

Quantitative and qualitative measures of student learning at university level

David B. Hay · Harvey Wells · Ian M. Kinchin

Published online: 1 October 2007
© Springer Science+Business Media B.V. 2007

Abstract This paper reports the use of quantitative and qualitative measures of university student learning during teaching in psychiatry. Concept mapping, pre-and post test scores and performance in written assignments were used to measure the quality of change in personal understanding and to show the ways that the knowledge-targets of the course were achieved. The data show that: (1) Concept mapping can be used to explore personal understanding because it facilitates discrete statements of meaning. (2) These personal statements can be compared through time to assess change. (3) Specific criteria can be used to measure the quality of the change from one statement to another so that the different qualities of change that occur can be made-visible in the course of teaching. The approach is discussed in the broader context of learning theory and teaching practice. We show in particular, that prior-knowledge is an important determinant of learning because it affects the sense that can be made of taught material.

Keywords Concept mapping · Learning and teaching · Higher education · Pedagogy

Introduction

There is now a broad literature that focuses on learning and teaching in higher education (see for example, books by Biggs 2003; Nicholls 2002; Ramsden 2002; Prosser and Trigwell 1999). Compared with other educational sectors, however, there is remarkably little research that documents change in students' personal understanding of specific concepts in particular disciplines (Laurillard 2002). This is probably because the process of change is deemed intractable and because so few frameworks have been developed from

D. B. Hay (✉) · I. M. Kinchin
King's Institute of Learning and Teaching (KILT), Kings College London,
James Clerk Maxwell Building, 57 Waterloo Road, London SE1 8WA, UK
e-mail: david.2.hay@kcl.ac.uk

H. Wells
The Institute of Psychiatry (IOP), De Crespigny Park, London SE5 8AF, UK

which criteria for the measurement of change are derived. The separate approaches of Jarvis (1992) and Novak (1998), however, are notable exceptions. Recently, Hay (2007) has shown that these models can be combined to provide criteria for the measurement of learning quality. Moreover Novak's concept mapping method (Novak 1998) is a means by which change in personal knowledge can be made visible during learning (see Hay, Kinchin et al. *in press*). This is explained here in a single discipline of higher education (psychiatry). The principles are general, however, and one of the purposes of the work described here is to show how concrete measures of student learning can be obtained in the course of teaching for any cognate discipline. This is because the approach (of this paper) is grounded in definitions of *learning as change* rather than convergence on any given target understanding (e.g. Jarvis 1992) and specific criteria for the measurement of the quality of change have been developed by Hay (2007 after Novak 1998)). These issues are explained below.

Jarvis' approach is located in a wider scene of adult and continuing education. It is preceded by the work of Kolb (e.g. Kolb and Fry 1975), but although it shares a focus on learning through experience it is also sufficiently different to claim an independent antecedence. Briefly, Jarvis' definition of learning states that there are only two possible outcomes of experience. The person (learner) is either changed (learning) or they are not (non-learning). This is a succinct and useful definition because it suggests how learning should be measured (as personal change). The ontology of this definition is clear; learning is personal change. The definition is therefore applicable in any discipline.

Novak's approach is a little more complex and focuses not on the ends-state of learning, but on the process by which learning is acquired. Novak's work is grounded in Ausubel's theory of assimilative learning (Ausubel 1963, 1968; Ausubel et al. 1978; see also Ausubel 2000) and suggests that all change in knowledge and understanding can be located somewhere between the two extremes of rote and meaningful learning. According to Novak (1998) these two ends of a single learning continuum are defined by the degree of integration of newly learnt material with prior-knowledge. Thus measurement of change (to qualify new knowledge integration) is an empirical basis for the study of learning quality (Hay 2007). The approach has an important corollary in the literature on deep versus surface student learning (Säljö 1975; Marton and Säljö 1976, 1984) but where the literature on approaches to learning is grounded in the phenomenographic method (Marton 1981, 1986) or psychometry (e.g. Biggs 1993, 2003; Entwistle et al. 2001, Entwistle et al. 1991), Novak's model is a basis for measurement of change in knowledge in particular disciplines (Hay and Kinchin 2007).

Using concept mapping to measure change

Concept mapping is one of a broad family of graphic organising tools that includes mind-mapping (Buzan and Buzan 2000) and spider diagramming (Mintzes et al. 1997). Yet Novak's method (Novak 1998) has some very specific rules that set it apart from other techniques and facilitates the measurement of learning quality. A concept map is a declarative model of personal understanding comprising concept labels to identify specific ideas and explanatory links that give these ideas meaning. Thus, concept maps comprise any number of propositional statements, each of them made up of paired and linked concepts. Each proposition is a statement of understanding and the validity of each assertion is laid bare. This is a powerful teaching tool since it facilitates the sharing of personal and target understandings (*sensu*: Entwistle and Smith 2002) among teachers and

students (Kinchin and Hay 2007). Where the same person maps the same topic repeatedly then a comparison of two or more such ‘snap-shots’ facilitates measurement of the quality of change (Novak and Mussonda 1991; Hay 2007). Those parts of the knowledge structure that are new can be differentiated from those that are old and the degree of integration (of new ideas and extant parts of the prior-knowledge structure) can be measured.

Novak’s theoretical work on rote and meaningful learning is also a significant framework for the measurement of learning and teaching quality. The literature on student learning ‘styles’ and ‘approaches’ is well developed but this is not the same. Work on deep versus surface student learning styles (Säljö 1975; Marton and Säljö 1976, 1984) and on approaches to learning more generally (Biggs 1993; Entwistle et al. 2001; Entwistle et al. 1991; Entwistle and Tait 1994; Kember 1996, 2000) are useful since they inform a wider understanding of issues in university learning and teaching, but they remain ‘proxies’ for learning quality. They measure approaches to learning or study strategies (and sometime motivations) (Entwistle and McCune 2004), but they do not show how student’s understanding is changed as a consequence of learning). Novak’s approach is different because it is based on taking successive measures of personal understanding of particular concepts as they are changed in the course of learning and teaching (Hay 2007).

The work we present here has four aims: (1) to develop and explain an approach to the measurement of learning quality that is grounded in the work of Novak (1998) and Jarvis (1992); (2) to report the use of the method in a particular discipline (psychiatry); (3) to explain the ways that concept mapping can be used to document knowledge-change in the normal course of teaching for any discipline, and; (4) to explore some of the weaknesses of the method as well as priorities for future research.

Methods

This study was done during a short-course in mental health taught at the Institute of Psychiatry, King’s College London. The teaching was done in twelve 6 h sessions over a period of 3 months. It comprised lectures, seminars and group activities in the classroom. Reading, independent study and project work in groups was also required in the students’ own time. The summative assessment was a 3,000 word essay due five weeks from the end of the taught course.

Eighteen students took part and all of them finished the course. All of them were health care professionals with at least 2 years of clinical practice behind them. Three ‘A’ level passes at grade C or above was also a course prerequisite.

The course was entitled ‘Dual Diagnosis’ and it dealt with concomitant issues of mental illness and substance abuse. Figure 1 is a simple summary of the course content. The cognitive topics for the course were as follows: substance abuse, mental health, personality disorder, treatment, social exclusion and the prevention of relapse. Practice and skills learning included: interview techniques, Bethient assessment, cognitive behavioural therapy, medical intervention, stress and vulnerability models and the wider aspects of health-care. Most of the teaching sessions combined theory and the development of practical skill through case study analysis and role-play.

For the purposes of the research, the first session began with an introduction to the concept mapping method. The students were taught the method as it was described by Novak (1998) (none of them had encountered the method before). They were then given half an hour to make a map describing their own understanding of dual diagnosis prior to the course. These maps were then collected and retained by the tutor. After this, the

Fig. 1 The course was entitled, 'Dual Diagnosis' and was focussed on the concomitance of mental illness and substance abuse

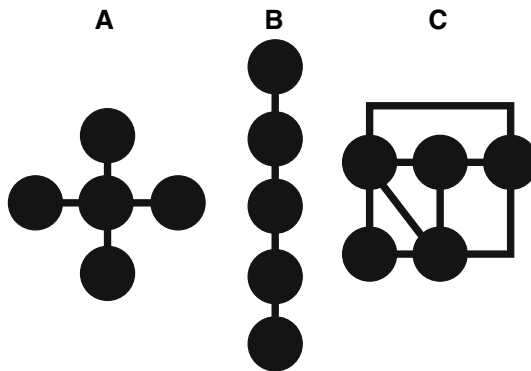
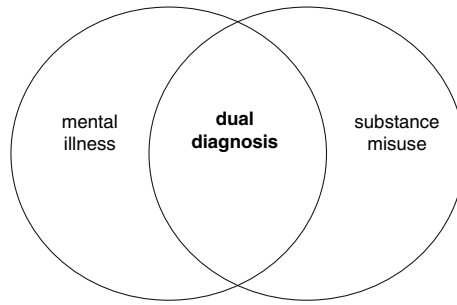


Fig. 2 Kinchin et al. (2000) report that the gross morphology of individuals' conceptual structures can be typified as spoke (a), chain (b) and network (c) structures. Spokes comprise concepts that are radial to a central organising principle (a); chains involve linear sequences of information (b); and networks (c) show a variety of alternative routes through what is known and understood. Spoke structures are usually indicative of novice or learning-ready status, chains are usually strategic and networks are broadly indicative of expertise (see Hay and Kinchin 2006; Kinchin and Hay 2007)

students' prior knowledge was further measured in a pre-course test. This comprised 15 short answer questions on the key-themes of the teaching that was to follow. Each question was worth a single mark and the students were informed of their score in the subsequent teaching session. The concept mapping activity and the short-answer test were repeated on the last day of teaching. Pre- and post-course test scores were compared and the changes in concept maps were assessed to document learning.

The concept maps that were produced individually before and after the course were analysed in two ways. First, the approach of Kinchin et al. (2000), was used to typify the gross morphology of each of the students' knowledge structures. This is summarised in Fig. 2. Briefly, the ways in which the concepts and links were arranged were classified as spoke, chain and network structures. The approach is now widely used in the literature and there are a number of papers that attest to its validity (e.g. Hay and Kinchin 2006). In particular, it offers a succinct measure of a student's propensity for learning (Hay and Kinchin 2006) and it avoids many of the pitfalls of quantitative concept map analysis (Kinchin et al. 2000).

The second approach was used to measure the quality of change in the course of learning. The method is summarised in Fig. 3 and was derived from work done previously by Hay (2007). Specific criteria were developed a priori from the literature on non-learning

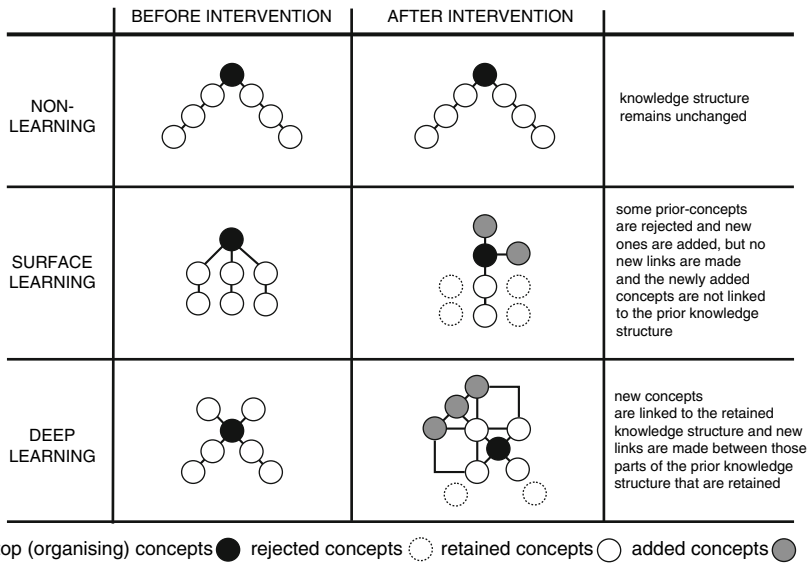


Fig. 3 Hay (2007) provides a framework for the measurement of change in the course of learning. This is indicative of different qualities of learning outcome. Non-learning (top) comprises the simple repetition of prior knowledge and constitutes reinforcement without change (Jarvis 1992). Surface learning (middle) is the addition (or deletion) of concepts in the absence of integration with extant parts of the prior knowledge structure. Deep learning (bottom) comprises the meaningful integration of prior knowledge and newly acquired information in ways that demonstrably increase understanding of the topic

(Jarvis 1992, 1998, 2006; Jarvis et al. 1998) and deep versus surface (Säljö 1975; Marton and Säljö 1976, 1984), strategic learning (Entwistle et al. 1991; Entwistle and Tait 1994) or rote-learning versus meaning-making (Novak 1998; Novak and Gowin 1984; Novak and Mussonda 1991; Novak and Symington 1982). The criteria were as follows:

1. *Non-learning*: was defined by an absence of cognitive change. Non-learning was therefore measured by the lack of new concepts in the second map and by an absence of new links in the extant prior knowledge structure.
2. *Rote learning*: was defined in two ways. First by the addition of new knowledge. Second by absence of links between the newly acquired concepts and those parts of the prior knowledge repeated in the second map.
3. *Meaningful learning*: was defined by a non-trivial change in the knowledge structure. Thus evidence of deep learning comprised the emergence of new links in parts of the prior knowledge structure developed in the course of learning and/or the meaningful linkage of new concepts to parts of the pre-existing understanding.

In the presentation of the results, all student identities are disguised by the use of pseudonyms. We show both qualitative and quantitative data for all 18 participants and a detailed case-by-case account for six of the students. The case studies are not used to extract generalities but to illustrate the ways in which the individuals learnt (or otherwise). Nevertheless, the cases that we have chosen are inclusive of all the types of learning change that we observed.

Results

Empirical measures of learning

Eighteen students took part in this study. Although no one dropped out during the teaching, two (Betty and Beth) failed to complete the assignment. Table 1 shows a summary of all the students' scores in the pre- and post-course tests and their respective essay grades. Most students gained a higher score in the post-course test than they had done in the pre-test, but the average gain score was less than one mark (in 15). The differences in pre- and post-tests were not significant. Class ranks were also largely unchanged. Mary, Steve and Tom were the only students to move by more than a single place. Mary improved more than any other student moving from 15th in the pre-test rankings to 6th in the post-tests. Her test scores were 8/15 (pre-course) and 12/15 (post-course); her final essay grade was 62%. Steve also moved upwards; from 5th to 3rd in the pre- and post-tests respectively. Steve achieved an essay mark of 57% and he scored 11/15 before the course and 14/15 at the end. Tom scored nine in both tests but fell from 11th to 14th in the class-rankings. His essay was a fail grade.

Concept map analysis showed that structural changes in students' knowledge were rare (Fig. 4). Of the eighteen students participating, in the study, only two drew concept maps of a different structural type before and after the course. One of these comprised change from a chain (prior to the course) to a network (afterwards). The other was a chain before the course but was unclassified afterwards because the concepts were not linked.

Table 1 Quantitative measures of learning

Pseudonym	Pre-test score (rank)	Post-test score (rank)	Difference	Essay grade % (rank)
Amy	14 (2)	14 (3)	0	66 (3)
Paul	9 (11)	11 (10)	+2	51 (10)
Mary	8 (15)	12 (6)	+4	62 (4)
Raquel	11 (5)	12 (6)	+1	52 (8)
Peter	10 (8)	11 (10)	+1	57 (5)
Harry	8 (15)	8 (16)	0	43 (13)
Simon	8 (15)	9 (14)	+1	71 (2)
Bella	5 (18)	6 (17)	+1	Fail
John	7 (16)	8 (16)	+1	45 (12)
Richard	9 (11)	9 (14)	0	52 (8)
Saul	10 (8)	11 (8)	+1	51 (10)
Rose	8 (15)	9 (14)	+1	Fail
Steve	11 (5)	14 (3)	+3	57 (5)
Tom	9 (11)	9 (14)	0	Fail
Vicky	6 (17)	5 (18)	-1	No Submission
Beth	10 (8)	11 (10)	+1	No Submission
Paul	11 (5)	12 (6)	+1	54 (7)
Tammy	15 (1)	15 (1)	0	76 (1)
Means	9.4	10.3	+0.8	

All 18 participants completed the same pre- and post- test before and after the course. The average gain score was less than one mark in 15 and the difference in pre- and post-course test scores was not significant (students 't' test). The essay grades that comprised the summative assessment were synonymous with either pre- or post-course test results (Chi-squared analysis)









before the course		 spoke	 chain	 network	 un- classified	totals
after the course	 spoke	3				3
 chain			3			3
 network			1	6		7
 un- classified			1		4	5
totals		3	5	6	4	18

Fig. 4 Knowledge structures before and after the course were largely homologous. Using the criteria of Kinchin et al. (2000), only two students were shown to depict their knowledge and understanding differently before and after the course. Among the structures that were identical before and after learning, three were spokes, three were chains and six were networks. Three maps before the teaching and four afterwards could not be classified because they were not linked or their links lacked linking statements

Overall, there were only two students that satisfied the criteria for meaningful learning (Amy and Tammy). Change in Amy’s maps is reported in the case studies below. Tammy’s knowledge change suggested a mix of rote and meaningful learning. She had learnt some parts of the course deeply and integrated new concepts with the prior knowledge, but other ideas had been added superficially and comprised chains unlinked to the rest of the structure. Nevertheless, Tammy came top of the class rankings before and after the course (scoring 15/15 on both occasions) and her essay grade (76%) was the highest in the class. There were seven cases where rote and non-learning outcomes were also combined. Rote learning occurred five times. Non-learning was the sole outcome on four occasions (in these cases the students simply replicated their prior knowledge maps in the second map-making exercise).

There were no obvious correlations between the quantitative and qualitative measures of student achievement. Rose for example, was one of only two students to produce an end-of-course concept map that was radically different from her first. She was in the lowest inter-quartile for achievement in her pre- and post-tests scores and her essay did not achieve the pass-mark. Tammy came top of the class in all quantitative measures but showed a combination of deep and surface learning outcomes. Amy made more meaningful sense out of what she was taught than any other student but her place in the post-tests ranking was not changed as a result and her essay grade was not different from a position in the class predicted by her pre-test result. Mary’s pre- and post-test scores showed more improvement than any other student, but the qualitative measures of her learning were indicative of superficial change. All of these outcomes are illustrated by the case studies below.

Case study analysis

Case one: Amy

Amy scored 14/15 in both pre- and post-course tests. Her final essay mark was 66%. Many of the concepts that Amy used in her map of the topic were common before and after the course and the overarching structure of her knowledge was the same (Fig. 5). Her first map was a network of nine concepts and the second, a network of ten. Moreover, these ten concepts were arranged on the page in positions that corresponded almost exactly to those used in the first map even where the concept labels themselves were changed. There was however, a considerable increase in linkage of her second map. There were 10 links in the first map and 20 in the second.

Apart from the top concept of *dual diagnosis* (which was the given starting point for the map), there were only three others that were carried forward verbatim from the first to the second map. These were *substance misuse*, *relapse* and *crime*. Four more concepts however, persisted despite alternative use of terminology. Thus for example, the concept of *mental health* was replaced by *mental illness* in the second map, but the two terms were apparently used synonymously. Similarly, the concept of *professional agencies* persisted in the guise of *psychiatric services*, *social life* as *social problems* and *activities and work* were replaced by the concept of *career* in the second. These substitutions were largely advantageous and helped the second map to be both more explanatory and succinct. Only one concept (the concept of *family and friends*) was included in the first map but neglected in the second. The second map comprised just two concepts that were entirely new (*drug services* and the concept of the *client*).

A comparison of Amy’s maps before and after the course shows evidence of meaningful learning. Overall, the second map was considerably superior to the first. Not only was it better linked it was also more parsimonious. New concepts were deeply integrated and the reorganisation of ideas was evidence of new and personal meaning. The definition of the top concept for the course (the concept of *dual diagnosis*) was considerably improved in the second map. The second map showed considerable evidence of individual meaning making in the course of learning. Where the concept of *activities and work* was included in the first map, this was done only as a single link from the concept of *dual diagnosis* (“[a

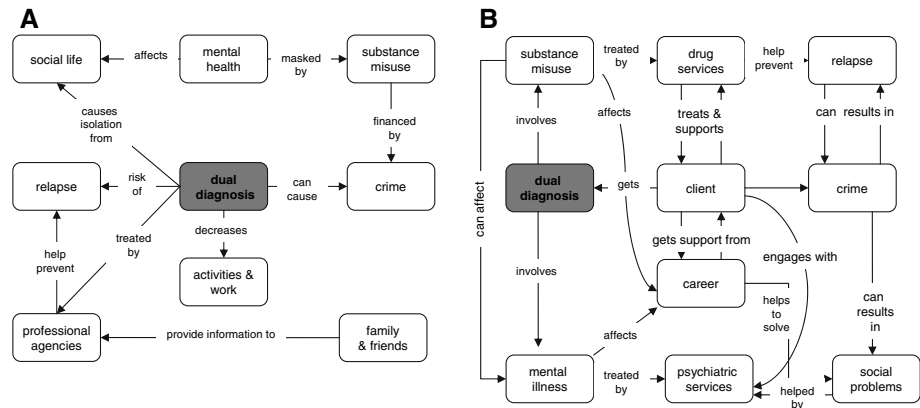


Fig. 5 Amy’s concept maps before (a) and after (b) the course

dual diagnosis decreases activities and work”). The link was tentative and may be deemed to imply undue causality. In the second map, however, the concept of [the] *career* was given much greater prominence. It had become an ‘organising principle’ and provided an anchor for the development of newly acquired meaning. In the second map, the impact of *mental illness* and *substance misuse* were described in terms of negative impacts on the *career of the client* and the role that a *career* might have for both the support of the individual (*client*) and the mitigation of *social problems* was explained. These are all themes that were introduced and discussed in the teaching of the module, but the ways that Amy chose to describe them in her second map was unique. It shows that she has compared the knowledge she brought to the subject and the material that she was taught to develop a new and personal understanding. Her learning was personal and the change from her first to second map was indicative of meaningful learning.

Case two: Mary

Mary’s pre- and post-course test scores (8/15 and 12/15, respectively) improved more than any other student. Mary obtained a mark of 62% for her essay. Like Amy, however, there was a negligible change in the gross structure of her concept maps before and after learning (Fig. 6). Both of Mary’s maps were networks and they comprised 9 concepts (with 14 links) and 13 concepts (with 20 links) respectively. Like Amy, Mary’s second map also included many concepts that were present in the prior knowledge structure (e.g. *mental illness*, the *client* or *Bethient*, *medication*, the *multidisciplinary team*). Some concepts were also retained but identified using different labels. The concept of *treatment* was replaced by the broader term *intervention* in the second map), and the idea of *drugs* was expanded to identify *drugs and alcohol*. Nevertheless, Mary added more new concepts than Amy in her second map. The concepts of *social problems*, *poverty*, *criminal activities*, *relapse* and

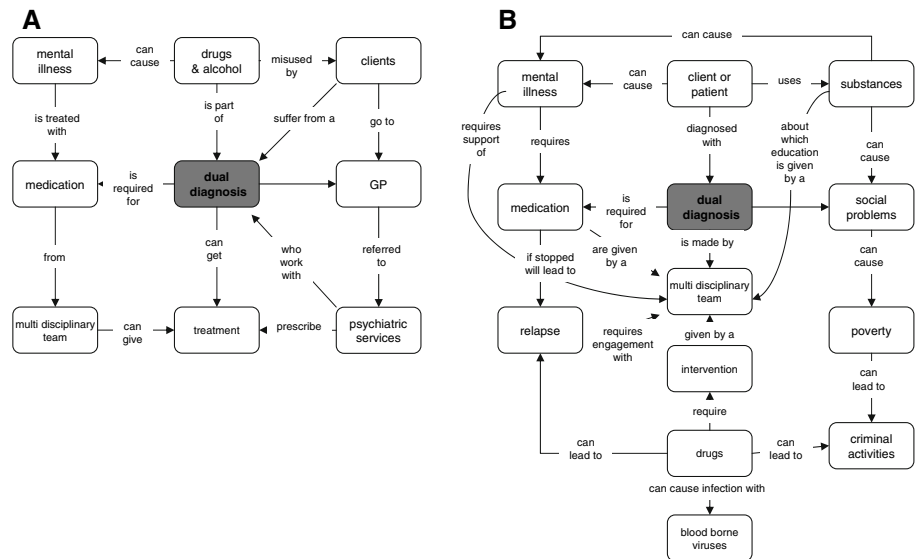


Fig. 6 Mary’s maps before (a) and after (b) the course

blood-borne viruses were all new in the second map. Conversely, the concept of the *GP* (general practitioner) appeared in Mary’s prior knowledge structure, but was neglected after the module.

The changes in Mary’s map provide evidence of learning but most of the changes are superficial and indicative of rote learning. This is because they constitute the simple addition of new concepts, linked one to another or to surface parts of extant prior knowledge. They are not deeply integrated with what was known before. There is little structural reorganisation (the same concepts tend to appear in the second map in the same places that they did in the first) and there is little new development of linkage among persistent concepts. The prior knowledge structure has not been changed meaningfully and addition has been made without evaluation. The new concepts of *poverty*, *social problems*, *criminal activities*, *relapse* and *blood borne viruses*, for example, share between them, only three links to concepts that were present in the prior knowledge.

Mary’s second map also comprises structural inconsistencies that are a barrier to the succinct explanation of meaning. Thus, for example, the terms *substances* and *drugs* are both used in the second map. This is probably an infringement of the rules for concept mapping (Novak 1998, requires that a concept is used only once in a map) and it decreases its parsimony. The problem is compounded by the use of the term *medication* as well. Mary’s second map is in need of structural reorganisation to resolve these issues and this is further evidence of surface learning.

Case three: Harry

Harry achieved a bare-pass mark for his essay (43%) and scored 8/15 in both his pre- and post-course tests. Figure 7 shows the changes in his knowledge and understanding before and after the module. His first map was a simple spoke structure in which four concepts (the concepts of *dual diagnosis*, *substance misuse*, *mental illness* and the *patient or client*), were arranged around the concept of the *nurse or doctor*. The map comprised a robust

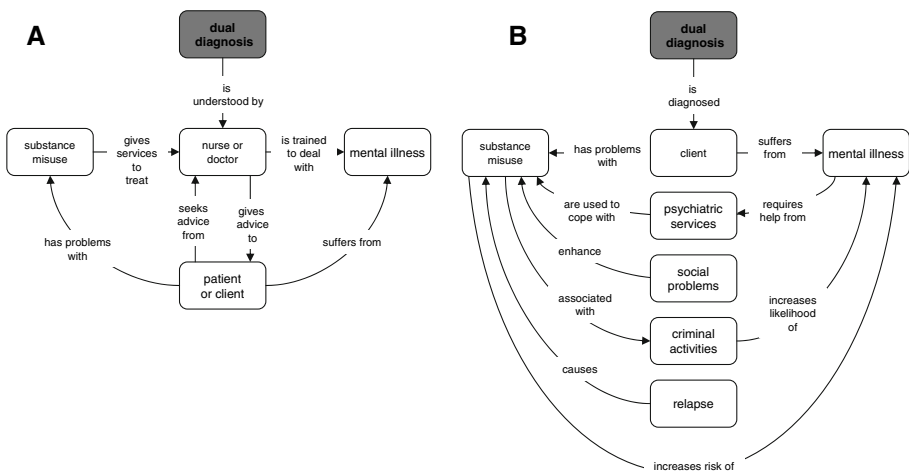


Fig. 7 Harry’s maps before (a) and after (b) the course

definition of the concept of *dual diagnosis*, however, because substance misuse and mental illness were described as simultaneous patient or client problems.

Structurally, the second map was almost identical to the first. This was despite the fact that the concept of the *nurse or doctor*, central to the first map, was absent in the second and the concept of the *client* now took this central and organising role. Three new concepts were added (*social problems, criminal activities and relapse*), but these were added peripherally and were not linked to the top concept (of *dual diagnosis*), nor to the new concept central (the *client*). These changes are indicative of a mix of non-learning outcomes and rote learning.

Case four: Bella

Bella was bottom of the student rankings in the pre-course tests (she scored only five marks of the 15 marks available). In the post-course tests she scored only one mark more (6/15) improving her position by just one place. Her essay was a fail grade. Before the course, her map was made of two chains (without linking statements), and afterwards these were longer (and linked), but not substantively different (Fig. 8).

The chains that Bella drew were initiated by the distinction between concomitant diagnosis of *mental illness* and *substance misuse*. This was appropriate for the definition of dual diagnosis that was developed by the course (Fig. 1). In the first map however, there

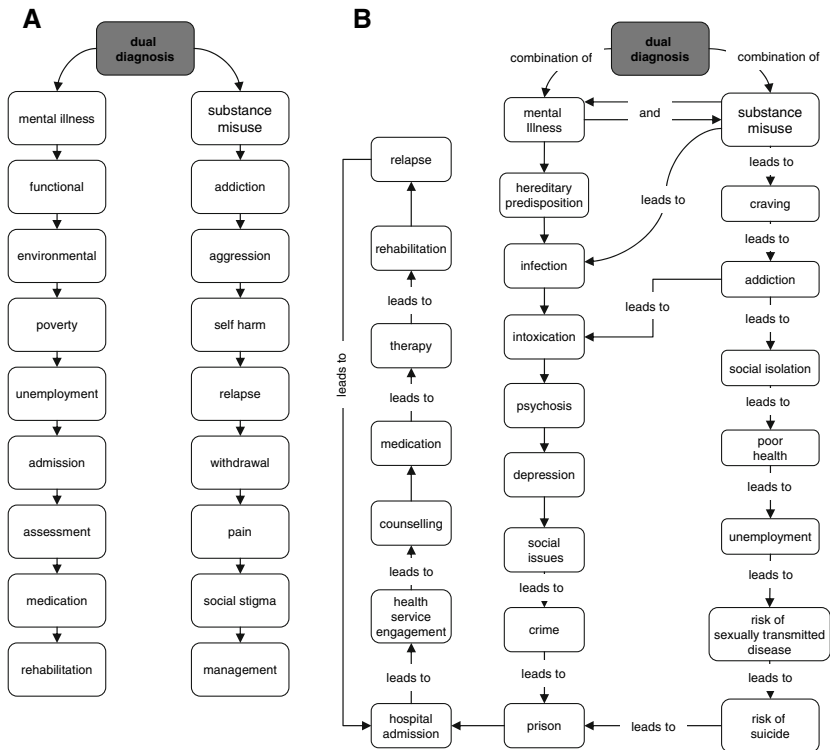


Fig. 8 Bella’s maps before (a) and after (b) the course

were no cross-links between the two chains, and in the second, there were only four (among a total of 24 concepts). Each chain in each of the maps was indicative of a temporal sequence of events rather than an exposition of the meaning of the subject. Furthermore, many of the successive links in the concept chains implied undue causality. Thus for example, *social isolation* was shown to lead to *poor health* and *unemployment* to an increased risk of *sexually transmitted disease*. The second map shows an important cross-link (newly acquired), between the concepts of *mental illness* and *substance misuse*, but this was not substantiated by linkage in other parts of the structure. Other cross-links in the second map were superficial because they were an obvious consequence of poor structural organisation (i.e. related concepts were unduly separated) or because the links were untenable (e.g. a *risk of suicide* leads to *prison*).

Bella’s concept maps before and after the module cannot be compared empirically (because the first map lacked links). Nevertheless, the second map is so trivial in its explanation of the topic that the change must be considered indicative of rote learning at best. More likely, it is non-learning.

Case five: Rose

Rose scored eight and nine in her pre- and post-course tests. Like Bella, she failed the course. She was also one of only two participants to show a significant change in knowledge structure. Figure 9 shows how the simple (and broken) chain that Rose drew for her first map was replaced by a complex network in the second.

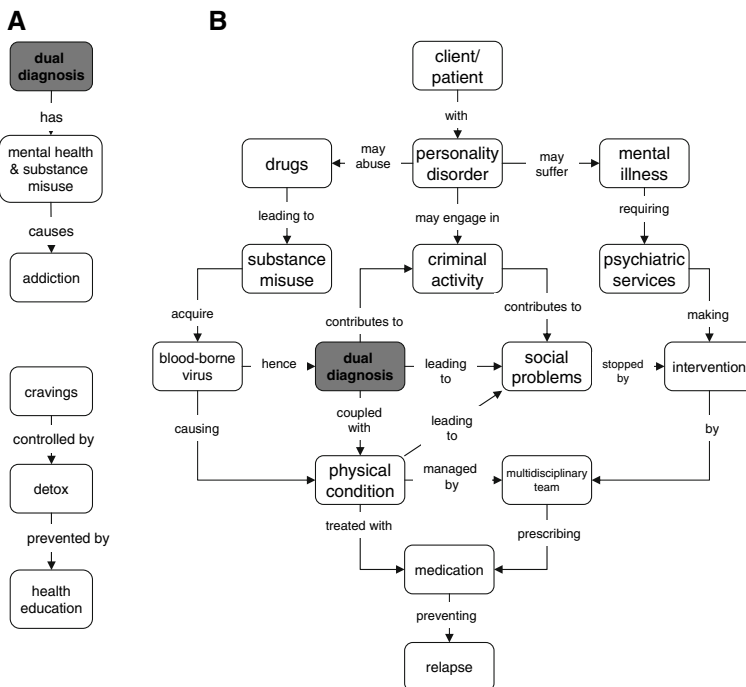


Fig. 9 Rose’s maps before (a) and after (b) the course

Rose's first map comprised 6 concepts and four links. It was an appropriate definition of the top concept (of *dual diagnosis*) but otherwise it was indicative of little prior knowledge of the subject. Her second map had 15 concepts and 19 links. At first glance these changes suggest considerable development of personal knowledge and understanding. Closer scrutiny, however, suggests that much of what Rose had newly acquired was poorly understood, and even her basic grasp of the subject had been lost by the end. Most importantly, before the module, Rose understood that a *dual diagnosis* comprised both *mental health* issues and *substance abuse*. By the end however, Rose was unable to define the concept in any succinct and appropriate way. Although she still apparently knew that both these issues were important, how this was so was confused and erroneously described. Thus for example, Rose's second map stated that the acquisition of a *blood-borne virus* was a key determinant of a dual diagnosis and that a *dual diagnosis* led directly (and causally) to *criminal activity* or other *social problems*. Moreover, the concepts of *substance misuse* and *mental illness* were related to the *dual diagnosis* concept only indirectly. The only route from the concept of a *dual diagnosis* to *mental illness*, for example, was through *medication* that was mediated *via* causal statements about *criminal activity*, other *social problems* and the treatment of *physical condition*. None of these were explanatory of the subject, nor were any of them sufficiently academic to merit inclusion.

Rose started with a superficial prior knowledge. By the end of the module her knowledge of terms had been increased but she had clearly failed to understand what she had acquired. She had added numerous misconceptions of the subject and was more confused than she had been before the course began.

Case six: Betty

Betty did not submit an essay for assessment. Her pre- and post-course test scores were 6/15 and 5/15 placing her at the bottom of the post-course test rankings. Her first map comprised only three linked concepts, one of which was the top concept of *dual diagnosis*. There was only one proposition in her first map. This asserted that *substance misuse and psychosis* (listed together as a single idea), cause *dependence and addiction*. Betty also listed the following terms in her first map: *disorder, detoxification, health education, therapy, Schizophrenia, epilepsy, disease, drug abuse and alcohol dependency*. None of these were linked, however. Her second map was not linked at all. Instead it comprised a list of eleven different concepts. Eight of these had been present in the first map and three were new (*criminal activity, relapse and a multidisciplinary team*). The concept of *dependency and addiction* had been dropped. The change in Betty's maps was therefore classified as non-learning. The second map explained less than the first and the concept list was not substantively changed.

A summary of structural change

Figure 10 is a summary of the structural changes documented in this study. Among the 18 students of this course there were only two who used different knowledge structures to describe their understanding before and after the course. All the others produced pairs of maps with remarkable structural homology, even where the content and its internal organisation were different (see Case One and Two, for example). The significant change in Rose's map (Fig. 10e) was not indicative of meaningful learning. Rose was clearly more

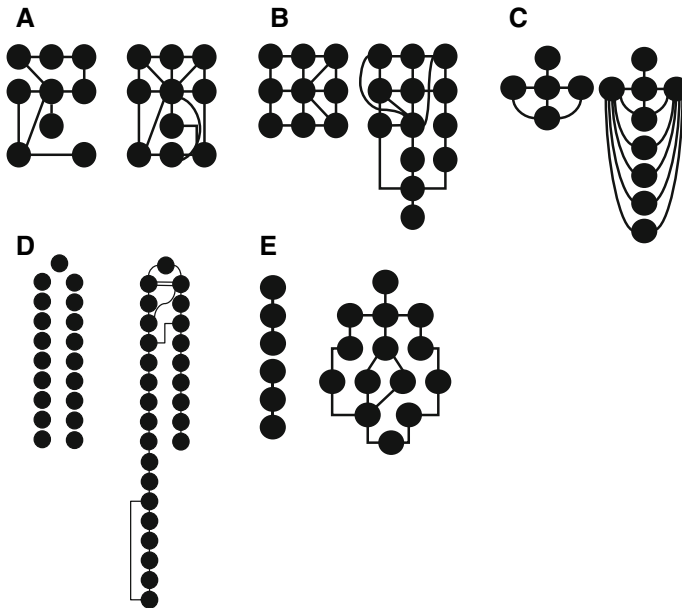


Fig. 10 Concept maps structures are shown for Amy, (a) Mary, (b) Harry, (c) Bella, (d) and Rose, (e) before (left) and after (right) the course. All except Rose (e) show high levels of structural homology even where the propositional composition of the maps had changed significantly (see the individual case studies). The gross structure of Rose's maps (case E) suggested meaningful change, but on closer inspection proved indicative of non- or surface learning

confused at the end of the course than she had been at the start, and she failed the assignment.

Discussion

Research of student learning

The data presented here suggest that Novak's description of rote and meaningful learning (Novak 1998) and Jarvis' model of learning versus non-learning outcomes (Jarvis 1992) provide robust frameworks for the assessment of learning quality. Moreover, that Novak's concept mapping method makes it possible for lecturer's to obtain sequential 'snap-shots' of their student's personal understandings so that issues of learning quality can be exposed and discussed.

This paper also shows how the use of simple labels of learning quality (non-learning (sensu Jarvis 1992) rote and meaningful learning (sensu: Novak 1998)) can be extracted from measures of change through concept mapping. These labels are developed out of complex theories of learning (Jarvis; Novak; Ausubel *ibid*); all of which emphasise issues of personal difference, but there is still a risk that 'labels' mask the underlying richness of learning as a continuum (Novak 1998; Kember 2000). It is important, however, that our approach shows how measures of learning quality can be attached to declarative statements of personal understanding (rather than personal 'styles' or 'approaches'). The

case-commentary that is a large part of the results in this paper is derived from the actual discussions that took place between the teacher and the students who were party to the study. It exemplifies the complexity of discussion of learning quality that is made possible where real measures of personal understanding are successively declared. ‘Labels’ and ‘definitions’ (of learning) were important (in these discussions) as they focused attention on particular *values*, but the real significance of the approach is better represented by the view that *concept mapping makes learning visible* (Hay, Kinchin et al. [in press](#)).

Entwistle and Smith (2002) argue that personal and target understandings are simultaneously important in student learning. But university teaching and assessment strategies while generally good at measuring student achievement (against the ‘targets’ of formal assessment), are less good at showing how personal knowledge and target performance are related (Kinchin and Hay 2007). As a consequence there is general concern that some ‘targets’ are achieved through rote learning (Kinchin et al. [in press](#)). One of the problems has always been that university teachers have not known how to acknowledge personal meaning-making; either in theory or in practice. Constructivism has led to widely used theories of learning that are personal, but if the theory is to be applicable in teaching then it must also find methods of practice that can be reconciled with the stated targets of teachers and the curriculum (Entwistle and Smith *ibid*). The data presented here show that lecturers *can* measure personal knowledge-change and that they need-not radically alter their teaching strategies to do so. Concept mapping is a practical method in the constructivist tradition; it can be used by teachers to *visualise* the ways that their students make sense of the knowledge targets of teaching but it also helps students incorporate new information towards personal understanding.

As we have shown, concept mapping can be used to document personal understanding and doing so may show how troublesome ideas (Perkins 2006) come into view or comprise thresholds of new understanding (Meyer and Land 2003, 2005). Concept mapping can be taught in about 20 min and satisfactory maps can usually be made in half-an-hour. This means that the method is realistic during normal university teaching. Some of the weaknesses and limitations of the approach are discussed below, but the results reported here show how one teacher (of psychiatry) was able to use the method to integrate research of their students’ learning as part of their normal teaching practice. Our work draws particular attention to the role of prior-knowledge. Each case of meaningful learning was prefaced by a rich understand of the topic before teaching and rote or non-learning occurred where prior-knowledge was poor. This is an important issue for higher education, but one that is often neglected by theories of learning ‘style or ‘approach’. Meaningful (or deep) and rote (or surface) learning *approaches* may be consequences of ‘proximity’ (or the lack of it) between prior-knowledge and the targets of teaching. Students who lack prerequisite knowledge may be compelled to adopt rote learning and learning approaches more generally may be *proxy* for understanding that is made possible (or denied) at a more fundamental level of cognitive integration.

Towards an authentic pedagogy of higher education

Concept mapping provides an epistemological framework for teaching that is grounded in theory (Ausubel; Novak; *ibid*). It is also a method of research that facilitates the empirical measurement of learning (this paper). As a result, it spans the common gap between theory and practice and is a basis for authentic pedagogy. In particular, it has the potential to engage students and teachers in a discourse of scholarship. At any point in time it can be

used to share an understanding of a topic and to explore misconceptions (e.g. Novak and Gowin 1984; Novak 1998, Kinchin and Alias 2005; Fisher 2000). If it is used as we have used it here then it combines processes of teaching with research of learning. For a sector that purports to foster both the process of knowledge transfer and its creation, this is important.

Other important issues might also be better understood through long term concept mapping studies. The timing and nature of assessment, for example, is a much debated topic (see reviews by Rowntree 1987; Brown and Knight 1994) and the case of Rose (this study) provides a focus for many of these concerns. Rose failed the course and analysis of her final concept map explains why. She started from a poor prior knowledge structure but was clearly even more confused at the end of the course than she had been at the beginning. However, she was also the only student to radically alter her knowledge structure during her learning. She was tested in the middle of this process of change but this appears to have been done at a time when her knowledge had not yet ‘broken through’ into new understanding. If she had been supported more, and tested later, then it is quite possible that she would have been able to restructure her learning sufficiently to pass the module. Her concept maps suggest that she was trying to learn in meaningful ways but that she had not had time for this in the duration of the course (see Scheja 2006, for a discussion of delayed understanding). Others students (like Tammy) finished the course knowing little more than they did to begin with, but they still did well. Probably Tammy could have written her essay on the day she started and she still would have passed. In the context of widening participation it is clear that a university education now needs to provide more opportunities for all and not just for those who will do well anyway. Concept mapping has an important role to play because it is a means of identifying students at risk of failure (as a consequence of their prior knowledge structures) and it can be used to show the ways that they go about linking new knowledge to old in their learning.

Weaknesses and limitations of the method

Thus far we have argued that concept mapping has a unique role to play in higher education. It is also important, however, that we describe some limitations and possible criticisms of the method. Concept mapping is useful compared with other graphic methods of knowledge declaration (like mind mapping) because it requires explicit statements of meaning. It makes visible the structure of knowledge as well as its content and it measures knowledge and understanding separately (since ideas are expressed through concept labels and meanings are defined by the linkage between concepts). In this regard, however, concept mapping may encourage explanations of meaning that are trivial compared with the rich expositions of personal understanding obtained through interviews or essay writing (Entwistle pers. com.). This is a valid point and one that should be addressed by studies using interviews, written narrative accounts and concept mapping together. We have begun to do this in our own research, but at this early stage, we can only suggest our findings are mixed. It is already clear, however, that some students with poor essay writing skills are sometimes better able to explain their understanding using concept mapping but the interpretation of this finding is complicated by several issues. (1) Comparison with interviews is difficult because interview transcripts can include contradictions that are largely prohibited in concept mapping. (2) Interviews rely on prompts from the researcher (where concept mapping does not). (3) Compared with essay writing, concept mapping emphasises personal constructions of meaning where narratives can be constructed in the

absence of understanding. These are issues that are likely to be rich avenues of research in the future.

It is also possible that concept mapping produces artefacts that arise from familiarity with the method (rather than being genuine indicators of learning). In this case students' second maps might be richer, not because they have learnt, but merely because they have mastered concept mapping. This is an important contention but it is not substantiated by our own research. To address this issue empirically, Hay (reported in Hay and Kinchin 2007), obtained concept maps from 277 students at the end of a 10 week course in statistics. Half of this student group had been taught concept mapping to begin with (and had made maps of their prior-knowledge), and half had not. At the end there were no quantitative differences (using measures of concept or linkage richness and map complexity), nor were there obvious differences in map quality (using measures of map structure (Kinchin et al. 2000) or the validity of propositions (Hay, Kehoe et al. *in press*). Concept mapping is so simple that there need-not be any significant learning-curve for the method. This does not mean that people don't learn through concept mapping, however. Making concept maps is a process of rehearsal and if it is done often enough then people learn from the reflection this entails. As a method for researching learning, concept mapping is not objective because the process of learning measurement (using concept mapping) is bound-up with the learning process itself. If this is true of concept mapping, however, it is true wherever we accept that verbal expressions are indicative of innate aspects of mind and abstract mental processes (Chomsky 1965). For use in teaching, however, this is also a strength, because it means that teachers can encourage student-centred meaning-making and record its outcomes simultaneously.

Interviewing students repeatedly to discover their mental constructions of the topics they are learning may be richer than concept mapping (and is probably embedded in the one-to-one teaching methods of elite institutions), but interviews are more time-consuming than concept mapping.

Conclusion

This paper provides empirical evidence for change that is indicative of different qualities of learning. We find that for most students the quality of learning is mixed (combining meaningful learning with rote learning or rote learning with non-learning), and that the richness of student-prior knowledge is a significant determinant of the quality of student learning.

Further work needs to track students' knowledge-change through time, and to develop and test concept mapping to these ends, but we suggest that the concept mapping method has a unique utility for longitudinal research of learning among university students. Concept mapping is a means of engaging students *and* teachers in transfer of target knowledge and the creation of personal understanding; it helps students to learn meaningfully and helps teachers research how the content of teaching is understood. It is therefore a bridge between processes of personal meaning-making and the acquisition of targets acknowledged by the curriculum (Entwistle and Smith 2002; Kinchin and Hay 2007). Furthermore we attest that concept mapping is a ground for a pedagogy common to all the cognate disciplines of higher education. This is because it rests on notions of personal change (as measures of learning) rather than psychological traits or 'approaches' to learning or the acquisition of targets per se. Novak's concept mapping method and his definition of meaningful learning (Novak 1998) acknowledged the simultaneous roles of

target knowledge in teaching and the personal meaning-making that must take place if teaching leads to learning.

Acknowledgements The work reported here was funded by a grant for the research of pedagogy from the Society for Educational Studies.

References

- Ausubel, D. P. (1963). *The psychology of meaningful verbal learning*. New York: Grune and Stratton.
- Ausubel, D. P. (1968). *Educational psychology: A cognitive view*. New York: Holt, Rinehart and Winston.
- Ausubel, D. P. (2000). *The acquisition and retention of knowledge: A cognitive view*. Dordrecht: Kluwer Academic Publishers.
- Ausubel, D. P., Novak, J. D., & Hanesan, H. (1978). *Educational psychology: A cognitive view*. New York: Holt, Rinehart and Winston.
- Biggs, J. (1993). What do inventories of students' learning processes really measure? A theoretical review and clarification. *British Journal of Educational Psychology*, 63, 3–19.
- Biggs, J. (2003). *Teaching for quality learning at University: What the student does*. Buckingham: Society for Research into Higher Education & Open University Press.
- Brown, S., & Knight, P. (1994). *Assessing learning in higher education*. London: Kogan Page.
- Buzan, T., & Buzan, B. (2000). *The concept map book*. London: BBC Worldwide Ltd.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge: MIT Press.
- Entwistle, N., McCune, V., & Walker, P. (2001). Conceptions, styles and approaches within higher education: Analytic abstractions, everyday experience. In R. J. Sternberg & L.-F. Zhang (Eds.), *Perspectives on thinking, learning and cognitive styles* (pp. 103–136). London: Lawrence Erlbaum.
- Entwistle, N. J., & McCune, V. (2004). The conceptual basis of study strategy inventories. *Educational Psychology Review*, 16, 325–345.
- 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.
- Entwistle, N. J., & Smith, C. A. (2002). Personal understanding and target understanding: mapping influences on the outcomes of learning. *British Journal of Educational Psychology*, 72, 321–342.
- Entwistle, N. J., & Tait, H. (1994). Approaches to studying and preferences for teaching in higher education. *Instructional Evaluation and Faculty Development*, 14, 2–10.
- Fisher, K. M. (2000). Overview of knowledge mapping. In K. M. Fisher, J. H. Wandersee, & D. E. Moody (Eds.), *Mapping biology knowledge* (pp. 5–23). Dordrecht: Kluwer Academic Publishers.
- Hay, D. B. (2007). Using concept maps to measure deep, surface and non-learning outcomes. *Studies in Higher Education*, 32(1), 39–57.
- Hay, D. B., & Kinchin, I. M. (2006). Using concept maps to reveal conceptual typologies. *Education and Training*, 48, 79–83.
- Hay, D. B., & Kinchin, I. M. (2007). Using concept mapping to make concrete measures of learning quality in higher education. *Proceedings of the European Learning Styles Conference*, Dublin, Ireland, June 11–14, 2007.
- Hay, D. B., Kehoe, C., Miquel, M. E., Kinchin, I. M., Hatzapinagos, S., Keevil, S. F., & Lygo-Baker, S. (in press). Measuring the quality of e-learning. *British Journal of Educational Technology*.
- Hay, D. B., Kinchin, I. M., & Lygo-Baker, S. (in press). Making learning visible: the role of concept mapping in higher education. *Studies in Higher Education*, 33(3).
- Jarvis, P. (1992). *Paradoxes of learning*. San Francisco: Jossey Bass.
- Jarvis, P. (1998). *From practice to theory*. San Francisco: Jossey Bass.
- Jarvis, P. (2006). *Towards a comprehensive theory of human learning: Lifelong learning and the learning society*, vol. 1. London & New York: Routledge.
- Jarvis, P., Holford, J., & Griffin, C. (1998). *The theory and practice of learning*. London: Kogan Page.
- Kember, D. (1996). The intention to both memorise and understand: another approach to learning? *Higher Education*, 31, 341–351.
- Kember, D. (2000). Misconceptions about the learning approaches, motivation and study practices of Asian students. *Higher Education*, 40(1), 99–121.
- 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, 569–591.

- Kinchin, I. M., & Hay, D. B. (2007). The myth of the research-led teacher. *Teachers and Teaching: Theory and Practice*, 33(1), 43–61.
- 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.
- Kinchin, I. M., Lygo-Baker, S., & Hay, D. B. (in press). Universities as centers of non-learning. *Studies in Higher Education*, 33(1).
- Kolb, D., & Fry, R. (1975). Towards and applied theory of experiential learning. In C. L. Cooper (Ed.), *Theories of group processes* (pp. 103–136). London: J. Wiley and Sons.
- Laurillard, D. (2002). *Rethinking university teaching: A framework for the effective use of learning technologies* (second ed.). London and New York: Routledge/Falmer.
- Marton, F. (1981). Phenomenography—describing conceptions of the world around us. *Instructional Science*, 10, 177–200.
- 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* (pp. 36–55). Edinburgh: Scottish Academic Press.
- 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).
- Meyer, J., & Land, R. (2005). *Overcoming barriers to student understanding*. London: RoutledgeFalmer.
- 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* (pp. 405–447). Orlando: Academic Press.
- 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*. Mahaw: New Jersey & London' Lawrence Erlbaum.
- Novak, J. D., & Gowin, D. B. (1984). *Learning how to learn*. Cambridge: Cambridge University Press.
- Novak, J. D., & Musonda, D. (1991). A twelve-year longitudinal study of science concept learning. *American Educational Research Journal*, 28(1), 117–153.
- Novak, J. D., & Symington, D. J. (1982). Concept mapping for curriculum development. *Victoria Institute for Educational Research Bulletin*, 43, 3–11.
- Perkins, D. (2006). Constructivism and troublesome knowledge. In Meyer, J. H. F., & Land, R. (Eds.), *Overcoming barriers to student understanding: threshold concepts and troublesome knowledge* (pp. 33–47). London: Routledge.
- Prosser, M., & Trigwell, K. (1999). *Understanding learning and teaching: the experience in higher education*. Buckingham: Open University Press.
- Ramsden, P. (2002). *Learning to teach in higher education*. London & New York: Routledge.
- Rowntree, D. (1987). *Assessing Students: How Shall We Know Them?* London: Kogan Page.
- Säljö, R. (1975). *Qualitative differences in learning as a function of the learner's conception of the task*. Gothenburg: Acta Universitatis Gothoburgensis.
- Scheja, M. (2006). Delayed understanding and staying in phase: Student's perceptions of their study situation. *Higher Education*, 52, 421–445.