INT. J. SCI. EDUC., 16 JULY 2004, VOL. 26, NO. 9, 1043–1064

Improving the usefulness of concept maps as a research tool for science education

Els Van Zele, Department of Solid State Physics, CRIPE; e-mail: Els.VanZele@UGent.be; Josephina Lenaerts, Department of Solid State Physics, CRIPE and Willem Wieme, Department of Applied Physics, CRIPE Ghent University, B-9000 Ghent, Belgium

The search for authentic science research tools to evaluate student understanding in a hybrid learning environment with a large multimedia component has resulted in the use of concept maps as a representation of student's knowledge organization. One hundred and seventy third-semester introductory university-level engineering students represented their understanding related to the atom in a concept map. A qualitative analysis of the data presented a more informative and complete picture of students' understanding as compared with most prominent scoring techniques in the literature. The detailed analysis method is illustrated by means of some students' concept maps and the overall results of the entire student population are discussed. The most striking difference between common and more insightful maps is the inclusion of additional clusters describing models of the atom, beyond the clusters that just describe the elements of atoms (e.g. protons, electrons, etc.) and their characteristics.

Introduction

In the 1980s, Novak outlined the potential use of concept mapping for the improvement of teaching and meaningful learning (Novak and Gowin 1984). A concept map represents a person's structural knowledge about a certain concept or subject. Crucial terms (named concepts) are related by means of explanatory links (named propositions) that declare the relationships between those concepts: for example, components, characteristics, model, and so on. Figure 1 shows an example of a student concept map for the atom. Since then, a substantial bibliography related to concept maps has been assembled (Novak 1990, White and Gunstone 1992, Wandersee et al. 1994, Mintzes et al. 2000), offering ample information about their usefulness as instructional tools (Trowbridge and Wandersee 1994) and learning tools in different disciplines (Dykstra 1992, Esiobu and Soyibo 1995, Pearsall et al. 1997, Wallace and Mintzes 1990). Although the potential use of concept maps for assessing students' knowledge structure has been recognized, concept maps are far more frequently used as instructional tools than as assessment devices (Lomask et al. 1993, Ruiz-Primo and Shavelson 1996, Wilson, 1993).

According to constructivist learning theory, the learner's growing understanding of content knowledge is considered a process of enlargement and enrichment of the interrelations established between different kinds of information and eventually their integration into his/her existing knowledge framework (Glaser and Bassok

International Journal of Science Education ISSN 0950–0693 print/ISSN 1464–5289 online ©2004 Taylor & Francis Ltd
http://www.tandf.co.uk/journals
DOI: 10.1080/1468181032000158336

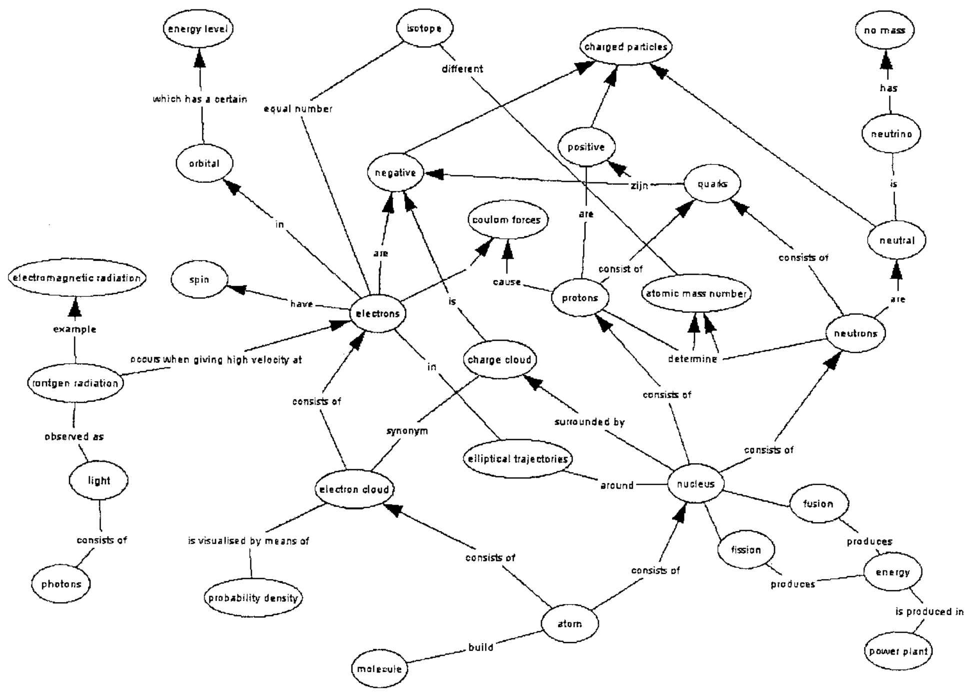

Figure 1. Example of a student's atom map.

1989). The quality of the learner's mental model either prevents or enables the accessibility of the processed information on later occasions (Baxter et al. 1996, Chi et al. 1981). Therefore learners should be carefully introduced to powerful meta-cognitive learning tools that enable them to build structural representations of the knowledge they are to acquire, such as concept mapping, in order to foster the transition from passive (rote) learning towards more engaged meaningful learning strategies. Misconceptions have proven to be very resistant to traditional instruction (Lenaerts and Van Zele 1998, McDermott and Redish 1999, Pfundt and Duit 1994, Van Zele et al. 2001), and even to inhibit further understanding if not appropriately dealt with. It will be demonstrated that concept maps can be useful tools to research student misconceptions.

Prominent concept mapping scoring systems

Distinct concept mapping scoring systems can be found in the literature. Novak and Gowin have proposed scoring criteria—on the specific request of teachers using concept maps as an assessment device—along with guidelines for constructing a concept map (Novak and Gowin 1984). While limited to hierarchical maps, this procedure is generally accepted as the most comprehensive scoring method (Ruiz-Primo and Shavelson 1996) and has been reported to work quite well as long as the assignment is well structured and of a 'closed format' (i.e. the map structure and the concepts are provided by the evaluator). Consequently, students merely 'fill in the map' (Shaka and Bitner 1996, Ruiz-Primo et al. 2001), which is indeed the strategy most frequently applied in research (McClure et al. 1999, Ruiz-Primo and Shavelson 1996).

Numerous other scoring schemes—many of them to some extent reducible to the Novak and Gowin scheme—have been devised. Generally, scoring systems for concept mapping can be categorized as either quantitative (i.e. making counts of characteristics; for example, McClure et al. 1999, Ruiz-Primo and Shavelson 1996) or qualitative methods (i.e. describing the content and quality of the map to some extent; for example, Hoz et al. 1990, Lomask et al. 1992, White and Gunstone 1992). In both types, individual maps may either be compared with a criterion map or not. For a detailed description of these scoring systems, we refer to the literature. A short description of these scoring systems, applied to a student's atom concept map (referred to as 'student X'), is provided later in table 5.

Quantitative scoring devices are generally more objective compared with qualitative analysis tools, which rely more on the expertise of the evaluator for interpreting the findings. But, due to the very nature of a concept map (i.e. a visual representation of one's knowledge structure related to a certain concept), a fairly objective analysis is possible. Compared with concept maps, the assumed knowledge structure interpreted by the instructor from reading a student's open response answer related to the same subject would be much less objective: there is a fair chance that the instructor intuitively adds and assumes links that are not mentioned or inadequately described in the student's response and that may not exist in the student's mind. Uncertainties and the nature of missing links (forgotten or unknown) emerging during the analysis of a student's concept map can only be clarified by means of a short interview with the student. The questions need to be very carefully chosen in order not to suggest relationships between concepts that are not shown or not entirely clearly explained in the student map.

The concept mapping assignment used in this study

The introductory quantum physics context

This research is placed in the context of a third-semester university-level introductory quantum physics course, in which students are to develop an understanding of basic quantum physics concepts, such as the atom. Concept mapping has been selected as a promising candidate to guide the investigation into students' structural representations of the atom.

Scope of this study

We wish to investigate whether and how concept mapping can be turned into an effective research tool from the perspective of Physics Education Research. For this purpose we have selected actual results produced by two students who we will conveniently call 'student X' and 'student Y'. This selection has been made with the view of obtaining the best possible illustration of the key features of our research. It is fair to say that student X's atom map (atom map X, shown in figure 2), is representative of the 10% of maps considered to be the best by three independent evaluators, while the map drawn by student Y (atom map Y, shown in figure 3) is representative of mediocre maps. The detailed data, derived from both atom maps, are presented later in tables 2 and 3 and will be used to materialize the discussion held in this paper. In addition, the general results from the 170 atom maps are shortly discussed.

Procedure

During the fall semester of 1999, 222 third-semester university-level engineering students enrolled in an Introductory Quantum Physics course were instructed to draw a concept map representing their ideas and knowledge structure about the atom. Effective concept mapping requires students to be familiar with the nature of concept maps, and therefore students were consecutively:

(a) instructed about the aims and the nature of the concept mapping technique, during a 45 min introduction by means of a sample map from the field of electricity—debating the importance of clearly stated propositions to describe the interrelationships between concepts—during a group discussion.

(b) provided with a non-exhaustive list with a fair number of suggestions for concepts related to the subject, leaving the students free to adopt these in their concept map, to leave them out or to include different ones (concept list reproduced in table 1). No restrictions were made concerning either the size or structure of the concept map.

- Students were provided with a computer program QMap 0.4 (for MS Windows) to facilitate repositioning of concepts and relationships dynamically, compared with its paper-and-pencil version equivalent. QMap 0.4 provides an electronic worksheet on which concepts can be put in ovals while propositions between any pair of concepts are indicated by means of labelled arrows. The format or layout of the concept map is not predetermined.
- Students were given a three-week time slot to develop and finalize their concept map. The instruction was given a fortnight before the subject was dealt with during the plenary lectures, and the assignment had to be handed in about one week after these particular lessons. The response rate to this optional assignment was 77% (170 student atom maps).

(c) interviewed (a limited number of students) in order to gain insight into the nature of missing links. A typical student interview consists of three different stages in which the questions posed become gradually more focused. First, students are asked whether and perhaps which additional links can be made between concepts used in the student's concept map. At this stage, students discuss the missing links they know about, but which they forgot to introduce in their map while making the assignment. In a second stage of the interview it is suggested that a proposition could be made with a specific concept (e.g. propositions with the concept [energy level]). In this stage, the interviewer and the student talk about relationships the student cannot describe accurately. In the final stage of the interview the student is asked about the missing link between two particular concepts (e.g. [photons] and [light]). Propositions that are discussed in this stage of the interview are fairly probably those the student is not confident enough about to mention or do not occur in the student's mind.

Table 1. Itemised incomplete, non-exhaustive list of concepts, part of students' atom map assignment.

The items (concepts) listed below are provided to help you construct a map describing the concept 'atom'. You do not need to use all items presented in the list, and you are free to add any other concepts that seem relevant

Circular trajectory	*Continuous spectrum*	*Coulomb force*	*Electromagnetic radiation*
Electron	Electron cloud	Elliptical trajectory	Energy level
Fission	Photon	Isotope	Nucleus
Crystal	Charge cloud	Light	Atomic mass number
Molecule	Negative charge	Neutral	Neutrino
Neutron	Uncharged	Orbital	Positive charge
Proton	Quarks	Roentgen radiation	Spin
State	Probability density	…	…

(d) given proper feedback about common misunderstandings and areas of clouded or unclear understanding.

The incomplete—non-exhaustive—list of concepts gives students some idea of how to start building their map, and starts all of them off from an (averaged) basic level of understanding, established from a pre-test at the beginning of the course. It further allows research into whether and how students use the provided concepts and whether they introduce additional concepts, not included in the list (e.g. [atom model]). For instance, a student using the concept [gluon] certainly must have read about it, as it was neither included in the list nor introduced during the lectures, while a student leaving out the listed concept [probability density] clearly failed to grasp the essence of our contemporary view on atomic models, all the more as this aspect was repeatedly emphasized during the lectures.

Collected data

Data from student X's and student Y's atom maps

Figure 2 presents student X's atom map and the data listing obtained from that map are reproduced in table 2. Atom map Y is reproduced in figure 3 and the corresponding data listing is given in table 3.

Student X's prior knowledge related to the atom

The students have been probed for their ideas related to the atom when first entering the Quantum Physics course. The answers from student X are used to better interpret the knowledge structure, demonstrated by atom map X, compared with these pre-test answers, reproduced in table 4.

Novak score for atom maps X and Y

The Novak score for student X's atom map (table 5, third row) has been calculated using the 12 different levels as replacement for the hierarchical levels because

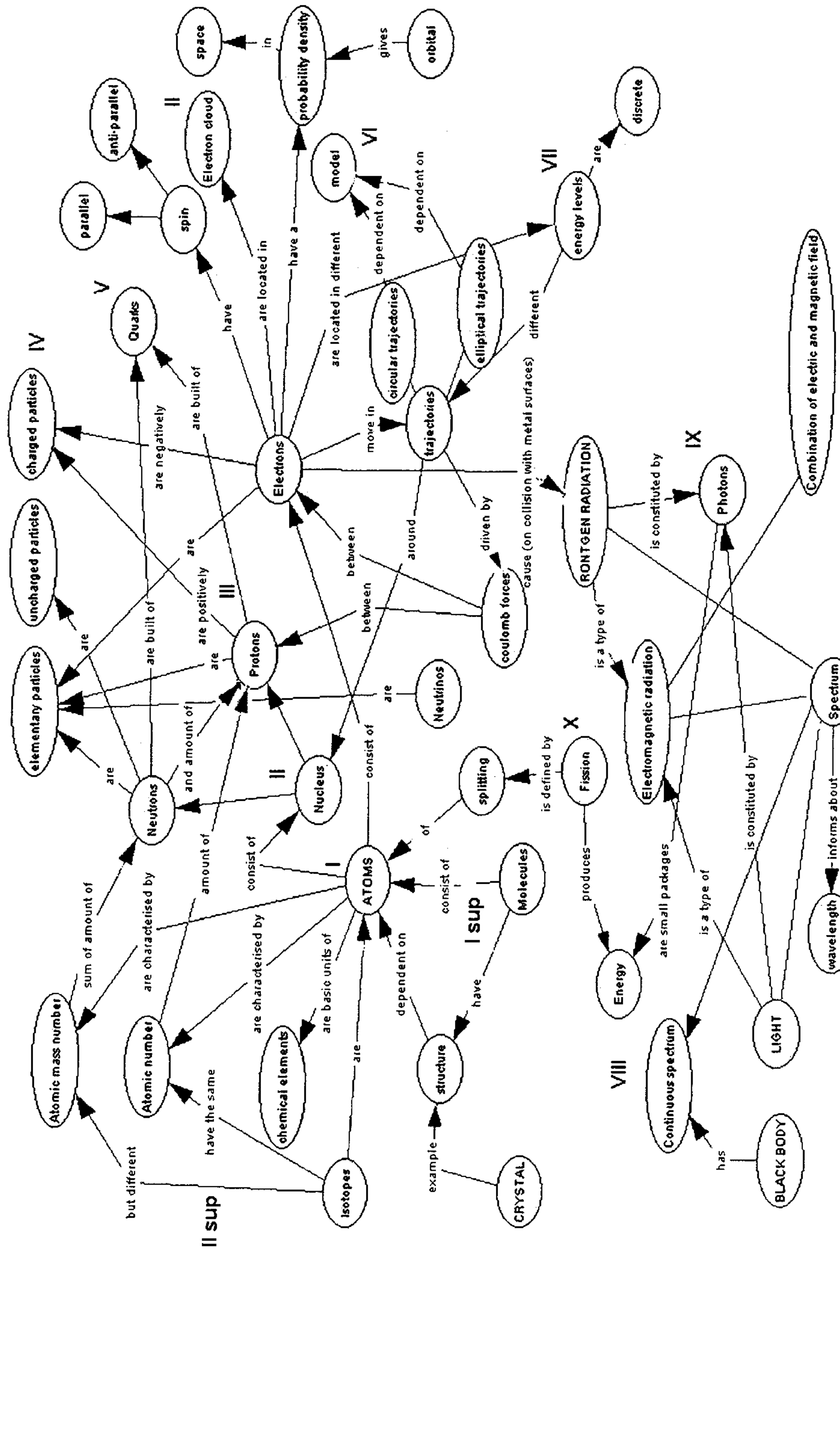

Figure 2. Student X's atom map.

Table 2. Data listing from student X's atom map.

Description			#
LISTED CONCEPTS USED			
all but two: [charge cloud] replaced by [electron cloud] and [state]			
WORDS and CONCEPTS ADDED			
[wavelength]—[energy]—[spectrum]—[combination of electric and magnetic field]—[discrete] [structure]—[chemical element]—[atomic number]—[proton]—[trajectory]—[model]—[space]—[parallel]—[anti-parallel]			
OTHER CONCEPTS, not introduced although logically expected			
[fusion], although [fission] is used			
[discrete spectrum], although [continuous spectrum] is used and the word [discrete] is added			
CONCEPTS			38
[black body] has been accepted as a concept			
parallel and *anti-parallel* have not been accepted as concepts			
discrete has been rejected as a concept, *discrete levels* would have been accepted			
crystal is categorized as an example, not accepted as a concept			
LINKS			56
wrong links	0		
No text links	9	e.g. [spectrum]—[light]	
weak links	0		
moderate links	4	e.g. [atom]—is a basic element of—[a chemical element]	
excellent vector links	43	e.g. [fission]—is—[splitting]—of—[atoms]	
missing links	6	e.g. [spectrum]—[energy levels], [wavelength]—[energy]	
NUMBER OF VALUABLE PROPOSITIONS			47
HIERARCHICAL (and semantic) LEVELS			12
Super-ordinate level 2:		Isotopes and their characteristics	
Super-ordinate level 1:		Molecule, structure	
Level 1:		Atom	
Level 2:		Nucleus and electron cloud	
Level 3—elements:		Protons—neutrons—electrons also neutrinos	
Level 4—characteristics:		Un/charged particles—elementary particles—spin	
Level 5—fundamental parts:		Quarks	
Level 6—modelling:		Trajectories—…—electromagnetic radiation	
Level 7—characteristics of modelling:		Spectrum, discrete energy levels	
Level 8 - functioning:		Roentgen radiation, wavelength, energy	
Level 9—characteristics of functioning:		Photons	
Level 10—technological applications:		Fission	
CROSS-LINKS			15
[atom]—[isotopes] and their characteristics (3 links)			
[fission]—[atom]			
[atom]—[electrons] since [electron cloud] is considered by the student as a characteristic rather than as a more general concept when compared to [electron]			

Table 2. Data listing from student X's atom map. (contd.)

Description	#
[protons, neutrons]—[quarks] (2 links)	
[electrons]—[model and trajectories] (3 links)	
[Coulomb forces]—[protons and electrons] (2 links)	
[electrons]—[roentgen radiation]	
[fission]—[energy]	
[trajectories]—[nucleus]	
EXAMPLES	1
crystal (example of a structure)	
[black body] has been accepted as a concept, and is therefore not an example	
CRUCIAL NODES	7
Atom—neutron—proton—electron—orbits—photon—spectrum	

this physics context (the concept of the atom) is not hierarchically structured. Consequently the propositions made between concepts that belong to different aspects (e.g. levels 3 and 5, levels 1 and 10) are used as cross-links. Propositions made between concepts of levels 3 and 4 (i.e. elements and their characteristics), levels 6 and 7 (i.e. modelling and their characteristics) and levels 8 and 9 (i.e.

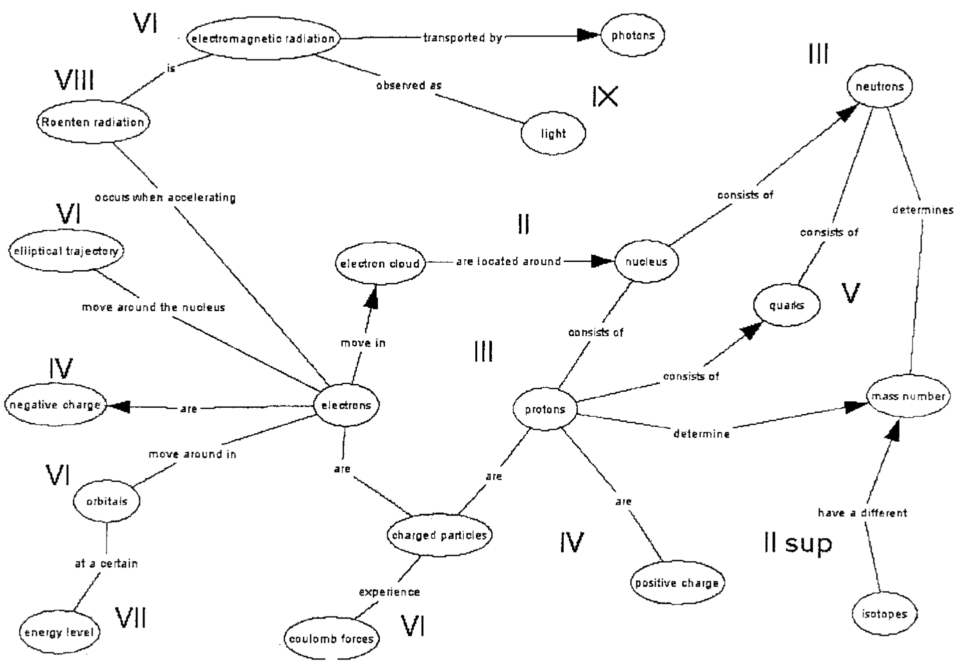

Figure 3. Student Y's atom map.

Table 3. Data listing from student Y's atom map.

Description		#
CONCEPTS		**19**
LISTED CONCEPTS USED		*17*
WORDS and CONCEPTS ADDED [charged particles] and [nucleus]		*2*
OTHER CONCEPTS, not introduced although logically expected [atom] , [energy] and [wavelength]		*3*
NUMBER OF VALUABLE PROPOSITIONS		**20**
LINKS		
Wrong links 0		
No text links 0		
Weak links 0		
Moderate links 1	[electron] —move in —[electron cloud]	
Excellent vector links 19		
Missing links 4	[elliptical trajectory] and [orbitals] and [electron cloud] [photons] and [light] [energy levels] and [electromagnetic radiation]	
HIERARCHICAL (and semantic) LEVELS		
Super-ordinate level 2:	Isotopes and their characteristics	
Level 2:	Nucleus and electron cloud	
Level 3—elements:	Protons—neutrons—electrons also neutrinos	
Level 4—characteristics:	Un/charged particles—elementary particles—spin	
Level 5—fundamental parts:	Quarks	
Level 6—modelling:	Trajectories—…—electromagnetic radiation	
Level 7—characteristics of modelling:	Spectrum, discrete energy levels	
Level 8—functioning:	Roentgen radiation, wavelength, energy	
Level 9—characteristics of functioning:	Photons	
CROSS-LINKS [proton] and [neutron]—[quarks] (2 links) [proton] and [neutron]—[mass number] (2 links) [electromagnetic radiation]—[photons] and [light] (2 links) [electron]—[elliptical trajectory] and [orbital] (2 links) [electron]—[roentgen radiation] (1 link) [charged particles]—[coulomb forces] (1 link)		**10**
CRUCIAL NODES proton—electron—neutron—electromagnetic radiation		**4**
EXAMPLES		**0**

functioning and their characteristics) are discarded as cross-links as these provide definitions for the relationships between the linked concepts. The Novak score for atom map X is 258, atom map Y scores 170, while all Novak scores range from 110 to 300.

Table 4. Student X's pre-test answers related to the atom.

Student X describes 'concepts and ideas about the atom':

'Atoms are very tiny basic elements of about everything. An atom consists of a nucleus (built of protons and neutrons) surrounded by electrons. Protons and electrons have opposite elementary charges (elementary charge = e). Atoms can bind and form complex structures that make up the macroscopic materials we know or use. The study of atoms has brought about many developmental stages in all sciences, e.g. medicine, …'

Student X's answer related to 'the magnitude of the atom':

'The atom has a magnitude of about 10^{-20} m'

In positioning the following four statements between right and wrong:

Statement 1:

'An atom has the same characteristics as the material it constructs'

'The characteristics of an atom may cause the macroscopic characteristics of the material that it constitutes'

Statement 2:

'An atom is not an existing entity, but merely a concept used to explain certain phenomena'

'I don't agree since it has been demonstrated by means of experiments that the atom is an existing entity'

Statement 3:

'An atom model provides a convenient description of an atom, but the atom itself might appear to be different than the model'

'This is something we will probably never know'

Statement 4:

'We describe atoms in some way because we can't see them'

'This is true'

Results and discussion

To research how concept mapping can be turned into an effective research tool from the perspective of Physics Education Research, a comparison is made between the data obtained from a quantitative analysis (e.g. the prominent scoring systems reported in literature) and from a qualitative analysis, based on phenomenological considerations (Marton 1981).

Scoring student X's atom map

The different scoring techniques (reported in the literature) applied to atom map X (table 5, results atom map X) all weigh the nature and quality of a concept map according to different measures. The quantitative scores obtained for atom map X range from 0.06 (congruence score by Ruiz-Primo and Shavelson) to 0.77 (salience score by Ruiz-Primo and Shavelson) for relative measures (i.e. calculating a fraction of different aspects) and vary between 8 (holistic map score according to McClure et al.) and 258 (typical Novak and Gowin score) for absolute measures (i.e. counting certain aspects). The obtained qualitative approaches for atom map X (table 5) equally result in a diversity of scores too, as different measures evaluate

Table 5. Various scoring techniques applied to student X's atom map.

Reference	*Scoring system*	*results atom map X*
Quantitative approach		
Novak and Gowin (1984)	(1 point x count of valid student propositions)	(1 × 47)
	+ (5 points × count of hierarchical levels)	+ (5 × 12)
	+ (10 points × count of cross-links)	+ (10 × 15)
	+ (1 point × count of examples)	+ (1 × 1)
	= weighted student sum	**= 258**
Ruiz-Primo and Shavelson (1996)	Prepositional accuracy: Sum of individual proposition scores obtained from the student map	47
	Convergence: amount of accurate propositions in the student map divided by the total number of possible propositions in the student map	43/703 (= **0.06**)
	Salience: Number of correct student propositions divided by the total of student propositions from the student's map	43/56 (=**0.77**)
	Alternative: Novak score divided by weighted sum of criterion map	258/(weighted sum of criterion map)
McClure et al. (1999)	triple sum of valid propositions	3 × 47 = 141
	structural map (identical to Novak & Gowin)	258
	holistic map: general understanding (0 - 10)	**8**
	relational map	146
Qualitative approach		
Hoz et al. (1990)	1) propositions	3
	2) map as a whole: division of count of student links by count of links in criterion map	56/(propositions in criterion map)
	3) division of count of valid student links by count of links in student map	47/56
	4) concept groups:	
	homogeneity (3 level ordinal scale)	2/3
	structure (4 level ordinal scale)	3/4
	title fit (4 level ordinal scale)	not provided
Lomask et al. (1992)	1 score based on a completeness – strength matrix	5
	strength / size: Strong, medium, weak, none	
	complete 5 4 3 2	
	substantial 4 3 2 1	
	partial 3 2 1 1	
	small 2 1 1 1	
	none 1 1 1 1	
White and Gunstone (1992)	(1) count of student propositions corresponding to propositions in the criterion map	resemblance with criterion map
	(2) additional points for insightful links (positive) or incorrect links (negative)	+ 3: electromagnetism cluster – 5: missed links

Lomask et al. (1992) completeness – strength matrix:

strength / size	Strong	medium	weak	none
complete	5	4	3	2
substantial	4	3	2	1
partial	3	2	1	1
small	2	1	1	1
none	1	1	1	1

different aspects of the concept map. This substantially obscures interpreting these results and particularly prohibits comparison between results obtained from distinct methods.

Some critical reflections

The actual performance of currently available scoring techniques (table 5) was felt unsatisfactory for analysing concept maps:

- *Concept mapping as an assessment tool*: the very nature of concept mapping seems to be conflicting with the way most prominent scoring techniques address concept maps as assessment tools. Traditionally, assessment mostly seems to be thought of as grading, deciding upon which score should be awarded to a student's concept map. Prominent scoring techniques, although all valuable in their own right, seem to overlook the actual knowledge structure as most base their score on counts of propositions and concepts. It is thought that as an assessment technique, the score should represent students' understanding and their subsequent knowledge structure.
- *Concept maps as a research tool*: it is felt that a quantitative approach does not truthfully represent the potential of a concept map as a demonstration of a student's knowledge structure while a more detailed picture of the student's understanding can be extracted from a qualitative analysis, describing explored ideas and accounting for mistakes, 'shortcomings' and possible misconceptions. Such a procedure is proposed later.
- Nevertheless, concept mapping may perfectly fit into the philosophy of *permanent evaluation*, similar for instance to portfolio assessment. Even when summative assessment is intended, a solid research method for concept maps will facilitate decision-making on which standards to assign the final grade.
- *Criterion or master map*: relative scoring techniques divide the score obtained from a student's map by the score of a criterion or master map. Two types of master maps have been proposed in the literature: an absolute master map, representing the essential knowledge a student should acquire throughout the course; and a relative map, composed from maps drawn by the top 5–10% best students (McClure et al. 1999). One could argue that absolute master maps are consistent with the setting of clear standards and that they allow positioning students' intermediate level of understanding, relative to the ultimate goals intended. Nevertheless, relative scores expect students to structure their knowledge in a preformatted way (i.e. a student gets no points when not applying a predefined link in the map, while he/she may describe the same relationship by means of other propositions, not made in the master map). For instance, this is the case with this test group as only 3% of the sample population has used the concept [model] to express information about the shape of the atom, while 91% of the students provided some kind of description without paraphrasing this concept. The use of a criterion map would only award 3% of the students for their insight, which makes a comparison with such a map at least awkward. Consequently, master scores have not been calculated for atom map X in table 5.

For all these discussed reasons, we feel that a different approach towards the analysis of concept maps from the perspective of Physics Education Research may be useful.

A possible alternative research method

As students progress, the structure of their concept map reflects the increasing complexity, (hopefully) resembling more and more the knowledge structure of an expert. It might be argued that this complexity could in the end prohibit an effective use of the concept mapping procedure. However, according to the theory of phenomenography (Marton 1981), the number of fundamentally different ways of understanding in the analysis of people's statements concerning the same phenomenon is strictly limited. While individuals' observations and reasoning about a phenomenon indeed create diversification, this leads to no more than five or six distinct 'lines of explanation'. Taking this into account a collection of concept maps will undoubtedly diverge, but only by a limited set of properties. Efficient coding of the concept map can then be guaranteed, allowing for objective analysis and interpretability. A quantitative data analysis allows drawing frequency tables of applied concepts and occurring propositions, giving a first impression of the size and density of a concept map, while a qualitative data analysis allows researching the students' content knowledge structure into more detail. In order to do this, we focus on its shortcomings and on the description of explored clusters. Each aspect will be discussed in detail by means of atom map X, and then the results from atom map Y will be interpreted.

Shortcomings

A concept map may include different kinds of shortcomings: straightforward mistakes, faulty or vague descriptions, misconceptions and either completely or partially deficient relationships. Knowledge of these shortcomings is important feedback for the instructor, enabling one to pinpoint students' difficulties and to track recurrent patterns or areas of vague or clouded student understanding. We focus the search for missing relationships on concepts present in the student map, which in the mind of expert scientists are interrelated but treated as separate, uncorrelated entities by the student, thereby revealing that only partial understanding has been achieved. Obvious examples from student X's map are the unconnected concepts [spectrum] and [discrete energy levels] or [wavelength] and [energy]. In atom map Y, [energy] is missing, as well as some kind of categorization for the concepts [elliptical trajectory], [electron cloud] and [orbitals].

Description of explored clusters

Every concept can be described by a set of properties such as its components and structure, its characteristics and dimensions, the scientific model used to represent it, and so on. The integration of such characteristics and properties reflects the quality of the student's knowledge structure. Collecting these in a database allows tracking changes over time and researching specific concepts and the way they are embedded in the students' knowledge structure. Furthermore the list enables comparison of specific characteristics between different maps. In atom map Y, the different levels are much more modestly described, sometimes only touched upon by means of a single concept (e.g. level 8, [Roentgen radiation]; and level 9, [light] and [photons]).

Application to atom map X

To illustrate the capabilities of the proposed qualitative research method, the summary from student X's atom map is discussed. Halfway through the Introductory Quantum Physics course, student X (hereafter referred to as 'he') has achieved quite a high level of understanding of the basic concepts relating to the atom. On the whole, he has demonstrated an excellent grasp of the requirements of a concept mapping assignment. He has added an impressive number of concepts not included in the list and he has made no clear-cut mistakes. Additional features worthwhile commenting on are:

- Some *missing connections,* particularly concerning concepts introduced by student X. As it is to be expected that students will introduce those concepts they feel sufficiently at ease with, one would expect relatively fewer shortcomings surrounding these concepts. Nevertheless student X made no propositions connecting the following concepts: [energy] and [wavelength]; [electromagnetic radiation] and [energy]; [photons], [energy levels] and [spectrum], [light] and [energy levels]; [discrete] and [continuous] energy levels; [trajectory] and [orbital] and [electron cloud] and [probability density].
- Some *potentially weak areas* show up in atom map X (reproduced in figure 2):
  - — at the left-hand side a partial electromagnetism cluster is located connecting the concepts [spectrum], [continuous spectrum] and [light], while at the right-hand side some additional concepts ([discrete (levels)] and [energy levels]) clearly belong to the same cluster—however, no propositions connect both sections;
  - — the concept [energy] is introduced by means of the propositions: [photons] are small packages of [energy], and [fission]—produces—[energy], but the student failed to link directly the basic concepts [energy] and [wavelength], although this is one of the basic ideas that shaped quantum physics.

While the foregoing are shortcomings in the eye of an expert, it might be understandable that students focus on the main assignment, subsequently forgetting to check introduced concepts for additional links. Student X commented during a short interview that he saw this assignment as part of a quantum physics course, addressing the atom as a central theme. He has also added relevant information from the Electromagnetism course, although he considered this as supplementary information, intended to convey the fact that his understanding transgresses the boundaries of different courses. This is notoriously difficult for students, and they should be forgiven if focus on the principal assignment is somewhat lost in that process.

- The *lack of non-cross-level links* (i.e. links made between concepts that belong to the same cluster) describing how electrons 'belong' to an atom nevertheless points to some degree of misconception. It suggests that student X has established a dual image concerning the way electrons behave, allowing the coexistence of different concepts such as [trajectories], [orbitals], [electron cloud] and [probability density]. At least five parallel 'lines of thinking' can be reconstructed from his atom map:

1. [electrons]—move around—[the nucleus] in [trajectories] controlled by [Coulomb forces]—acting between—[protons] and [electrons];
2. [electrons] are located in an [electron cloud];
3. [electrons] display a [probability density] defined in [space] caused by [an orbital];
4. [electrons]—are located in different—[discrete]—[energy levels];
5. The shape of the trajectories is defined by the model: [electrons]—move in [trajectories] being either [circular] or [elliptical]—dependent on [the model].

In this area of atom map X (right-hand side of figure 2) all concepts expected in the cluster have been included, but the lack of interconnectivity nevertheless points to incomplete understanding. Given a free response format assignment, such shortcomings would probably go unnoticed as student X masters the right vocabulary.

The qualitative analysis, along with its interpretation—based on both the student map and the subsequent interview held with the student—provides the researcher with a source of evidence on the knowledge structure demonstrated in the atom map and, consequently, enables a clearer view on the capabilities of the student. This approach could enhance the usefulness of concept mapping as a tool for Science Education Research and even strengthen the use of concept maps as a formative assessment tool. It is further believed that such an analysis is more suitable to be used when it finally comes to grading a concept map compared with scoring devices suggested in the literature (table 5).

Application to atom map Y

In atom map Y, the knowledge structure is less pronounced: some levels are merely introduced by only a few concepts (e.g. level 8, [Roentgen radiation]). The electromagnetism cluster is positioned in two corners of the map: [energy levels] at the bottom on the left-hand side, while [electromagnetic radiation], [photons] and [light] are located on top of atom map Y. The different representations of how electrons 'belong' to an atom are not connected either. These are merely seen as different ways of presenting the position of an electron around the nucleus, without classifying these as atom models. The concept [elliptical trajectory] has been adopted, whereas [circular trajectory] has not. According to student Y—in the subsequent interview—elliptical trajectories give a much more accurate description, compared to circular trajectories. Therefore student Y did not mention [circular trajectories] in atom map Y.

General summary of the 170 atom maps

A quantitative as well as a qualitative analysis has been conducted to all 170 atom maps, yielding a massive amount of data. From every concept map all words (concepts, examples, etc.) and descriptive relationships (propositions) have been extracted. These data have been numbered and organized in a data matrix amenable to statistical evaluation using SPSS 10.0. The collective data from all maps yield frequency tables of the concepts as well as of the propositions. As illustrated in

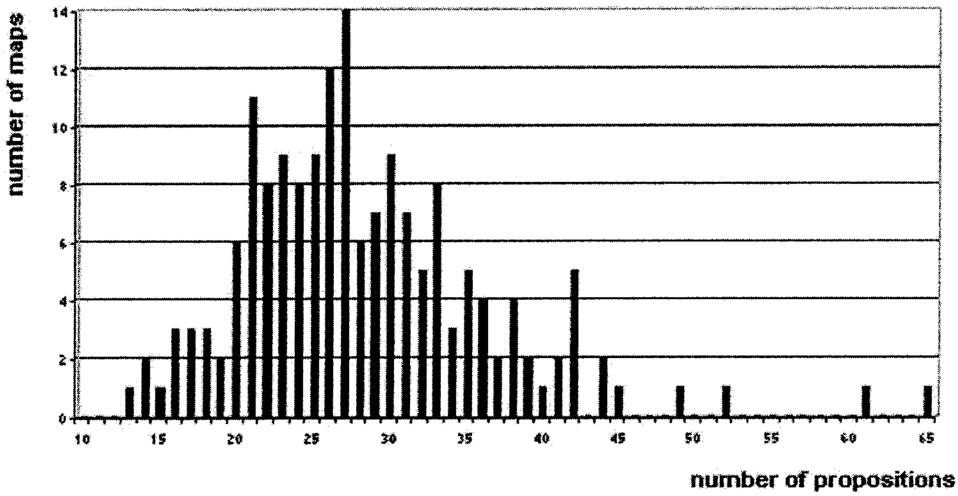

Figure 4. Distribution of number of propositions used in each student's atom map.

figure 4, the number of propositions in each map varies considerably. All together, 4790 propositions have been recorded, equivalent to a mean value of 28 (standard deviation $\sigma = 8.2$) propositions per atom map.

The qualitative analysis of atom map X fairly well represents the different aspects found in the collection of 170 students' atom maps, especially the different 'lines of thinking' related to how electrons 'belong' to an atom. A particular shortcoming, evident from all 170 atom maps, is the absence of any information concerning the size of the elementary particles. This finding has been corroborated during subsequent interviews with students as most of them would not even hazard a guess as to the order of magnitude. As the same students have proven to be proficient rote learners on many other occasions, this can only mean that instruction has miserably failed to impress on them the expert's view concerning the importance of such data.

Some observations related to the students' use of concepts such as the listed concept [molecule], the unlisted concept [atom model] and the particular way in which students describe the atom model without introducing the concept [atom model] are now described.

- Since the concept [molecule] has been provided in the list, one could expect students to feel compelled to use this concept, regardless of whether they are fully aware of its relevance with respect to the overall assignment:
  - — 131 students (77%) did use the concept [molecule];
  - — 205 propositions have been made using the concept [molecule].
  - — the propositions mostly relate to the concepts [atom] (59%) or [crystal] (24%); and
  - — other concepts (e.g. amorphous state, emission spectrum, electron, electron cloud, energy level, nucleus, etc.) have been linked to the concept [molecule], but only infrequently (two or three times for each map).

With regard to the concept [molecule], the 131 students' responses are surprisingly homogeneous, most often repeating a rather vague description that can be found in elementary treatments of properties of matter in high school physics or chemistry courses. This finding is suggestive of unproductive rote learning.

- Only 3% (five students out of 170) introduced the non-listed concepts [atom model] or [model] in their map. This suggests these students classified the historical models as [atom model], a higher level of classification encountered in taxonomies of educational objectives (Bloom 1956).
- A much larger group of 155 students (91%) did, however, expand on the listed concepts [circular trajectory] and [elliptical trajectory] to arrive at some kind of either concept or commentary describing how atoms could be conceptualized. This observation is believed to be critical, because an entire chapter of the Introductory Quantum Physics course deals with the crucial aspect of conceptualizing and modelling the atom.
- With only 3% of the students spontaneously introducing the concept [model], one might reasonably assume that at least some students have paraphrased it without using the specific wording 'model'. Propositions including [atom], [electron] or [electron cloud] would appear to be good candidates in this respect. Indeed, 155 students used such propositions. The most often made propositions are [an atom]—has a—[nucleus] (126 students), [electron clouds]—are made of—[electrons] (115 students), [an electron]—[moves in/ is located in]—[an orbital] (92 students), and [an electron]—[is located on]—[a circular orbit] (74 students).

These descriptions of how electrons are arranged around a nucleus convey some kind of impression about the shape or the appearance of an atom but, without reference to the concept [model], this remains unscientifically vague. More importantly, it points to the fact that students failed to classify this description as a model. These findings may imply that students structure the content at hand as it is presented, without wondering to which abstract category the studied topics belong, although the latter action is a rather crucial step in meaningful learning. This finding is corroborated by interviews and examination results, revealing that most students are tuned to study and remember for the specific purpose of passing the examinations, and only very few aspire to interpret what they learn.

The most typical difference between rather small and more substantive maps is that smaller atom maps—containing between 10 and 20 propositions—show five to seven semantic levels (superordinate level 1 and six subordinate levels, i.e. level 1–6), while the larger atom maps—containing 20–40 propositions—include all clusters and use much more propositions per cluster. Small maps quite completely explore the atom and its elements, their characteristics and its components, but only demonstrate small parts of the modelling level, nevertheless referring to an [electron cloud] model (the semantic levels are presented in table 2: datasheet for atom map X in the section on hierarchical and semantic levels).

Many maps show propositions made with electron clouds and probability density plots, raising the impression that the students have firmly grasped these concepts but, similar to student X's atom map, these propositions are only seldom categorized as distinct atom models. Typically, students have been found to only seldom introduce the concept model themselves during a short oral explanation (i.e. a small interview) about their atom map. Most students only used the concept [model] once the evaluator introduced it.

Concept mapping compared with open response questions

A comparison is made between the data obtained from student X's pre-test entries in the database (table 4) related to the atom model and the data that can be distracted from concept map X (table 2) in order to understand what makes concept mapping so interesting in respect to a student's knowledge structure. The free response answers from the pre-test more resemble a small collection of facts related to the atom: the properties of the atom are described by means of levels 1 (atom) and 2 (nucleus), and levels 3 (protons, neutrons and electrons) and 4 (the corresponding charges). The atom is mentioned as a basic particle of materials (level 1_{sup}) and the idea that the study about atoms has helped evolve science in many ways (level 10) is introduced. The question related to magnitudes is not answered approximately correct and the relation between microscopic properties of the atom and the characteristics of materials is not clear to the student. The statements related to whether an atom really exists, whether it can be seen, and how this relates to an atom model leaves the student confused. On the one hand, student X knows that some effects of the existence of atoms have been experienced (e.g. Rutherford experiments, etc.) and he must have read somewhere that no one has yet 'seen' an atom, where 'seen' is interpreted as completely distinct from all adjacent objects. Typically, some beliefs and some confusion about the very nature of the atom leave the student incapable of identifying the usefulness and the typical characteristics of a model. Therefore, student X's free response questions barely reveal any significant information about the knowledge structure at that moment in time compared with the knowledge structure demonstrated by atom map X.

It can be argued that a pre-test will never be as informative as a formative or post-test question, but this argument does not explain the distinct nature of both tools: a free response question on the one hand registers in a linear order all aspects that come to the mind of the student and risks to reflect a line of associative thinking, leaving it to the reader to structure the described content knowledge. A concept map, on the other hand, is a graphical tool, specifically aiming at constructing a two-dimensional representation of the links between different concepts. Concept mapping combines two ways of presenting content knowledge: visual (i.e. the knowledge structure) and verbally expressed information (i.e. propositions) (Mayer and Sims 1994, Paivio 1986). Therefore, the concept mapping task better demonstrates and consequently promotes the development of a coherent knowledge structure—a crucial aspect of meaningful learning.

Critical analysis of this study

It is felt that a typical quantitative analysis, yielding overviews of occurrences of various concepts and showing histograms of counts of concepts and amounts of propositions, yields too little information, compared with the time spent to gather it, and therefore a qualitative analysis might be better suited to research students' knowledge structure.

Most physics contexts are not hierarchically structured and therefore the Novak score for student X's atom map has been calculated with the semantic levels as replacements for the hierarchical levels. As a result of this choice, cross-links (propositions between different semantic levels) are overemphasized with an additional weight (factor 10), compared with all other propositions (factor 1) in the Novak score. This choice does not influence the occurring problems reported in this study related to quantitative scoring devices.

Although the performed qualitative analysis is relatively objective—due to the nature of a concept map—it is left to the evaluator to interpret the findings. For instance, the lack of coherence in the electromagnetism cluster may be perceived as a result of the task setting (i.e. focus on quantum physics, not on electromagnetism) and therefore is much more acceptable compared with the lack of coherence between the different representations and atom models and the lack of information about sizes and magnitudes. The aspects on which to focus while conducting a qualitative analysis lie in the hands of the experienced lecturer or researcher (dependent on goals and objectives for the course).

Effective concept mapping requires students to be familiar with the concept mapping method, in order to have them focus on meaningful and insightful propositions, and therefore they should be taught how to draw concept maps. The benefit of the time spent is rather substantial: once students get acquainted with this meta-cognitive learning tool, they can readily use it to structure any amount of information, enabling them to put large structures of knowledge into perspective. In addition, the dual coding (i.e. visually and verbally expressed knowledge; see Paivio 1986) of information held within concept maps allows students to better access knowledge structures and more accurately retrieve useful information to enable transfer to new situations. Furthermore, concept mapping provides an attractive basis for collaborative brain-storming and discussion, enabling groups to build a shared understanding of a domain.

Concept mapping in secondary education

Concept mapping was originally developed (Novak and Gowin 1984) to research the growth of children's knowledge structure related to a certain subject. On the specific request of teachers—for the use of concept mapping as an assessment device—Novak proposed a quantitative scoring system. This paper debates a qualitative research method—used in a context of introductory physics courses at university—developed to reach behind the vocabulary and to research students' knowledge structures. These conditions described certainly also apply to the secondary education teacher, who is the expert in studying and assessing his/her students' performances in the classroom. Therefore, the outlined qualitative analysis might be suitable for a secondary education context. The performed analysis can easily focus on only a few aspects rather than on a complete analysis, to some extent similar to, for instance, Action Research. Choosing some crucial points to look into can easily focus the teacher's analysis and enable positioning students' understanding in relation to the intended goals for a particular series of lessons. For instance, a teacher might focus on the following questions:

- How do students deal with the numerous historical representations of the atom?
- How do students relate these different representations to the concept of an [atom model]?

Both suggestions are examples from the investigation reported in this paper that may perfectly fit into both the time schedule and the interest of the secondary education teacher.

Secondary education teachers associated with the teacher training programme at our university have had their pupils draw concept maps for years, which are then analysed by the prospective teacher students. In this manner, concept mapping is added to prospective teachers' instructional strategies.

For all these reasons, we feel that the outlined qualitative analysis of concept mapping might provide a valuable alternative to the numerous scoring techniques to which teachers are directed by literature.

Summary and conclusion

In this study, concept maps have been used to research students' ideas and knowledge structure related to the atom. A qualitative research procedure is proposed to research students' subsequent understanding.

While concept mapping has been a well-known meta-cognitive learning tool for more than two decades now, it is also undervalued as most scoring systems are rather one-dimensional.

The purpose of this paper has been to find proper ways to research the knowledge structure demonstrated by a student's concept map, and from that research to suggest a valuable method to research concept maps. From the concerns raised from a literature review of prominent ways of addressing concept mapping assignments, a research protocol for concept mapping has been outlined. A more complete picture of each student's understanding emerges from incorporating aspects such as the shortcomings and considering the content and structure of the concept map. This supplementary information is made accessible by providing students with a non-exhaustive list of concepts where some rather crucial concepts are deliberately omitted. As a result, an opportunity is created for evaluators to investigate which concepts students decide as appropriate or necessary to use.

Students at introductory university level have a fairly well established baseline understanding of the concept of the atom. More insightful maps incorporate atom models, although only seldom categorized in that way. The electromagnetism cluster is rather well documented in these maps, although sometimes fragmented while additional links are commonly introduced by the students during subsequent interviews. Rather weak atom maps merely describe the basic knowledge about the components of the atom by means of well-known definitions, without revealing much additional information, described by other semantic levels.

Comparison with free response questions reflects the alternative angle of approach of the concept mapping technique especially interesting for demonstrating the knowledge structure.

The general results from our engineering students test population and the extensive data extracted from student X's and student Y's atom maps show that concept mapping indeed is workable from the perspective of Science Education Research, both for large-scale research studies as for classroom situations, lifting concept mapping as a research tool far beyond its usually exploited potential.

References

BAXTER, G.P., ELDER, A.D. and GLASER, R. (1996). Knowledge-based cognition and performance assessment in the science classroom. *Educational Psychologist*, 31, 133–140.

BLOOM, B.S. (ed.) (1956). *Taxonomy of Educational Objectives: The Classification of Educational Goals: Handbook I, Cognitive Domain* (New York: Longmans).

CHI, M.T.H., FELTOVICH, P.J. and GLASER, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science,* 5, 121–152.

DYKSTRA, D.I. (1992). Studying conceptual change in learning physics. *Science Education,* 76, 615–652.

ESIOBU, G.O. and SOYIBO, K. (1995). Effects of concept and Vee mapping under three learning modes on students' cognitive achievement in ecology and genetics. *Journal of Research in Science Teaching,* 32, 971–995.

GLASER, R. and BASSOK, M. (1989). Learning theory and the study of instruction. *Annual Review of Psychology,* 40, 88–96.

HOZ, R., TOMER, Y. and TAMIR, P. (1990). The relationship between disciplinary and pedagogical knowledge and the length of teaching experience of biology and geography teachers. *Journal of Research in Science Teaching,* 27, 973–985.

LENAERTS, J. and VAN ZELE, E. (1998). Testing science and engineering students: the force concept inventory, *Physicalia Magazine,* 20(1), 49–68.

LOMASK, M., BARON, J.B., GREIG, J. and HARRSION C. (1992). *ConnMap: Connecticut's use of concept mapping to assess the structure of students' knowledge in science*. Paper presented at the annual meeting of the National Association of Research in Science Teaching, Cambridge, MA, 21–25 March.

LOMASK, M.S., BARON, J.B. and GRIEG, J. (1993). Assessing conceptual understanding in science through the use of two- and three-dimensional concept maps. *Proceedings of the 3rd International Seminar on Misconceptions & Educational Strategies in Science and Mathematics* (Ithaca, NY: Misconceptions Trust).

MARTON, F., (1981). Phenomenography—describing conceptions of the world around us. *Instructional Science,* 10, 177–200.

MAYER, R.E. and SIMS, V.K. (1994). For whom is a picture worth a thousand words? Extensions of the dual-coding theory of multimedia learning. *Journal of Educational Psychology,* 46, 389–401.

MCCLURE, J.R., SONAK, B. and SUEN, H.K. (1999). Concept map assessment of classroom learning: reliability, validity, and logistical practicality. *Journal of Research in Science Teaching,* 36, 475–492.

MCDERMOTT, L.C. and REDISH, E.F. (1999). Resource Letter PER-1: Physics Education Research. *American Journal of Physics,* 67, 755–767.

MINTZES, J., WANDERSEE, J. and NOVAK, J.D. (eds.) (2000). *Assessing Science Understanding: A Human Constructivist View* (San Diego, CA: Academic Press).

NOVAK, J.D. (1990). Concept mapping: a useful tool for science education. *Journal of Research in Science Teaching,* 10, 923–949.

NOVAK, J.D. and GOWIN, D.R. (1984). *Learning How to Learn* (New York: Cambridge Press).

PAIVIO, A. (1986). *Mental Representations: A Dual Coding Approach* (Oxford: Oxford University Press).

PEARSALL, N.R., SKIPPER, J.E.J. and MINTZES, J.J. (1997). Knowledge restructuring in the life sciences: a longitudinal study of conceptual change in biology. *Science Education,* 81, 193–215.

PFUNDT, H. and DUIT, R. (1994). *Bibliography: Students' Alternative Frameworks and Science Education*, 4th edn (Kiel: IPN).

QMAP 0.4 (for MS Windows). *Quantum Science Across Disciplines* [http://qsad.bu.edu/Installers/QMap04Inst/install.htm].

RUIZ-PRIMO, M.A. and SHAVELSON, R.J. (1996). Problems and issues in the use of concept maps in science assessment. *Journal of Research in Science Teaching,* 33, 569–600.

RUIZ-PRIMO, M.A., SCHULTZ, S.E., LI, M. and SHAVELSON, R.J. (2001). Comparison of the reliability and validity of scores from two concept-mapping techniques. *Journal of Research in Science Teaching,* 38, 260–278.

SHAKA, F.L. and BITNER, B.L. (1996). Construction and validation of a rubric for scoring concept maps. *AETS Conference Proceedings,* ERIC Document Reproduction Service No. ED. 398060.

TROWBRIDGE, J.E. and WANDERSEE, J.H. (1994). Identifying critical junctures in learning in a college course on evolution. *Journal of Research in Science Teaching,* 31, 459–473.

VAN ZELE, E., LENAERTS, J. and WIEME, W. (2001). Evaluation of a new instructional model and learning ware for an introductory quantum physics course. In H. Evans, A. Moller Andersen and H. Sorensen (eds.) *Bridging Research Methodology and Research Aims, The 5th European Science Education Summerschool,* Gilleleje, Denmark, 6–13 September 2000 (Copenhagen: Danish University of Education), 326–331.

WALLACE, J.D. and MINTZES, J.J. (1990). The concept map as a research tool: exploring conceptual change in biology. *Journal of Research in Science Teaching,* 27, 1033–1052.

WANDERSEE, J.H., MINTZES, J.J. and NOVAK, J.D. (1994). Research on alternative conceptions in science. In D.L. Gabel (ed.) *Handbook of Research on Science Teaching and Learning* (New York: Macmillan), 77–210.

WHITE, R.T. and GUNSTONE, R.F (1992). *Probing Understanding* (New York: Falmer Press).

WILSON, J.M. (1993). The predictive validity of concept mapping: relationships to measures of achievement. *Proceedings of the Third International Seminar on Misconceptions & Educational Strategies in Science and Mathematics* (Ithaca, NY: Misconceptions Trust).