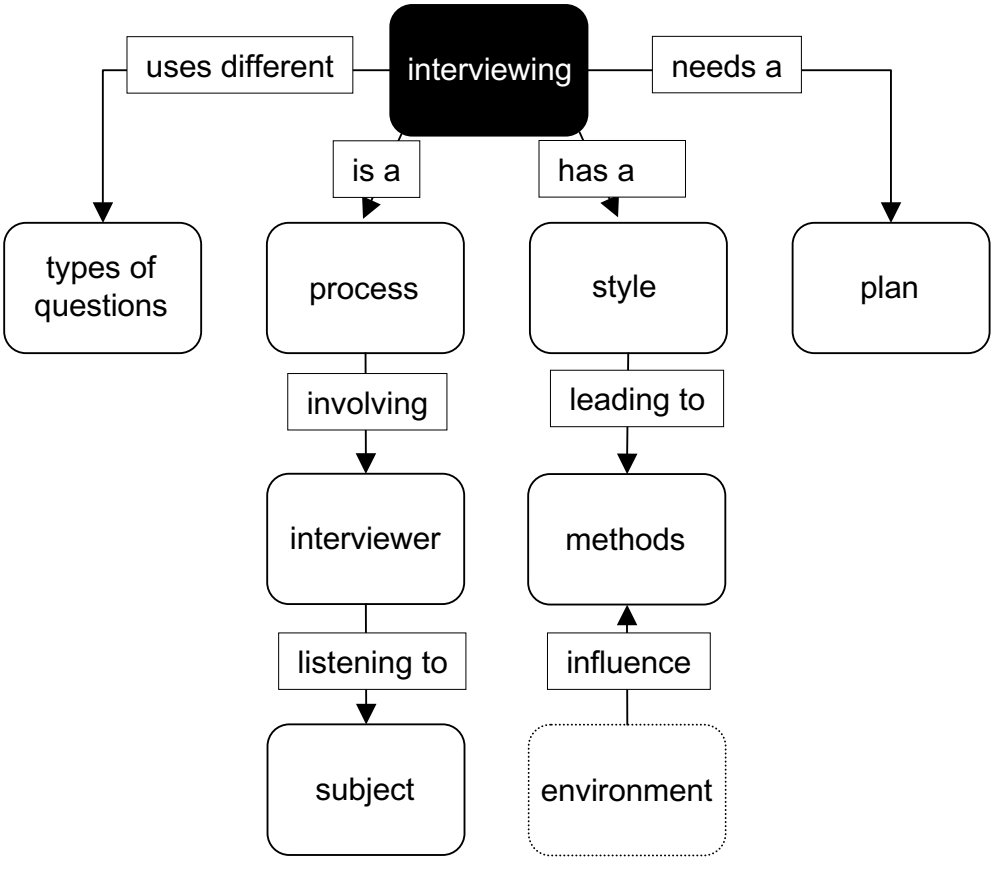


Figure 2. Asha's first map. The top concept is shown in black and the concepts Asha used to explain the topic are in white boxes. The directional arrows suggest the ways in which the appropriate linking statements are to be read. One concept (the 'environment') is shown in a shaded box since it in not anchored in the map and is superficial

knowledge of the subject, but her understanding was not deep. In particular, the concept of 'planning' was not placed under the 'processes' heading, nor was it linked to any other concept in the 'process' hierarchy. Similarly, the 'types of question' were listed as a distinct concept, suggesting that she knew that different question types could be used in the interview process, but not how or why this should be so. Furthermore, she did not understand how the selection of question types could be linked to the planning phase. Concepts such as 'style', 'process' and 'question types' were all ill defined in Asha's first map, and it is likely that her understanding of these terms was confused at this stage of her learning.

After the teaching session and the reading done in the intervening week, however, Asha's map was much more explanatory (Figure 3). In her second map, Asha again used only eight concepts to describe the topic, but overall, the knowledge was much better structured. It was also significantly improved in explanatory content and

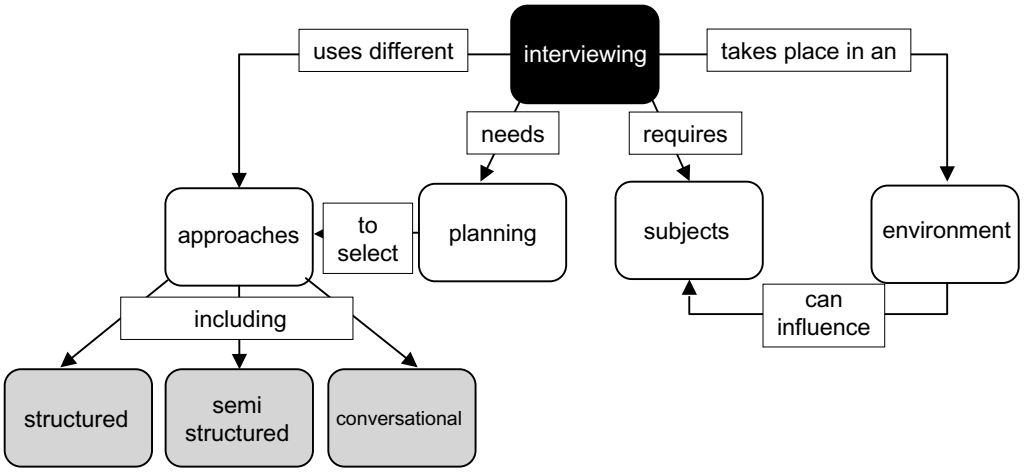


Figure 3. Asha's second map. New concepts (ones that did not occur in Asha's first map) are shown as shaded boxes

showed a more comprehensive grasp of meaning. Potential confusion between interview 'processes', 'styles' and 'question types', for example, was resolved by using the collective term 'approaches', and this was linked to 'planning' in a meaningful way. The link between the 'environment' and the interview 'subject' was also explained sensibly, and in a way that it was not in the first map. The second map included three new concepts—'structured', 'semi-structured' and 'conversational' interview 'approaches'. These concepts were added in ways that suggest deep learning, because they were integrated within the overarching knowledge structure and linked to prior knowledge. Here, Asha was able to show how these 'approaches' were related to the interview 'planning' process, for example.

Overall, the gross structural change in Asha's conception of the topic was considerable. Her first map was a simple spoke structure with four radiating knowledge chains. Only two of these chains were extended beyond a single concept, and the two chains that did progress comprised only three concepts each. There was no cross-linking in the map at all. In her second map, however, Asha was able to construct a linked network from her knowledge.

Analysis of Andy's maps also showed evidence of deep learning. In contrast with Asha, however, Andy's first map (Figure 4) was already rich. The map comprised 12 distinct concepts and it is richly linked. Clearly, Andy's view was process orientated, and interviewing was also placed in the broad overarching context of research. The 'conclusions' in Andy's map played a key role: providing the bridge that linked the interview process to the 'revision and development' of the 'research hypothesis', and to the development of 'techniques' and interview 'schedules' through 'feedback'. Although there was considerable linearity in Andy's map, cross-linking that spanned a variety of hierarchical levels achieved integration and made the overall structure a simple network.

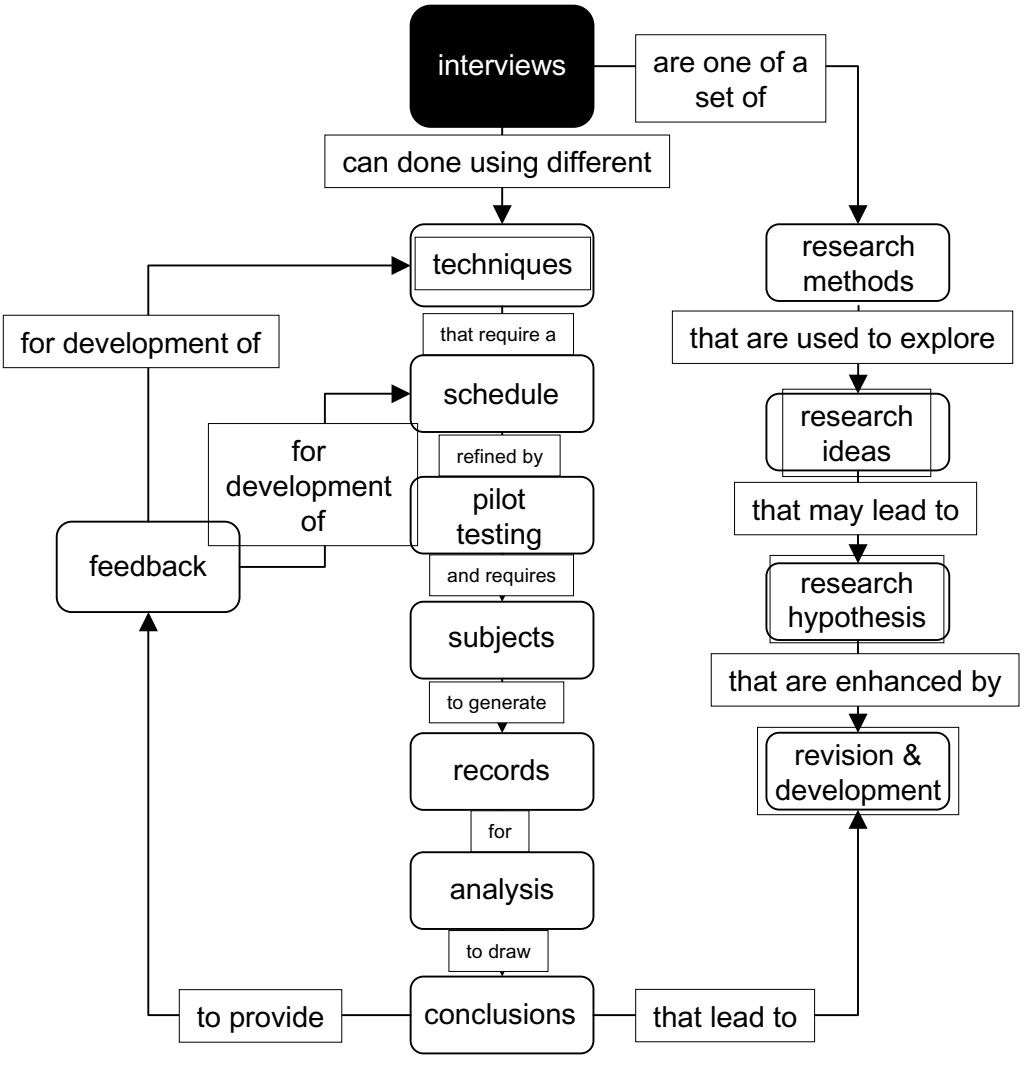


Figure 4. Andy's first map

The interlinking of the topic, however, was considerably developed in Andy's second map (Figure 5). This suggests that Andy's understanding of the topic had increased significantly as a consequence of his learning. Like Asha, Andy retained many concepts in the course of his learning, but he was also able to incorporate many new ones. He achieved a considerable increase in conceptual and explanatory richness in his second map as a result. Many of the new concepts were linked to those previously held, and comparison of the *before* and *after* maps suggests an increase in understanding that has been acquired in meaningful ways. Description of the 'recording' and 'analysis' processes, for example, have been widened in scope to include 'sampling', 'coding' and 'validity'. The topic is still grounded in the wider

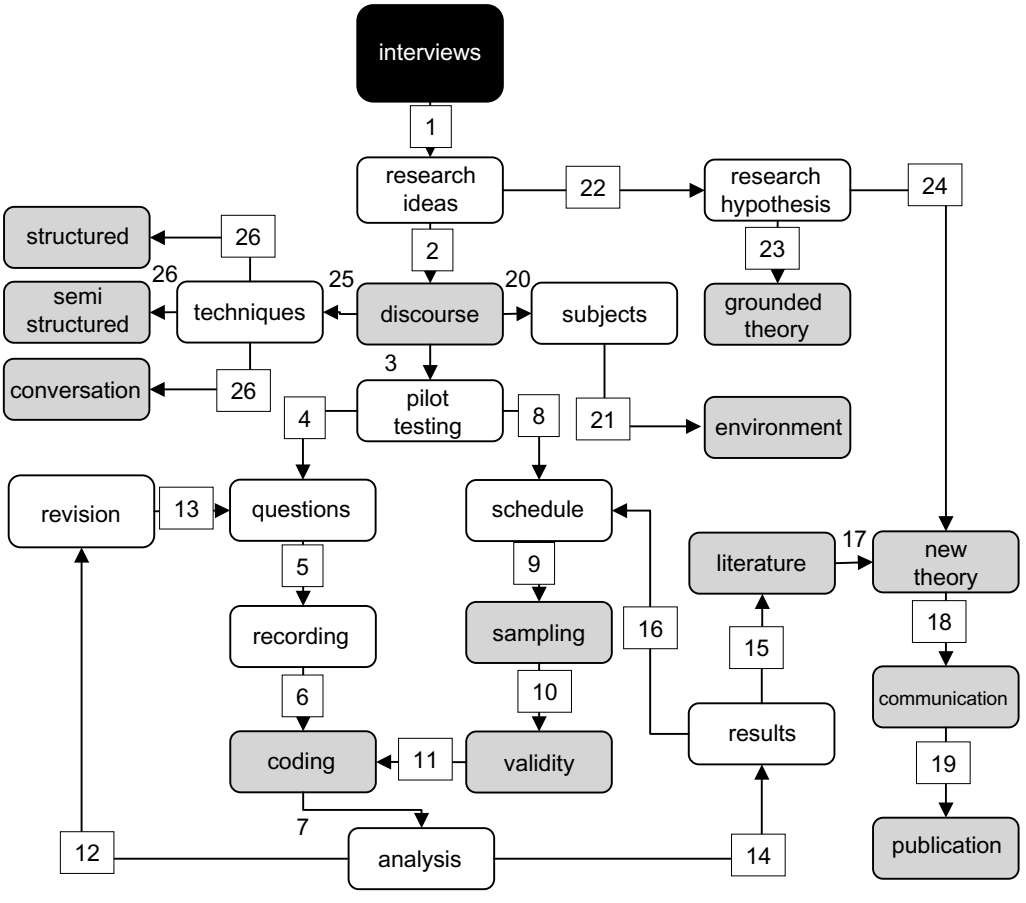


Figure 5. Andy's second map. Linking statements were as follows: 1 are a means of exploring, 2 through, 3 and require, 4 to develop appropriate, 5 and methods of, 6 so that, 7 leading to, 8 to produce a, 9 and approach to, 10 that ensure, 11 of, 12 leading to, 13 and rewriting of, 14 which produces, 15 that can be compared with, 16 that informs revision of, 17 to develop, 18 to be explained to others by, 19 for example by, 20 between interviewer and, 21 who may be influenced by the, 22 can be explained by testable, 23 may be lacking so that other approaches are necessary such as, 24 may be refined to develop, 25 using different, 26 such as

context of research, but this too has been expanded to show the role of the 'literature', 'theory' and 'communication' (which is exemplified by 'publication'). 'Discourse', 'grounded theory' and the 'environment' for the interview have also been introduced as new concepts, but are integrated well with the prior knowledge structure. Three new examples of interview 'technique' have been introduced; 'structured', semi-structured' and 'conversation' approaches, but, in contrast to the other additions, these are not well integrated and are symptomatic of surface learning. Andy has learnt that there are a variety of interview methods, but he has not yet been able to discuss their use in the context of other issues in his map: issues of 'validity', 'question' design or the interview 'schedule', for example.

Overall, Andy, was clearly someone with a good background knowledge of the subject to begin with, but he has also been able to use his new learning in meaningful ways. The rich complexity of his second map is a significant (and meaningful) advance upon the map that he first produced.

*Surface learning*

Andy's maps provide strong evidence of deep learning, despite some peripheral illustrations being added in a surface manner. Sally, on the other hand, showed a clear propensity for surface learning. Her first map (Figure 6) comprised lists of concepts loosely grouped or clustered under conceptual headings, but without explanatory links.

Sally appears to have aimed at a spoke and chain structure (she produced five distinct chains branching separately from the top concept at the hub). Nevertheless, she was unable to add the explanatory links that would have completed this structure. This suggests that she had a superficial knowledge of the range of topics and issues comprising the subject, but had little grasp of the underlying principles necessary for her to make meaning of the topic. Her second map (Figure 7) comprised many new concepts, but like her first, her attempt to explain it was entirely without linking statements.

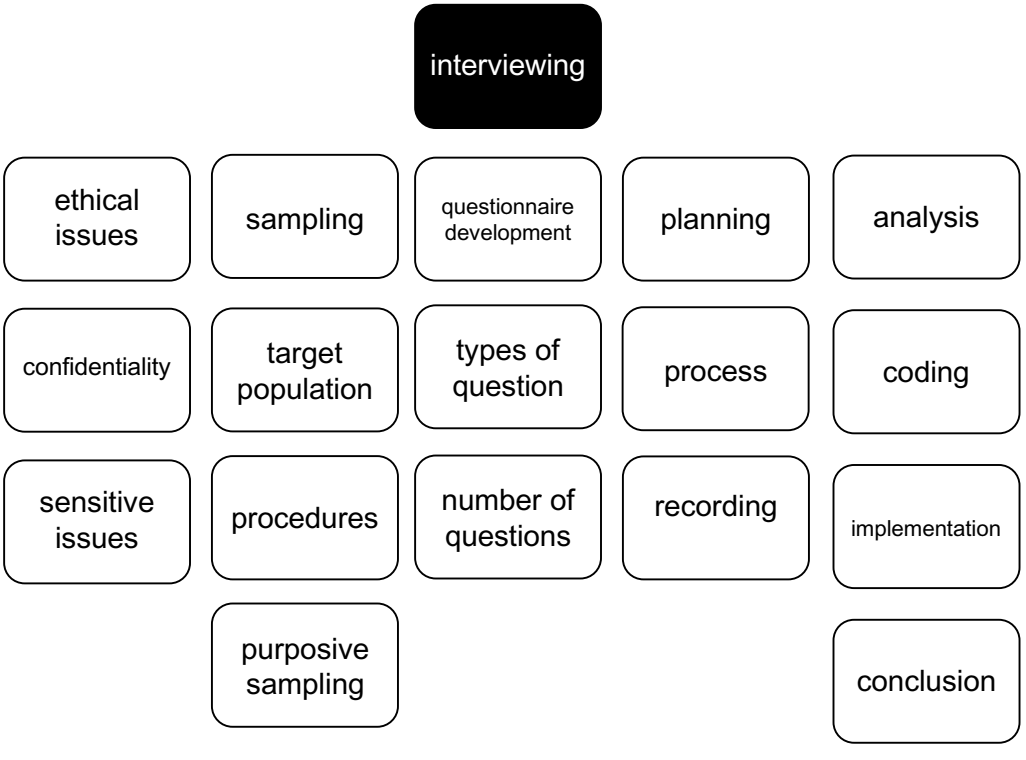


Figure 6. Sally's first map

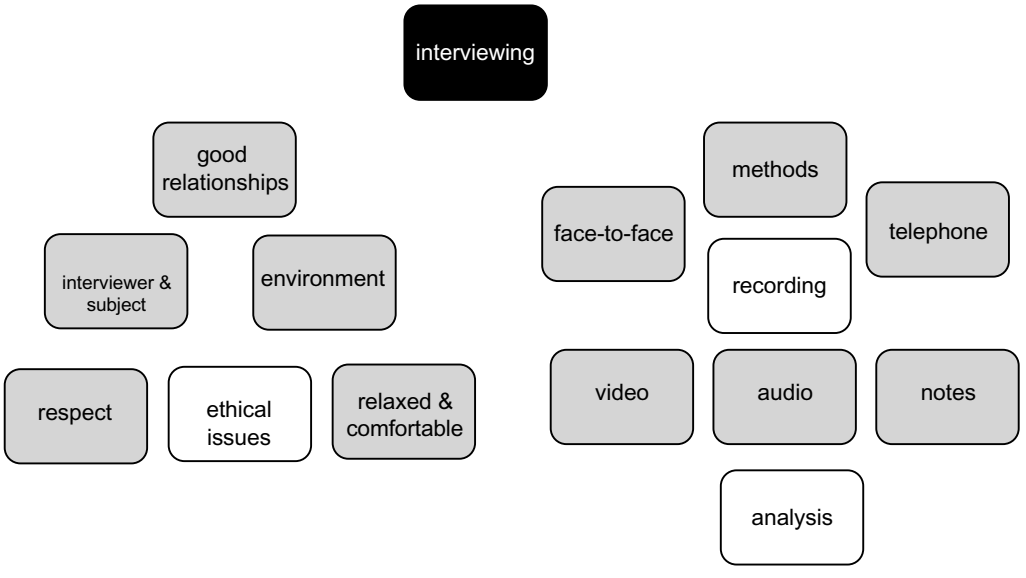


Figure 7. Sally's second map

In her second map, Sally re-used only three concepts from her first map, 'analysis', 'recording' and 'ethical issues'. Otherwise she rejected all that she knew before and replaced it with new content. This is clearly evidence of surface (or rote) learning. There has been little (or no) attempt to integrate the newly presented information with prior knowledge. Furthermore, the second map (like the first) is without links. Sally has replaced one set of knowledge with another, but has done so without any increase in understanding. What she has learnt appears to be little more than a new set of terms. Nevertheless, she has learned *something*; her conception of the topic is different even if she remains unable to make real sense of it. Furthermore, the hierarchy of her second map is better developed than her first. This suggests that some change in understanding has probably occurred, even if it has yet to be demonstrated in meaningful ways.

*Non-learning*

Richard's first map (Figure 8) was a trivial description of the topic. The map was annotated with linking statements, but these were largely superficial and did little to illustrate the meaning of the linkage. Furthermore, there was no cross-linking from one part of the map to another. Thus, for example, 'pilot testing' was not described as a part of 'questionnaire design', nor was 'sampling' included as part of interview 'planning'. 'Sampling' and 'reporting' were described as part of the interview 'process', but 'planning', 'questionnaire design' and 'pilot testing' were not.

In Richard's second map the organisational structure was not significantly improved (Figure 9). Indeed, there was very little conceptual development from the

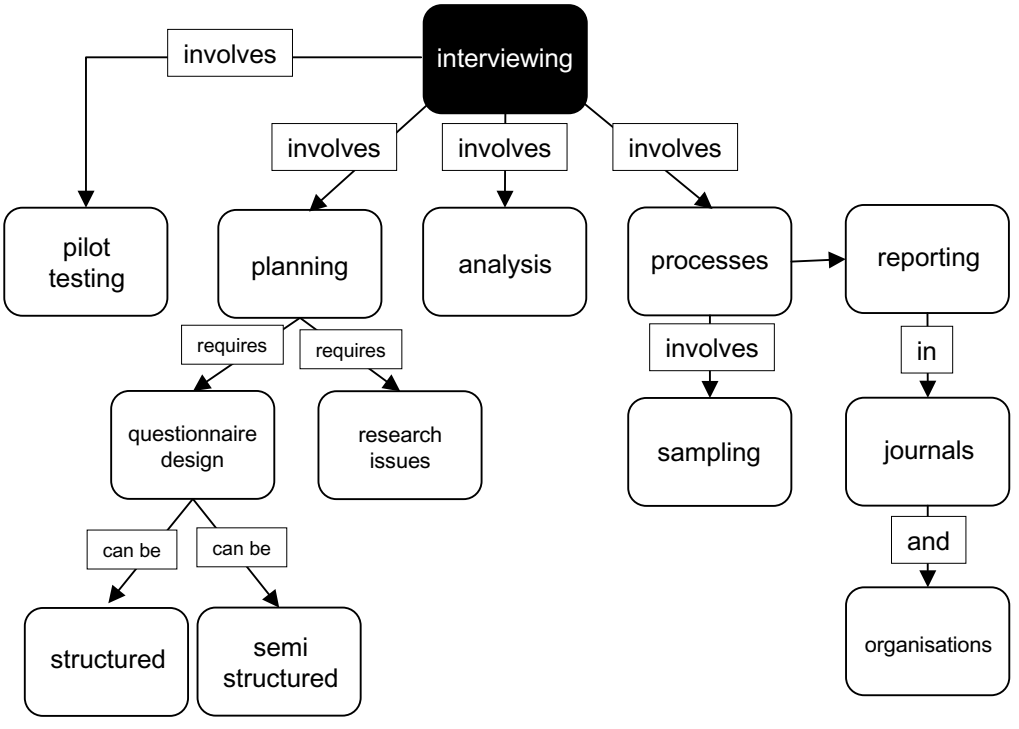


Figure 8. Richard's first map

first map. Some new concepts were introduced: 'ideas', the 'published research', a research 'question', and issues of 'data organisation' or 'data collection'. Three of these were linked together to form a new chain ('ideas' *developed by*, 'published research' *that helps form* 'a question') but this chain was not related to any other knowledge in the structure. It was probably the result of surface learning. The rest of the map was largely unchanged and is indicative of non-learning. The concept of 'collecting data', for example, was new and linked to the prior one of 'processes', but 'organising data' floated without linkage to anything else. Concepts of 'structured' and 'semi-structured' interviews and of 'organisations' (as targets for results dissemination) were likewise shown without linkage, even though they had been integrated in the first concept map.

The second map was not an advance on the first. The conceptual richness of the map was not improved and neither the hierarchy, nor the linkage richness, showed evidence of significant learning. Overall Richard's conception of the topic was reinforced but not changed in any meaningful way.

### *Analysis of the maps*

Of the eight sets of maps (not illustrated here), two were clearly examples of deep learning. These were not significantly different from the first case illustrated here

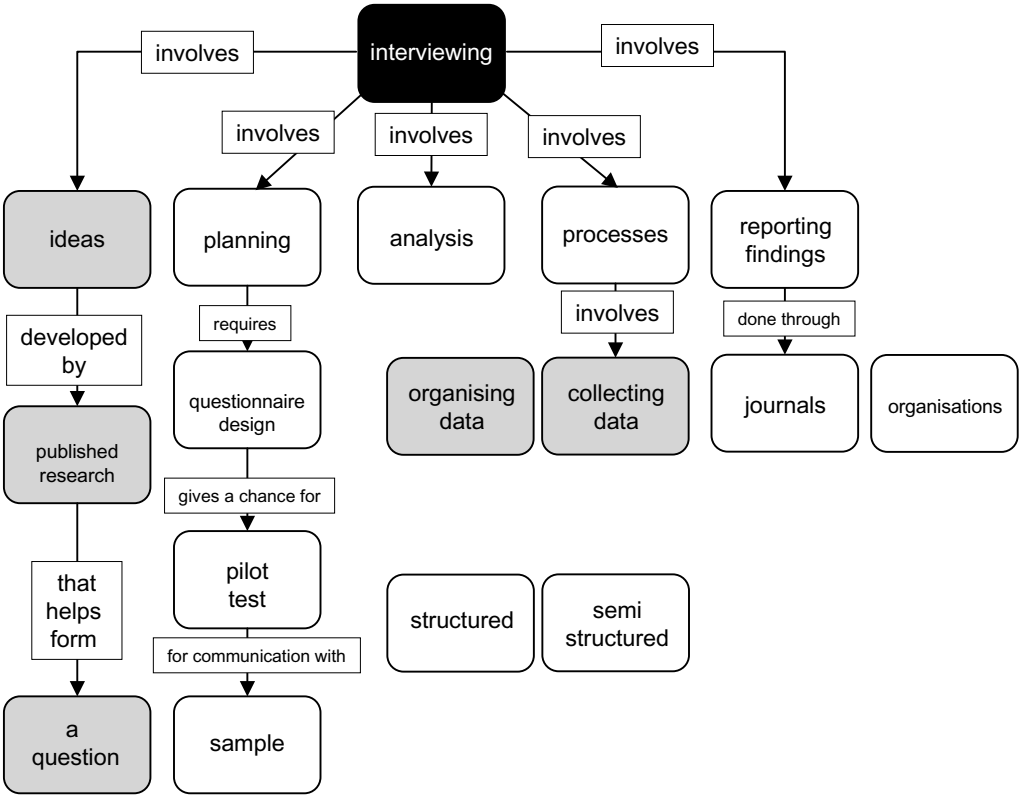


Figure 9. Richard's second map

(Asha); evidence of deep learning included an increase in structural complexity (networking), and the meaningful integration of newly acquired knowledge with the pre-existing knowledge structures. Three others were similar to the second case (Andy), and showed both deep and surface learning simultaneously. Two were predominantly non-learning cases (where knowledge structures changed little in the *before* and *after* maps). The final case was indeterminate. Here, the second map comprised *only* concepts that were listed in the first, but they were linked in ways that were clearly more meaningful (and better understood) than before. This suggests deep learning (because the structural complexity was increased and was indicative of newly developed understanding), but it failed to meet the criteria for meaningful learning (defined beforehand) because new concepts were not added in meaningful ways.

## Discussion

The data presented here provide empirical evidence for the meaningful distinction of deep, surface and non-learning. They suggest that concept mapping can be used to measure and typify the *quality* of learning (or the absence of it). Table 1 shows the attributes of deep versus surface learning. Surface learning clearly has utility in a

Table 1. The attributes of deep and surface learning (after Marton & Säljö, 1976)

| | Surface learning | Deep learning |
|---|---|---|
| Knowledge | An increase in knowledge or information about a subject acquired by gathering unrelated facts and without integration with what is already known. | An increase in understanding of a subject involving grasp of underlying principles. |
| Application | An ability to apply new knowledge to particular tasks and problems but without transferability. | An ability to apply newly understood principles in a variety of different contexts and situations. |
| Endurance | An ability to recall new information but usually only short term. | Long-lasting personal change. |

variety of educational contexts (Säljö, 1975; Marton & Säljö, 1976), and as a learning strategy it has benefits as long as it is used appropriately (Entwistle & Tait, 1994). Nevertheless, it is differentiated from deep leaning in qualitative terms.

Novak (1998) explores the issues of learning quality in his differentiation between meaningful and rote learning. He describes the utility of rote learning in precise terms. He states that 'sometimes it is useful to recall knowledge learned in precisely the same form as the original message. Phone numbers, for example, cannot be approximate' (Novak, 1998, p. 61). He also explains, however, that rote learning is limited in terms of recall and transfer and subsequent change, 'because [concepts learnt by rote are] stored arbitrarily and non-substantively in cognitive structure, [they] soon cannot be recalled and confer interference with new, related learning and recall of related information' (Novak, 1998, p. 62). Illustration of these issues is commensurate with the distinction between deep and surface learning (Säljö, 1975; Marton & Säljö, 1976). It is clear that different people, learning different things at different times, do so in ways that are more or less richly understood, have different endurance and are more (or less) transferable. This is probably the first article, however, to show that 'the making of meaning' can be reduced to directly observable phenomena, measurable against defined sets of criteria. That this can also be equated with measures of endurance and transferability, however, remains to be tested.

### *Knowledge change*

There is now a well documented literature on the role of 'threshold concepts' in learning (see, for example, Meyer & Land, 2003, but see also, Clement *et al.*, 1998). This reports that certain concepts (or overarching understandings of principle) can be 'troublesome' to learn, but that once they are acquired, represent new openings to the understanding of a topic not possible before. It is an important contention; first, it implies that learning is likely to proceed by increment (not continuous progression), as 'threshold concepts' are acquired; second, it is amenable to empirical assessment. Where student expositions of understanding lack meaningful (and structural) integration of 'threshold

concepts', their descriptions are likely to be trivial; when they do, individual expositions of the topic should be demonstrably improved. The approach reported here may make an important contribution to this strand of research. If 'threshold concept' acquisition is really indicative of 'breakthrough' into new patterns of understanding, then concept mapping the changes across such 'thresholds' is likely to make the process transparent. It might also promote the process by enhancing the student–teacher dialogue before and after concept acquisition (Kinchin, 2003).

Concept mapping is a powerful educational research tool, but it also has considerable utility in other learning, teaching and research contexts. Table 2 shows some of the basic uses of the method. This article has been presented in order to assess the validity of deep, surface and non-learning approaches at a process level, but it also exemplifies the contribution of the concept mapping method in the research and in the development of learning and teaching practice. The use of concept mapping in the broader contexts of higher and adult education is made explicit by the work of Daley (e.g. Daley, 2002a, b).

*Weaknesses and suggested areas for future research*

This article represents a new approach to the study and measurement of deep, surface and non-learning. It is a significant departure from the methods prevalent in the field, and as a consequence the methodology requires careful scrutiny. The research presented here is not without flaws, and raises issues that should be addressed in the future. The assertion that meaning is measurable in personal terms (without external tests of validity) is contestable. Nevertheless, there is a well-developed body of

Table 2. Some of the common educational uses of concept mapping

| Use | Description | References |
|---|---|---|
| Studies of learning | A tool for the identification of student knowledge and understanding in the course of learning | Novak & Symington, 1982; Novak, 1998 |
| | A method of identifying misconceptions in the course of learning | Kinchin, 2000, 2001 |
| Teaching practice | A method of promoting teacher–student dialogue | Kinchin, 2003 |
| Lesson planning | A means of prescribing content in meaningful ways | Kinchin & Alias, 2005; Martin, 1994 |
| Assessment | Quantifying understanding | Edmondson, 2000 |
| Cognitive typology | A means of showing ways of thinking and learning | Kinchin *et al.*, 2000 |
| Learning style | An approach for the identification of student learning preferences | Kinchin, 2004 |
| Expert identification | A means of distinguishing expert and novice knowledge structures | Novak & Godwin, 1984; Kinchin, 2001 |
| Teamwork | To achieve the integration of contrasting views among the different stakeholders in the educational design process | Hughes & Hay, 2001 |
| | A diagnostic approach to the investigation of student roles in learning | Kinchin & Hay, 2005 |

literature that explores this issue (Mintzes *et al.*, 1997, for example). In this article, deep learning has been defined only in ways that are personal and individual. Nevertheless, the criteria used here to define deep, surface and non-learning might be applied in ways that include external and 'expert' verification in the future, and this is likely to be informative.

This article also uses prior knowledge (and, in particular, evidence of the integration of new knowledge with old) as a criterion for the definition of deep learning. The work may, therefore, be criticised on the grounds that what is demonstrated in prior-knowledge maps is not necessarily indicative of all that was understood before instruction. Many concepts and knowledge structures may have remained undisclosed when the prior-knowledge maps were made. This is not an uncommon problem for work of this type, but it is also probably resolvable in the future. Interviews, for example, could be done to discriminate between changes that are attributable to the integration of newly understood concepts, and understanding that is developed by reorganisation or alternative description of that which was previously undisclosed. Being able to differentiate between meaningful change that is attributable to new knowledge acquisition, as distinct from reorganisation of existing knowledge is likely to be a significant contribution. This will be particularly relevant if it can be done in the context of 'threshold concepts' (Meyer & Land 2003).

This article does not test the (tentative and qualified) assertion that deep learning leads to achievement in ways that surface learning does not. This is an important area for future research. The distinction between deep and surface learning has significance only where it can predict differences in approaches indicative of different outcomes. What is done here provides a (measurable) bridge between the literature on conceptions of learning and outcomes. Importantly, it provides a framework for the definition of criteria by which meaningful learning can be observed and assessed. Nevertheless, research is still needed to show how such an approach might be instrumental in demonstrating that deep learning has utility in ways that surface learning does not.

## Conclusions

Concept mapping (Novak, 1998) is a powerful tool, both as a method of learning and teaching and as a research instrument. That concept mapping is underpinned by an established body of theory makes it appropriate for the empirical assessment of learning 'style', where many other methods are not. The strict definition of meaningful learning (Novak, 1998) means that learning *quality* is measurable (and perhaps even quantifiable), through integration with the definitions of learning developed by Säljö (1975), Marton and Säljö, (1976) and Jarvis (1992). Thus, criterion-based identification of deep, surface and non-learning is possible at the process level. Future studies should aim to further test the validity of a classification of learning in deep, surface and non-learning terms (as defined here). Work is required of longer duration, and it will be highly illuminating if future studies can provide empirical evidence of transition from surface to meaningful understanding. It is implicit in the work of both Jarvis and Novak, for example, that what is first learnt superficially, may be later incorporated

in knowledge structures in ways that are indicative of meaningful learning. If this can be related to changes that can be reliably interpreted in the context of 'threshold knowledge' acquisition, this will be a signal contribution to the field.

The work reported here uses only 'measures of meaning' to distinguish deep, surface and non-learning. This is done by assessment of knowledge-structure change and measures of conceptual richness. Deep learning, however, is distinguished from the surface approach in other ways; in terms of endurance and transferability, for example (Säljö, 1975; Marton & Säljö, 1976). Future work that includes these measures too will be important.

## References

Cassata, A. E., Himangshu, S. & Luli, J. J. (2004) 'What do you know?' Assessing change in student conceptual understanding in science, in: A. J. Canas, J. D Novak & F. M. González (Eds) *Concept maps: theory and practice.* Proceedings of the first international conference on concept mapping, Pamplona, Spain.

Clement, J., Zeitsman, A. & Brown, D. E. (1998) Not all preconceptions are misconceptions: finding 'anchoring conceptions' for grounding instruction on students' intuitions, *International Journal of Science Education,* 11(5), 554–565.

Coffield, F. J., Moseley, D. V., Hall, E. & Ecclestone, K. (2004a) *Should we be using learning styles? What research has to say to practice* (London, Learning and Skills Research Centre/University of Newcastle upon Tyne).

Coffield, F. J., Moseley, D. V., Hall, E. & Ecclestone, K. (2004b) *Learning styles and pedagogy in post-16 learning: a systematic and critical review* (London, Learning and Skills Research Centre/ University of Newcastle upon Tyne).

Daley, B. J. (2002a) The scholarship of teaching and learning: facilitating adult learning, *Journal of Scholarship of Teaching and Learning,* 3(1), 14–23.

Daley, B.J. (2002b) Facilitating learning with adult students through concept mapping, *Journal of Continuing and Higher Education,* 50(1), 21–31.

Edmondson, K. M. (2000) Assessing science understanding through concept maps, in: J. J. Mintzes, J. H. Wandersee & J. D. Novak (Eds) *Assessing science understanding: a human constructivist view* (San Diego, CA, Academic Press), 15–40.

Entwistle, N (1990) *Teaching and the quality of learning in higher education,* in: N. Entwistle (Ed.) *Handbook of educational ideas and practices* (London, Routledge), 103–136.

Entwistle, N. J. & Tait, H. (1994) Approaches to studying and preferences for teaching in higher education, *Instructional Evaluation and Faculty Development,* 14, 2–10.

Entwistle, N., McCune, V. & Walker, P. (2001) Conceptions, styles and approaches within higher education: analytic abstractions and everyday experience, in: R. J. Sternberg & L-F. Zhang (Eds) *Perspectives on thinking, learning and cognitive styles* (London, Lawrence Erlbaum), 103–136.

Entwistle, N. J., Meyer, J. H. F. & Tait, H. (1991) Student failure: disintegrated perceptions of studying and the learning environment, *Higher Education,* 21, 249–261.

Garner, I. (2000) Problems and inconsistencies with Kolb's learning styles, *Educational Psychology,* 20(3), 341–348.

Hoz, R. & Gonik, N. (2001) The use of concept mapping for knowledge-oriented evaluation in nursing education, *Evaluation and Research in Education,* 15(4), 207–227.

Hughes, G. & Hay, D. (2001) Use of concept mapping to integrate the different perspective of designers and other stakeholders in the development of e-learning materials, *British Journal of Educational Technology,* 32, 537–569.

Jarvis, P. (1992) *Paradoxes of learning* (San Francisco, CA, Jossey-Bass).

Jarvis, P. (1993) The learning process and late modernity, *Scandinavian Journal of Educational Research,* 37, 179–190.
Jarvis, P. (1998) *From practice to theory* (San Francisco, CA, Jossey-Bass).
Jarvis, P., Holford, J. & Griffin, C. (1998) *The theory and practice of learning* (London, Kogan Page).
Kinchin, I. M. (2000) Using concept maps to reveal understanding: a two-tier analysis, *School Science Review,* 296, 41–46.
Kinchin, I. M. (2001) Can a novice be viewed as an expert upside-down? *School Science Review,* 83, 91–95.
Kinchin, I. M. (2003) Effective teacher–student dialogue: a model from biological education, *Journal of Biological Education,* 37(3), 110–113.
Kinchin, I. M. (2004) Investigating students' beliefs about their preferred roles as learners, *Educational Research,* 46, 301–312.
Kinchin, I. M. & Alias, M. (2005) Exploiting variations in concept mapping morphology as a lesson planning tool for trainee teachers in higher education, *Journal of In-service Education,* 31(3), 569–591.
Kinchin, I. M. & Hay, D. B. (2005) Using concept maps to optimise the composition of collaborative student groups: a pilot study, *Journal of Advanced Nursing,* 51(2), 182–187.
Kinchin, I. M., Hay, D. B. & Adams, A. (2000) How a qualitative approach to concept map analysis can be used to aid learning by illustrating patterns of conceptual development, *Educational Research,* 42(1), 43–57.
Kolb, D. (1999) *The Kolb learning style inventory* (Boston, MA, Hay Group).
Kolb, D. & Fry, R. (1975) Towards an applied theory of experiential learning, in: C. L. Cooper (Ed.) *Theories of group processes* (London, John Wiley & Sons), 103–136.
Martin, D. J. (1994) Concept mapping as an aid to lesson planning: a longitudinal study, *Journal of Elementary Science Education,* 6(2), 11–30.
Marton, F. (1986) Phenomenography—a research approach to investigating different understandings of reality, *Journal of Thought,* 21(3), 28–49.
Marton, F. & Säljö, R. (1976) On qualitative differences in learning: I. Outcome and process, *British Journal of Educational Psychology,* 46, 115–127.
Marton, F. & Säljö, R. (1984) Approaches to learning, in: F. Marton, D. Hounsell, & N. Entwistle (Eds) *The experience of learning* (Edinburgh, Scottish Academic Press), 36–55.
Meyer, J. & Land, R. (2003) Threshold concepts and troublesome knowledge: linkages to ways of thinking and practicing within the disciplines. *Occasional Report 4, Enhancing Teaching and learning environments in undergraduate courses* (Edinburgh, ETL Project).
Mintzes, J. J., Wandersee, J. H. & Novak, J. D. (1997). Meaningful learning in science; the human constructivist perspective, in: G. D. Phye (Ed.) *Handbook of academic learning* (Orlando, FL, Academic Press), 405–447.
Nicholls, G. (2002) *Developing teaching and learning in higher education* (London, RoutledgeFalmer).
Novak, J. D. (1998) *Learning, creating and using knowledge: concept maps as facilitative tools in schools and corporations* (Mahwah, NJ, Lawrence Erlbaum).
Novak, J. D. & Godwin, D. B. (1984) *Learning how to learn* (Cambridge, Cambridge University Press).
Novak, J. D. & Symington, D. J. (1982) Concept mapping for curriculum development, *Victoria Institute for Educational Research Bulletin,* 43, 3–11.
Ramsden, P. (1992) *Learning to teach in higher education* (London, Routledge).
Säljö, R. (1975) Qualitative differences in learning as a function of the learner's conception of the task (Gothenburg, Acta Universitatis Gothoburgensis).
Sternberg, R. J. (2001) Epilogue: another mysterious affair at styles, in: R. J. Sternberg & L-F. Zhang (Eds) *Perspectives on thinking, learning and cognitive styles* (London, Lawrence Erlbaum), 103–136.