Exploring Formative Feedback on Textual Assignments with the Help of Automatically Created Visual Representations

Open Universiteit
www.ou.nl

Exploring Formative Feedback on Textual Assignments with the Help of Automatically Created Visual Representations

Adriana J. Berlanga, Peter van Rosmalen, Henny P.A. Boshuizen & Peter B. Sloep

Centre for Learning Sciences and Technologies (CELSTEC)

Open University of the Netherlands, P.O. Box 2960, 6401 DL Heerlen, The Netherlands

{adriana.berlanga, peter.vanrosmalen, els.boshuizen, peter.sloep}@ou.nl

Abstract

Learners, particularly lifelong learners, often find it difficult to determine the scope of their expertise. Formative feedback could help them do so. To use this feedback productively, it is essential to then suggest to them the remedial actions they need to overcome the gaps in their knowledge. This paper presents the design considerations of a support tool that aims at providing formative feedback on textual assignments. It does so by facilitating comparisons between learner's input texts and group input texts with respect to the intended learning outcomes. Using language technologies, the tool automatically extracts the concepts and relations of input texts; it then creates visual representations that can be put side by side to identify conceptual overlaps and missing concepts. The paper first introduces the theoretical underpinnings of the tool – specifically those concerning expertise development, knowledge creation and assessment of knowledge. It then draws up design considerations and clarifies how the tool should work. Next, it discusses the results of an initial study in which word clouds and concept maps have been applied to generate graphical visual representations. These help learners identify overlapping and missing core concepts, both in individual texts and in a compiled group text. Finally, the paper provides conclusions and directions for future work.

Keywords: *Learning, formative feedback, expertise development, language technology, concept map, word-cloud*

1. Introduction

When developing their expertise, it is imperative that learners have a good grasp of their level of understanding. And they need formative feedback on how their knowledge development progresses. Providing this kind of feedback requires that the learners be positioned with regard to the learning objectives, that knowledge gaps be identified and that remedial actions be suggested. Learners cannot carry out these tasks entirely by themselves, and for tutors (or teaching assistants as they are called in U.S. universities) these tasks are time consuming. This is particularly so in lifelong learning contexts in which learners may have different learning backgrounds and goals, and may follow quite different learning paths (Sloep et al., 2011).

Van Rosmalen et al. (2008) reported that stakeholders in lifelong learning identify four types of key student support activities; as it happens, these also easily lead to staff work overload: (i) assessment of student contributions, (ii) answering questions of students, (ii) monitoring and assessment of study progress, and (iv) group support. So, typically, tutors will provide students with *summative* feedback on their learning at the end of a module. However, since tutor time is limited, the kind and frequency of *formative* feedback students receive along the way, if they receive it at all, fully depends on what little additional time still is available, and often therefore on the enthusiasm of the tutor alone.

In informal learning situations as well as in some distance learning contexts, learners have a high risk to feel isolated as they are not part of a cohort of peers with whom they can learn and interact, or from whom they can receive formative feedback. This may be problematic, since individual success in learning activities depends on the extent to which learners perceive themselves as participants of a group or community (Rovai, 2002; Dawson, 2008). For lack of tutors, however, in such settings, information about peers and how they are

doing will be the only means of guidance that learners have access to: they thus need to rely on their peers to discuss and understand the study domain.

Formative feedback is the information that is communicated to the learner with the aim of modifying his or her thinking in order to improve learning (Shute, 2008). It is also useful for tutors, to guide learners, identify outliers and, if appropriate, reconceive the learning design. According to Shute (2008) formative feedback can reduce learners' uncertainty on how well they are performing, can promote learning, and can provide useful information for correcting misconceptions. Particularly for novices or struggling learners, formative feedback can reduce their cognitive load by helping them construct problem-solution schemas and by redirecting their attention to cognitive processes (Sweller, Van Merrienboer, & Paas, 1998). Formative feedback strategies help learners move forward in their conceptual development, highlight weaknesses and strengths, allow learners to reflect on their own work, activate learners as the owners of their own learning as well as activate them as 'instructional resources' for other learners (Underwood & Tregidgo, 2006; Black & Wiliam 2009).

Lack of regular formative feedback could imply that learners might not be aware of the limitations of their expertise and might not be able to address the issues this raises. It is our ambition to support (lifelong) learners by providing them with formative feedback. For this, we intend to use a tool that relies on language technologies and, therefore, demands minimal human intervention. More specifically, we offer learners a tool that provides feedback regarding their textual assignments, set-side-by side with the intended learning outcomes and the textual assignments of their peers.

This tool has rests on three (theoretical) underpinnings. The first one is based on research on expertise development (Boshuizen & Schmidt, 1992; 2008). It has been shown that

learners' conceptual development is closely reflected in the textual utterances they express in the course of the evolution of their domain knowledge. As a result, the concepts used and the relations expressed by novices and experts change systematically over time in a fashion that reflects their increasing level of expertise. The second underpinning is the knowledge construction process. Learners develop their expertise by combining cognitive and social processes (such as interacting with their peers or colleagues) to build their individual knowledge 'store' (Stahl, 2006). Learners need, therefore, information about their individual conceptual development, and supposedly benefit from information on how their peers are conceptualizing the same topic.

The third underpinning refers to language technologies. These can automatically extract the concepts and relations that are mentioned in a text, enabling the identification of a learner's progress against the backgrounds of the intended learning outcomes and a group of peers or colleagues. Language technologies analyze (raw) texts and extract concepts with related meanings, thereby enabling the generation of reference models. At a later stage, these models can be contrasted with each other to obtain meaningful information on, for example, conceptual development.

This paper is about the design considerations of the formative feedback tool envisioned by us. We present these, first, by discussing in some detail the theoretical underpinnings and design requirements of the tool. We then continue with the results of a first study that was meant to explore how the reference models can be (semi-)automatically generated, using tools for concept maps and word cloud construction. Third, we present how the proposed tool, and particularly the information it will provide to lifelong learners, has been perceived in an initial study. We wrap up our discussion with some conclusions and thoughts about future work.

2. Theoretical background

Expertise development

Research on expertise (e.g., Chi, Feltovich, & Glaser, 1981; Arts, Gijselaers, & Boshuizen, 2006) has shown that experts and novices differ in their problem-solving skills, knowledge use, information processing, and in the organization of their knowledge structures. Experts have elaborated, well-structured and well-organized mental frameworks, that they activate to interpret information and problems, and to create a suitable solution (Boshuizen & Schmidt, 1992). In contrast, novices do not easily activate their mental frameworks, which are, furthermore, less accurate, incomplete, ill organized and ill structured. Expertise development findings - in domains such as law (Nievelstein, Van Gog, Boshuizen, & Prins, 2008), management (Arts et al., 2006), and medicine (Van de Wiel, Schmidt, & Boshuizen, 2000) - have shown that knowledge becomes more hierarchically structured with increasing expertise; in contrast, novices' knowledge appears to be highly fragmented and its concepts loosely connected.

Boshuizen and colleagues (Boshuizen, Bromme, & Gruber, 2004; Boshuizen & Schmidt, 1992) have divided the learning process in expertise development into three levels: knowledge accretion, knowledge encapsulation and formation of scripts. The levels correspond to the capabilities of novices, intermediates and experts, respectively (see Table 1. For a detailed description see Boshuizen et al., 2004). Novices structure knowledge in networks, and take small steps of reasoning. They rely on knowledge networks only, which are relatively poor, require much information and, as a result, must be reasoned step-by-step. Intermediate students make 'big steps': they increase the step size by gathering a multitude of detailed concepts 'encapsulated' in one higher-order concept. Finally, knowledge experts use experience-based structures, called illness scripts. When they deal

with a case they activate entire, ready-made scripts, no small steps between them are taken anymore. This activation depends on information about conditions, faults, and consequences (Boshuizen & Schmidt, 2008).

Table 1. Expertise level, knowledge structure, learning and reasoning process (Boshuizen, Bromme, & Gruber, 2004; Nievelstein, 2009)

Expertise Level	Knowledge structure	Learning	Reasoning process
Novice	Networks (incomplete, and loosely linked)	Knowledge accretion, integration and validation	Step by step process
Intermediate	Networks (tightly linked and integrated)	Encapsulation	Big steps (but still one at the time)
Expert	Illness scripts	Illness script for formation	Groups of steps activated as a whole
	Memory traces of previous cases	Instantiated scripts	

Appropriate instructional strategies need to take into account learners' knowledge and the gaps therein (Ertmer & Newby, 1993; Jonassen, McAleese, & Duffy, 1993). Instruction intends to fill those gaps, and the question is how learners can be supported through formative feedback. The feedback should be based on their current, context-bound level of expertise, and it should take into consideration the learning context and the knowledge building process the learner is taking part in.

Knowledge creation theories

Knowledge creation theories focus on how individuals and groups develop knowledge that is new to them. Such theories explicitly deny that knowledge is transmitted untouched and unchanged from one – knowledgeable – person to another – non-knowledgeable - person. Indeed, they emphasize that knowledge is constructed in a dialectical and social process; that is, the construction process not only relies on explicitly stated knowledge and information but rather on a much bigger reservoir of tacit knowledge, that is accessible

through querying and discussion (Boisot, 1998; Brown & Duguid, 1991; Nonaka, Nayana, & Konno, 2000).

In the area of collaborative knowledge creation, Stahl (2006), following a social epistemological perspective (Brown & Duguid, 1991; Lave & Wenger, 1991), claims that learning comprises mutual construction of individual and social knowledge. In his view individuals generate personal beliefs from their own perspectives, but they do so on the basis of socio-cultural knowledge, shared language and external representations. These beliefs become knowledge through social interaction, communication, discussion, clarification and negotiation.
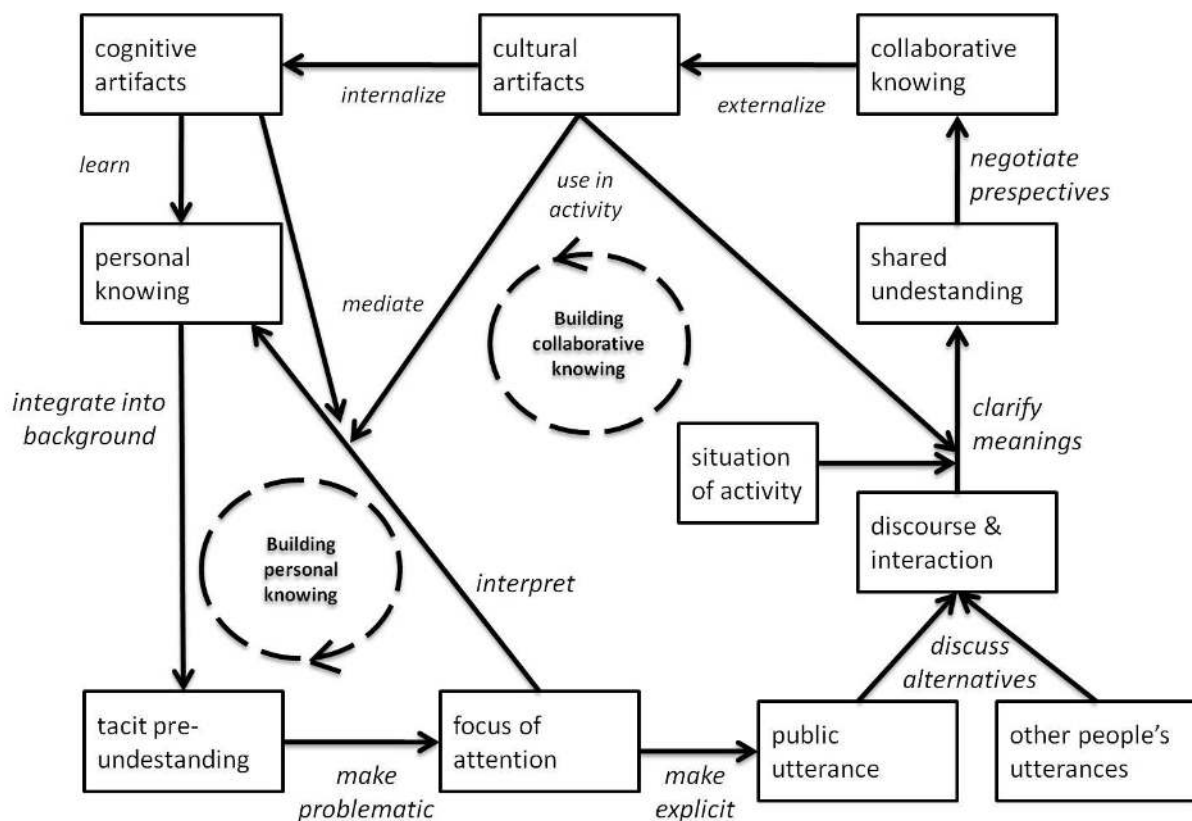
Figure 1. Stahl's cycle of knowledge building

Figure 1 shows Stahl's cycle of knowledge building. The diagram depicts how the personal and the collaborative knowing building cycles interact. The lower left corner shows the cycle

7

of personal understanding, which might start with a tacit pre-understanding influenced by personal knowing. This understanding may change if someone tries to solve conflicts or fill gaps by reinterpreting her meaning structures. If she cannot resolve the problematic character of her personal understanding herself, then she needs to enter into an explicitly social process and create new meaning collaboratively. If she decides to do so, she needs to express herself in public statements and by doing so she enters into the cycle of social knowledge building. This cycle –shown in the right part of the diagram (Figure 1)– represents how the social interaction process with people and with their shared culture influences the individual's understanding. This process consists of an interchange of arguments that provides rationales for different points of view, which eventually may converge on a shared understanding.

Knowledge building, furthermore, relies on cultural and cognitive artefacts. Cultural artefacts represent a shared understanding among a group of people. As soon as they are internalized, they become cognitive artefacts and re-enter the cycle of building personal knowing. Through this learners integrate the new knowledge into their previous knowledge and can eventually formulate new problems or questions. Of particular importance is that Stahl's diagram also includes activities for providing feedback in both cycles, to support personal understanding and social knowledge building.

The provision of formative feedback in this knowledge building cycle requires that learners be informed on how they are developing their expertise, not only from a cognitive point of view but also from a social point of view. From a cognitive point of view, learners need information so they can compare their own understanding of the topic with the intended learning outcomes. From a social point of view, they need cues so they can compare and discuss the differences in how they conceptualize a topic with respect to how others (e.g.,

peers, colleagues, members of a group, etc) do so. This then might provide learners with a way to reflect on how they are i) conceptualizing the topic, ii) rethinking the reasons behind differences and similarities between their conceptualization and that from others, iii) finding new concepts and links, iv) reflecting on what they know, and v) using those reflections to construct knowledge bases (Jonassen & Carr, 2004). This could also help learners to find peers that could support them clarify their ideas.

The upshot of this discussion is that learners' conceptual development needs to be identified. And for that it is imperative that the process of measuring conceptual development is charted out.

Measuring conceptual development

Goldsmith, Johnson, & Acton (1991) propose a structural approach to assess an individual's existing knowledge of a particular domain at a specific point in time. The approach consists of analyzing how she organizes the concepts of such a domain, and involves three steps: knowledge elicitation, knowledge representation, and evaluation of her knowledge representation.

Knowledge elicitation techniques measure the learner's understanding of the relationships between the various concepts in a set (Jonassen, Beissner, & Yacci, 1993). Methods that support this activity include, for instance, card sorting, concept maps, or essay questions. The second step is to define some representation of the knowledge elicited that reflects the underlying organization of the data (Goldsmith et al., 1991). Advanced statistical methods (e.g., cluster analysis, pathfinder nets) are used to identify the structural framework that underlies the set of domain concepts.

The third step is to evaluate the individual's knowledge representation relative to some standard (e.g., reference model). Normally, researchers follow one of three approaches

(Goldsmith et al., 1991): qualitative assessment of derived representations; quantifying the similarities between a student representation and a derived structure of the content of the domain; or comparing the cognitive structures of experts and novices.

3. Design considerations of a Tool that Provides Formative Feedback on Learner's Text Assignments

The theoretical underpinnings discussed in the previous section lead us to conclude that any design for providing formative feedback on learner's text materials should take into account that:

1. Some learner's level of expertise is reflected in the way she uses, links and relates concepts when she express her knowledge

2. Learners develop their knowledge in a building process, which encompasses cognitive and social perspectives

3. The assessment of a learner's conceptual development requires elicitation, representation and evaluation of her knowledge.

The assessment could be done through supporting comparisons between learner's textual utterances and reference models. These reference models have been defined as either one of two kinds (Berlanga et al., 2009a):

- *Predefined reference model*, built out of intended learning outcomes described in, for instance, course materials, tutor notes, curriculum, etc.

- *Group reference model*, built out of the concepts and the relations a group of people (e.g., peers, participants, co-workers, etc.) used the most.

The formative feedback tool, using language technologies, processes text materials (such as learner essays, learning material, etc.), extracts the concepts they contain and identifies

how these concepts are related to each other. The tool should display these concepts as a graphic (e.g., reference concept map) and as a list of concepts. The tool should also allow one to compare different text inputs.

As an example of how the tool would work, let us assume a lifelong learner, Efke, who is enrolled in a course about Learning Networks for professionalization in the European Open University. By way of assignment, she needs to write an essay that answers some questions the tutor has posted.

Efke writes her first draft essay. She then uses her text as input for the formative feedback tool and generates a concept map (see Figure 2). The concept map indicates what the most important concepts mentioned in her input text are as well as how these concepts are related.
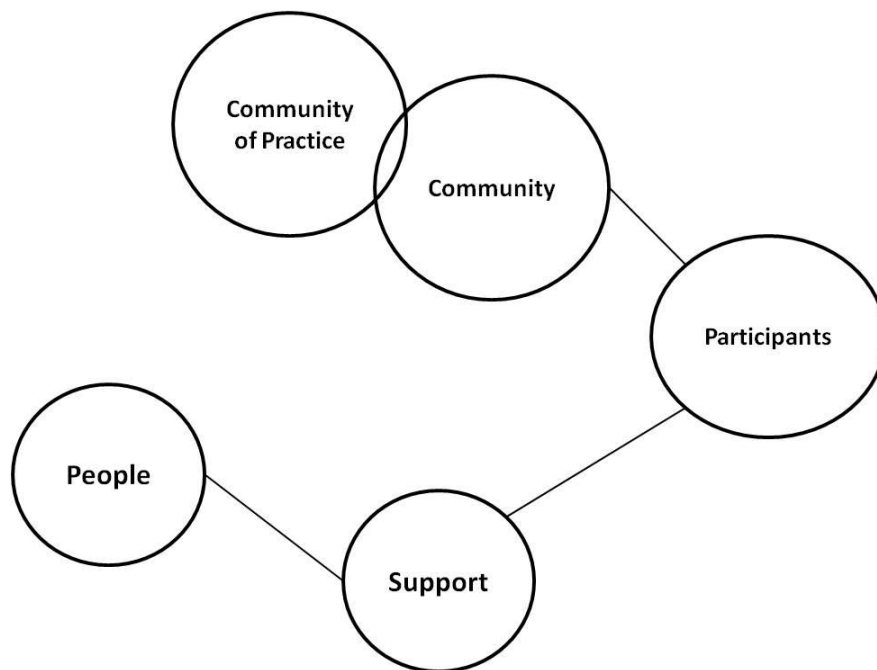
Figure 2. Example: learner concept map

The tool also makes the different concept maps created by Efke's tutor available: a predefined concept map, which was created from the learning materials of the course, a group concept map, which was created by aggregating in a single text the essays of a group

11

of students (those that already followed the course and submitted their essays). Again using the tool, Efke first selects her map and the predefined map, and creates a combined single map (see Figure 3). It then becomes evident to her what concepts she failed to mention in her essay, concepts that are present in the learning materials but not in her draft. In the visual representation they have differently shaped outlines (e.g., learners, services, content, resources). Concepts that are covered in both maps have the same shading, so Efke can easily see the overlapping concepts.
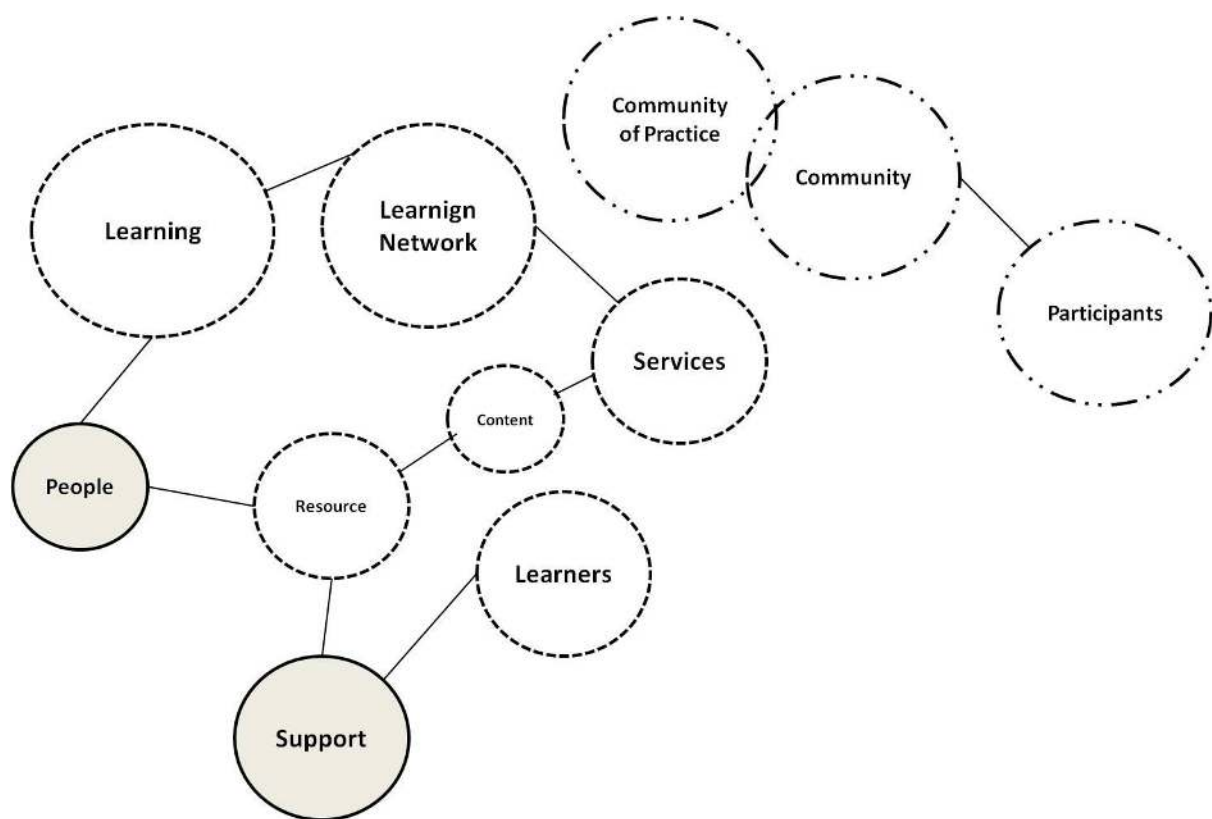
Figure 3. Example: learner concept map vs. predefined reference map

She observes that there are too many concepts her essay does not mention, so she decides to see whether she mentions the same concepts as her peers did. She compares her essay to the group concept map and discovers that her peers cover concepts she did not think of, thus identifying areas on which she may need help from others. With this in mind, she goes back to the study materials, contacts one of her peers to ask some questions, and writes a

new version of her essay. Then she uses the tool again to check her new version of the assignment.

In short, using the service learners receive prompt formative feedback about their learning texts. Through the feedback they may identify, for instance, shared concepts and relations between them, concepts they did not (yet) use or missed. In this way they will have an indication that they are on the right track or that they should update or discuss their work. Moreover, learners could monitor their own expertise development as the service would also provide comparisons of the learner's knowledge evidences submitted previously. For tutors, the service will provide a means to monitor the current progress of learners on a topic and to identify outliers; it will also allow tutors to take proactive actions to improve learners' conceptualization of the topic.

4. Language technology tools for eliciting and representing learner's knowledge

We have been exploring if (and how) it is possible to (semi-) automatically analyse input texts in order to generate the above mentioned reference models (predefined reference model, group model) and provide a comparison between them. The generation of such models requires to process text-based materials (i.e., learner and tutor materials), to identify the concepts they contain and their mutual relations, and to provide a visual representation that reflects how these concepts are connected. Learners can thus compare their visual representations with those of their peers and with a predefined reference model; the later supposedly represents the intended learning outcomes.

To make such a comparison different tools and methods are available, ranging from simple tools for term extraction and word cloud generation; via tools that assist in creating concept maps –e.g., (Clariana, Koul, & Salehi, 2006; Shute, Jeong, Spector, Seel, & Johnson, 2009); to

more complex technologies such as Latent Semantic Analysis (LSA) (Clariana & Wallace, 2007; Landauer, McNamara, Dennis, & Kintsch, 2007) that generate a mathematical model where both the domain knowledge and the knowledge of the learner can be projected.

Tools for concept map generation

The concept maps method provides a suitable and appropriate measure of the learner's representation of the subject matter structure. Concept maps, furthermore, are one of the most common ways of representing cognitive structures. Research demonstrates the appropriateness of concept maps in eliciting knowledge (Nesbit & Adesope, 2006) and their superiority for the evaluation of learners of different ages (Jonassen et al., 1997; Novak, 1998).

There are already a number of tools for the automatic construction and support of concept maps: Knowledge Network Organizing Tool (KNOT, PFNET) (Clariana et al., 2006); Surface, Matching and Deep Structure (SMD) (Ifenhaler & Seel, 2005); Model Inspection Trace of Concepts and Relations (MITOCAR) (Pirnay-Dummer, 2006); Dynamic Evaluation of Enhanced Problem Solving (DEEP) (Spector & Koszalka, 2004); jMap (Jeong, 2008); ProDaX (Oberholzer, Egloff, Ryf, & Läge, 2008), and Leximancer (Smith & Humphreys 2006) (see Table 2). These tools have some common characteristics: (a) they can (semi-) automatically construct concept maps from a text; (b) they use a kind of distance matrices; (c) they propose a quantitative analysis of the maps; and (d) most of them are concerned with conceptual development of learners (for a comparison see Berlanga et al., 2009b).

Table 2. Existing tools for construction of concept maps (adapted from (Shute, Jeong, Spector, Seel, & Johnson, 2009)

	Data Collection	Analysis	Comparison
KNOT	Concept pairs/Propositions	Quantitative Analysis	Direct comparison of networks with some statistical results

SMD	Concept map or natural language	Quantitative analysis is calculated using tools	Unlimited comparison
MITOCAR	Natural language	Quantitative analysis included multiple calculations using tools	Paired comparisons for semantic and structural model distance measure
DEEP	Annotated causal maps	Quantitative/qualitative analysis is done mostly by hand	Unlimited comparisons, showing details relative to concepts
jMap	Concept maps, causal maps, or belief networks	Quantitative analysis (calculated using tools)	Superimposes maps of individual (n=1) and group of learners (n = 2+) over a specified target map
ProDax	Association Data, Cross-Tables, Two-Way Two-Mode Data, Coordinates	Non-Metric Multidimensional Scaling/Cluster-Analysis	Comparison of maps based on Procrustean Transformation
Leximancer	Concept maps	Content analysis and relational analysis (proximity, cognitive mapping)	Imposes tags in a single map over user-defined tags (names, concepts, files, etc.)

Most existing approaches in this area set up the reference model in advance (Clariana & Wallace, 2007), or include a phase of sampling and negotiating amongst participants and peers aimed at detecting which concepts the reference model should include (Shute et al., 2009). Our work, however, strives to go beyond these approaches and derive the reference models post hoc, with minimal human intervention.

We have conducted, therefore, proof-of-concept tests (Berlanga et al., 2009a) to demonstrate the use of tools for diagnosing conceptual development and the automatic generation of reference models. In the first of these a predefined reference model was constructed (based on course and tutor materials) and compared with the concept map of a student. For this purpose, a combination of Leximancer and Pathfinder was used. The second proof of concept, in which only Leximancer was used, explored the generation of a predefined reference model that identifies the concepts and relations mentioned by participants in a small-scale pilot. Results showed that, since it contained mainly high-level

concepts, the predefined reference model seemed to provide little information that can be used to generate a formative feedback report. This could be so because the theoretical information is written by experts and thus encapsulates the knowledge through high-level concepts only. These will be too abstract to understand for students in an early stage of their studies (see Section 2).

The group reference model seems to solve this issue. It presents the set of concepts that is used by most people at a specific point of time and seems to better capture the level of abstraction and relations between concepts. To tutors, this model provides information on which individual learners can identify key concepts and which cannot, making it easier for tutors to identify outliers. However, it will require a better appreciation of the learner's knowledge representation – by contextualizing both the learner's knowledge and the situation in which the knowledge will be applied – and requires mechanisms to keep the model updated.

5. Research questions

Taking into account theoretical underpinnings (Section 2), the initial outline of the tool (Section 3), and the exploration of existing tools (Section 4), we designed and prepared a functional mock-up of a tool we felt would best provide the formative feedback sought. The mock-up was used to explore the following questions:

A. Using simple, standard term extraction and word cloud tools only, is it feasible to build a visual, 'individual' representation of a text on a selected topic that according to a learner (who is the writer of the text) covers the *core* concepts of that text?

B. Similarly, is it feasible to build a visual, 'group' representation of set of selected texts on a selected topic that according to a learner, being the writer of one of the texts, covers the *core* concepts of the consolidated text?

C. According to learners, are the resulting representations useful for comparing and contrasting the individual versus the group perspective on the selected topic?

Method and data

To explore these questions, we carried out an study with researchers, in the role of learners, of the Learning Networks research programme (Sloep, 2008) of the Open University of the Netherlands. The participants were selected from one research programme only to assure that they possessed largely the same background knowledge and used roughly the same concepts. We selected six participants, together a representative sample of different levels of experience in this programme, i.e. one PhD student, one Post-doc, and two assistants, one associate and one full professor. The participants (5 male, 1 female; aged from 29 – 59) had been active in the programme for at least 3 years.

We asked each of the participants to provide us with one of their existing articles, one that they deemed illustrative of their research in Learning Networks. Each of the articles (average size around 5000 words) and the combined article, i.e., the text of all articles copied together, was used to create two different visualizations, a word cloud and a concept map. A word cloud representation does not fulfil all requirements of our tool, i.e., it does not show the relations between concepts. However, given that it is easier to create and to understand, we were interested if it would suffice for our purpose of providing feedback. Wordle (www.wordle.net) was used to create the word-cloud– and Leximancer (www.leximancer.com) to create the concept map, in the following way. The text of each

article was manually 'pre-processed' in particular to expand domain specific abbreviations (e.g., LN was expanded with Learning Network) and to collate multiple word concepts into single words (e.g., Learning Network by "LearningNetwork"). Next, the resulting text was processed by Wordle and Leximancer. This included common operations such as stopping (for a definition see http://en.wikipedia.org/wiki/Stop_word), i.e., filter out numbers and common English words (e.g., words such as 'the', 'such', 'but') and specify the maximum numbers of words to be presented, in our case 25. 25 was chosen in order to focus on the core concepts (Ruiz-Primo, Shavelson, & Schultz, 1997) and to avoid making the visualisation of the concepts and therewith also its understanding unduly complex. Finally, the resulting visualization was fine tuned manually; i.e., any remaining not meaningful or redundant words or abbreviations in the top 25 were removed. In a similar way, two group visualizations were created, each based on the text of the six articles lumped together. Figure 4 shows the visual representations: Figures 4.a, 4.b show word cloud (Wordle) representations: individual (4.a) and group (4.b); Figures 4.c, 4.d show (Leximancer) representations: individual (4.c) and group (4.d). To collect the opinions of our participant researchers, we created a questionnaire and asked them:

1. To what extent the two kinds of *individual* representations covered their work. For this purpose the questionnaire asked participants to indicate (A) if they recognized their articles in the concept map, and (B) if their individual concept map represented a fair coverage of the core concepts mentioned in their article. The same questions (C, D) were asked regarding the word cloud representation.

2. To what extent the two kinds of *group* representations covered the topic of research 'Learning Networks'. The questionnaire asked participants, therefore, to indicate (E) if they recognized the Learning Network research in the group concept map, and (F)

if the group concept map represented a fair coverage of the concepts encompassed by the Learning Networks programme. Likewise, the respondents were asked the same questions (G, H) regarding the word-clouds.

3. To what extent a comparison between their *individual* representation and the *group* representation gave rise to similar and missing concepts. To this end the questionnaire asked participants to indicate (I) if they could identify similar and missing concepts when comparing the group concept map against their individual concept map, and again the same question for the word cloud representation (J). Please note that "missing" is not absolute. First, the comparison between the representations shows both concepts not discussed and concepts only minimally discussed. Second, the comparison points to concepts that might deserve more attention. "Might" since each paper emphasises its specific attention points. So there may be good reasons for why some concepts end up missing in some cases.

4. Finally, which of the two visualizations they preferred, the concept map or the word cloud representation.

The data collected consisted of a combination of ratings on a five-point Likert scale (1: strongly disagree; 2: disagree"; 3, neither agree nor disagree; 4: agree; 5: strongly agree). The questionnaire also contained open questions that allowed the participants to explain their choices. Finally, it contained two questions through which the participants were asked to select the three most significant overlapping concepts and the three most significant concepts of the group visualization that they missed in their own text.

For the study, the participants did not need to prepare themselves other than sending us their article. Given their experience, it was assumed that, even though they may not have read all articles, they were familiar with all key concepts of the programme. This should

resemble a regular course situation in which not all students necessarily read the same background literature but do have the same shared course background. After preparation of the visualisations, the participants were invited to a session in which each of them received an individual set of visualisations, consisting of the word cloud and the concept map of their article and of the group article. They received the assignment to carefully study the visualisations and fill in the above mentioned questionnaire. Participants were asked to work alone; if necessary, support was provided by one of the authors.

Figure 4a. Word cloud (Wordle): individual representation
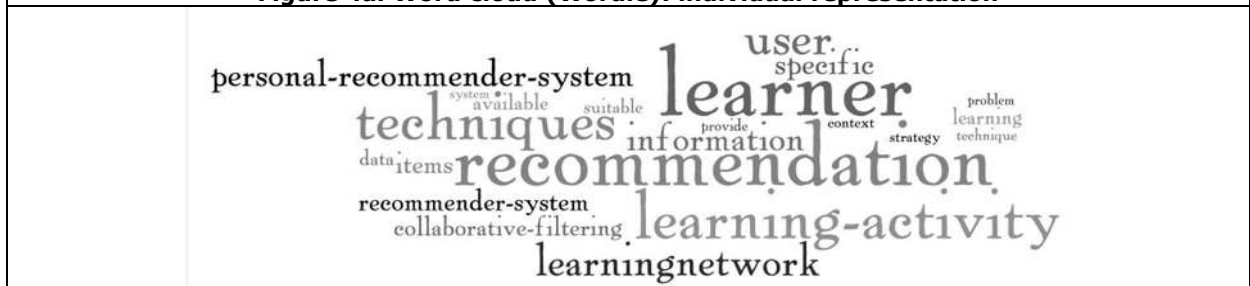
Figure 4b. Word cloud (Wordle):group right representation.
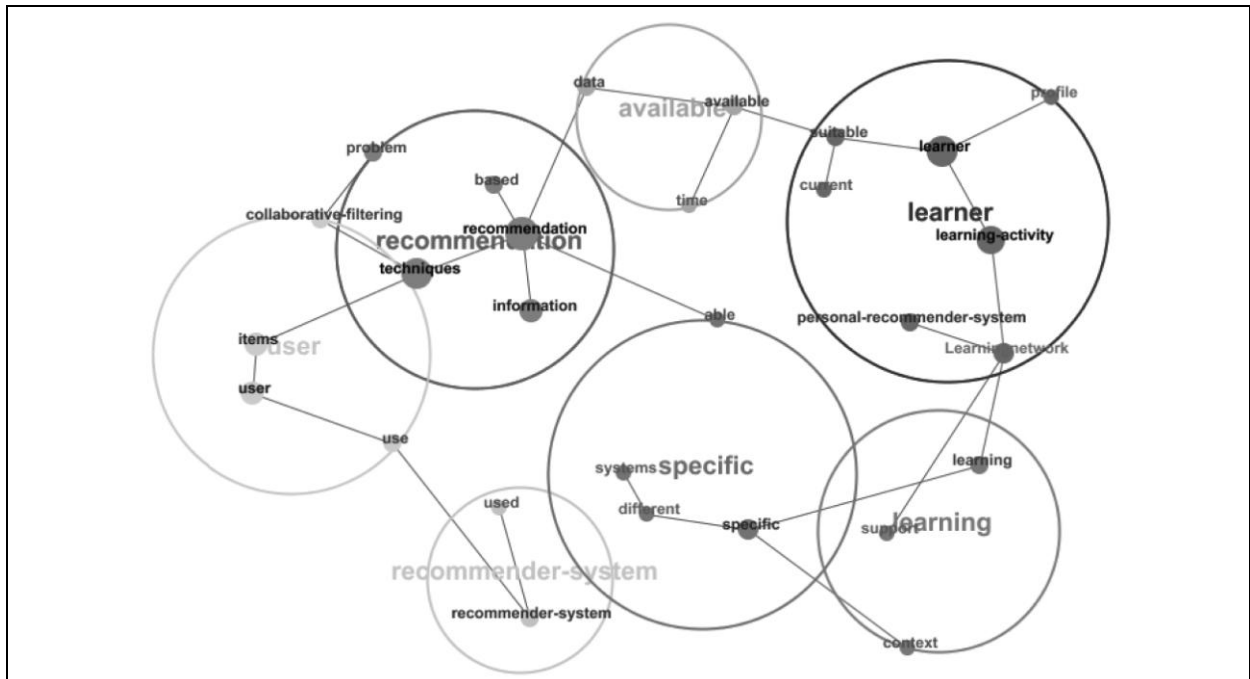
Figure 4c. Concept map (Leximancer): individual (left) representation
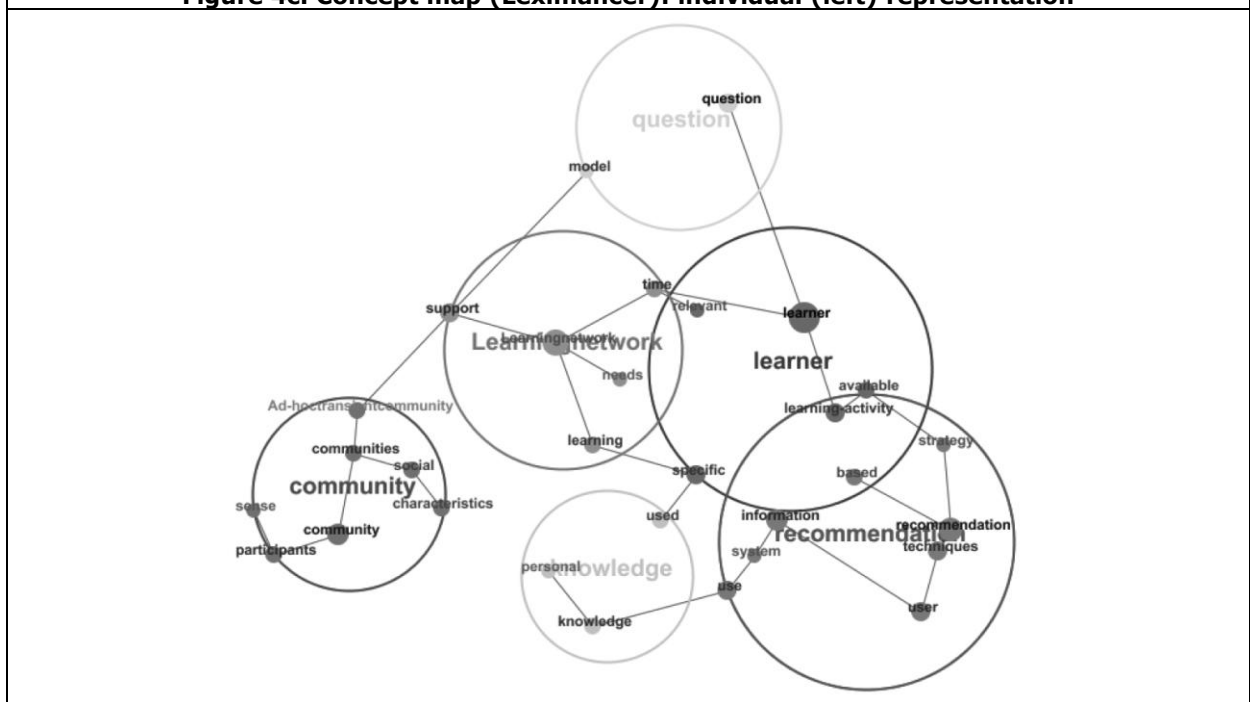
Figure 4d. Concept map (Leximancer): group representation

Results

The quality of the individual representation

Respondents rated the word cloud representation slightly better than the concept map

representation (Table 3). Most respondents (5 out of 6) "agree" or "strongly agree" that in

the word cloud representation they could recognize their article (avg. 4), and also most of them (5 out of 6) "agree" or "strongly agree" that the word cloud represents a fair coverage or the concepts their article was about (avg. 4.3).

Table 3. Perceived Quality of the individual representations (1=strongly disagree – 5 strongly agree)

	Concept map		*Word cloud*	
	recognition	**coverage**	**recognition**	**coverage**
P1	2	4	3	5
P2	4	5	5	5
P3	2	3	4	5
P4	4	4	4	4
P5	4	4	4	3
P6	2	4	4	4
Avg.	3	4	4	4,3

The quality of the group representation

Results show (see Table 4) that most respondents recognize the Learning Network research in the group representations, and that the representations show a fair coverage of concepts. There is a small difference between the respondent's opinion regarding the group concept map and the group word-cloud. In this case the concept map was perceived as the better representation.

Table 4. Quality of the group representations (1=strongly disagree – 5 strongly agree)

	Concept map		*Word cloud*	
	recognition	**coverage**	**recognition**	**coverage**
P1	4	4	4	5
P2	4	4	4	4
P3	4	5	4	3
P4	2	4	2	4
P5	4	3	2	3
P6	3	4	4	4
Avg.	3,5	4	3,3	3,8

The value of an individual and a group representation to compare and contrast individual vs.

group perspectives

Results show (see Table 5) that most respondents (5 out of 6) "agree" or "strongly agree" on the idea that they could identify similar and missing concepts while contrasting the individual and group concept map representations. Contrariwise, only half of the respondents indicated that they could identify similar and missing concepts while contrasting the individual and group word-clouds.

Table 5. Value of comparisons between group representation and individual representation (1=strongly disagree – 5 strongly agree)

	Concept map	*Word cloud*
P1	5	5
P2	4	4
P3	2	2
P4	4	2
P5	4	2
P6	4	4
Avg.	3,8	3,2

Results show also that respondents could fairly easily identify the overlapping and missing concepts between their individual and group representations. As expected, the overlapping concepts in the six articles are almost identical. The missing concepts show a spread depending on the article. Table 6 shows the concepts the respondents identified with help of the concept maps (the word cloud case shows similar results). Interestingly, while only one paper covered recommendation as its central topic, three other papers discussed it at a less prominent level. Therefore (on average) in the group model the concept "recommendation" was identified. This fits, at the same time, with the importance of the concept recommendation in the Learning Networks research programme.

Table 6. Individual and group concept maps: overlapping and missing concepts

	Concept Map					
	Overlapping			*Missing (or only covered minimally)*		
	1st	2nd	3rd	1st	2nd	3rd
P1	learning network	learner	community	recommendation	question	Support
P2	learning network	learner	question	recommendation	adhoc	community

					transient community	
P3	personal needs	knowledge	community	learning network	question	recommendation
P4	learning network	learner	recommendation	community	question	model
P5	learning network	learner	community	recommendation	knowledge	needs
P6	learning network	knowledge	community	recommendation	question	Learner

Preferences

The results show that there is no clear preference for either a concept map or word cloud representation. Half of the respondents prefer the concept map, and half the word cloud representation. Reasons for preferring concept map representations alluded to being able to see the links between the concepts and to see how concepts are grouped in sub-concepts; reasons for preferring word-clouds were simplicity and the lack of information regarding the semantic value of concept map links.

Interestingly, some respondents mentioned that they would find it useful to have a tool as the one we propose to help them to check if their articles cover the core concepts they are supposed to cover and to check inconsistencies. Also, one respondent mentioned that performing comparisons between a particular article (which could be for instance a "benchmark" of a particular topic) and a group model gives a good impression of whether the article and the group model cover the same ground.

6. Discussion

The results of this initial study lend support to affirming question A: it is possible to build an individual representation of a text that, according to its author, adequately covers the core concepts of the text. All participants indicated that they could recognize their article in the individual word cloud representation. Only half of the participants indicated that their individual map representation depicted their article fairly well. However, their main objection was that a concept map does not represent the argument structure of the article.

One of the participants admits, however, that '*perhaps it is asking too much of a concept map to also represent an argument structure*'.

Likewise, results support to affirmatively answer question B: it is possible to build a group representation of a set of selected texts on a selected topic that, according to the authors of the texts, covers the core concepts of the consolidated text. Results also show that participants find the quality of the group concept map representation, in terms of coverage and recognisability of concepts, slightly better than the word cloud representation. Some participants argue that the word cloud contains concepts that are not connected, and that they are either too general or too detailed.

Overall, participants perceive the group representation as useful input to compare and contrast the individual representation with the group perspective on the selected topic. In general, therefore, question C can be answered affirmatively. However, half of the participants do not perceive the word cloud representation as a proper way of identifying similar and missing concepts when comparing the group word cloud with their individual word cloud. Some mention that it is much harder to pinpoint differences in the word cloud representation. To compare, it may be necessary to adapt the word cloud visualization (probably also the concept map) so it accentuates the overlaps and differences.

Answers also show that visual representations, concept map and word cloud are useful to identify similar and missing concepts. Results also show that participants are positive on the idea of comparing their individual representations with the group perspective on the selected topic. They perceived, furthermore, both concept map and word cloud representations as useful inputs. Interestingly, there is no consensus on which representation is best. From this initial study only, we cannot derive conclusions about the reasons for this.

In general, the concept maps seem to provide more meaningful information than do the word cloud representations. First, concept maps cluster terms in themes and make the links between them explicit. Second, although the concept maps are more difficult to interpret than the word-clouds, they seem to provide a view on the expertise of the writer of the article.

From a technical point of view, results of the initial study show that it is possible to build individual and group visual representations that cover the concepts in the input text. Particularly, the use of word cloud tools is quite easy; with a minimum amount of work one may generate visual representations and compare different text inputs.

7. Conclusions and future work

This paper presented the theoretical considerations, design and requirements of a tool that can provide formative feedback on learner's textual assignments. The tool has been based upon the premise that enabling the comparison of different reference models, which should be generated automatically (or semi-automatically but at least with minimal human intervention), will allow learners to understand their position in a knowledge domain.

Results of a proof of concept of how existing tools can be used to provide learners with information about their conceptual development, indicate that relatively simple and widely accessible tools for term extraction and word cloud creation already provide useful information. However, more advanced tools are needed to identify the relations between the concepts. Results also show that the text input is essential to ensure the tool is working properly. It is evident that the generation of the group model has to rely on text inputs that discuss the same topic at a similar level of depth. If this is not the case the concepts mentioned in the texts will be so different that the group model will be meaningless. One

other issue to be further explored is how much expertise is required to be able to use this kind of feedback. Novices, for instance, may not yet share a sufficient number of concepts to do a useful comparison. The creation of the group model, finally, should be done in such a way that the aggregation in a single text of the text input from a group of people is done automatically, or at least, does not require much work.

The original idea was to consider all text inputs from the participants and generate a single concept map, in fact the group model. However, the results suggest that considering all text inputs of all participants in one and the same concept map will not provide a meaningful way of comparison. The input of each participant has the same impact on the group model, while participants might have widely differing levels of expertise. Therefore, a possible solution would be to ask an expert to pre-score input texts and see how this evaluation could be considered to generate the group model.

Furthermore, not only the text quality and length of the input, but also the topics the text mentions should be in line with what is expected. This constitutes a serious risk, if learners do not express in writing all their knowledge about the topic, the tool will not provide meaningful formative feedback. Of course, there is knowledge that cannot be expressed in writing, but learners, or users of the tool, should be aware of the fact that the tool does require appropriate input text of high quality to provide useful feedback. A possible solution would be that learners use the reference model as example and outline for writing their text assignment. In this way they would have an idea of what is expected.

Participants were positive about the idea of using models (individual and group visual representations) as means of comparing and contrasting what they had written with what a group of people had written, but also as a means of identifying what one knows.

Undoubtedly, more research is needed to establish how learners would benefit most from comparing their conceptual development with the proposed models (predefined or group model): whether it is a good strategy for learners to see comparisons with both models or, whether, depending on their level of expertise, comparisons with different models will be made available. The type of reference model to be used may depend on the level of a learner's development. The group reference model would most likely be of use for an individual learner at a novice level, as at this stage it would correspond to her Zone of Proximal Development (Vygotsky, 1978). As expertise develops, the group reference models may still be appropriate, depending on the development stage of the group as a whole, but predefined reference models may be more suitable to advanced learners.

8. Acknowledgment

9. References

Arts, A. J., Gijselaers, W. H., & Boshuizen, H. (2006). Understanding managerial problem-solving, knowledge use and information processing: Investigating stages from school to the workplace. *Contemporary Educational Psychology, 31*(3), 387-410.

Berlanga, A. J., Brouns, F., Van Rosmalen, P., Rajagopal, K., Kalz, M., & Stoyanov, S. (2009a). *Making Use of Language Technologies to Provide Formative Feedback.* Paper presented at the AIED 2009 Workshops Proceedings Volume 10. Natural Language Processing in Support of Learning: Metrics, Feedback and Connectivity. Proceedings of the 14th International Conference in Artificial Intelligence in Education, Workshops Proceedings Brighton, United Kingdom.

Berlanga, A. J., Kalz, M., Stoyanov, S., Van Rosmalen, P., Smithies, A., & Braidman, I. (2009b). *Using Language Technologies to Diagnose Learner's Conceptual Development.* Paper presented at the The 9th IEEE International Conference on Advanced Learning Technologies, Riga, Latvia.

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability, 21*(1), 5-31.

Boisot, M. (1998). *Knowledge assets: securing competitive advantage in the information economy*. New York, NY: Oxford University Press.

Boshuizen, H. P. A., Bromme, R., & Gruber, H. (2004). On the long way from novice to expert and how travelling changes the traveller. In H. P. A. Boshuizen, R. Bromme & H. Gruber (Eds.), *Professional learning: Gaps and transitions on the way from novice to expert* (pp. 3-8). Dordrecht: Kluwer.

Boshuizen, H. P. A., & Schmidt, H. G. (1992). On the role of biomedical knowledge in clinical reasoning by experts, intermediates and novices. *Cognitive Science, 16*, 153-184.

Boshuizen, H. P. A., & Schmidt, H. G. (2008). The development of clinical reasoning expertise: Implications for teaching. In J. Higgs, M. Jones, S. Loftus & N. Christensen (Eds.), *Clinical reasoning in the health professions* (Vol. 3rd rev. ed.). Oxford: Butterworth-Heinemann/Elsevier.

Brown, J. S., & Duguid, P. (1991). Organizational Learning and Communities-of-Practice: Toward a Unified View of Working, Learning, and Innovation. *Organization Science, 2*(1), 40-57.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5*(2), 121-152.

Clariana, R., Koul, R., & Salehi, R. (2006). The criterion-related validity of a computer-based approach for scoring concept maps. *International Journal of Instructional Media, 33*(3), 317-325.

Clariana, R., & Wallace, P. (2007). A Computer-Based Approach for Deriving and Measuring Individual and Team Knowledge Structure from Essay Questions. *Journal of Educational Computing Research, 37*(3), 211 - 227.

Dawson, S. (2008). A study of the relationship between student social networks and sense of community. *Educational Technology & Society, 11*(3), 224-238.

Ertmer, P. A., & Newby, T. J. (1993). Behaviorism, cognitivism, constructivism: Comparing critical features from an instructional design perspective. *Performance Improvement Quarterly, 6*(4), 50-70.

Goldsmith, T. E., Johnson, P. J., & Acton, W. H. (1991). Assessing structural knowledge. *Journal of Educational Psychology, 83*, 88–96.

Ifenhaler, D., & Seel, N. M. (2005). The measurement of change: learning-dependent progression of mental models. *Technology, Instruction, Cognition and Learning, 2*(4), 317-336.

Jeong, A. (2008). jMap v. 104. from http://dev22448-01.sp01.fsu.edu/ExcelTools/jmap/

Jonassen, D. H., Beissner, K., & Yacci, M. (1993). *Structural knowledge: techniques for representing, conveying, and acquiring structural knowledge*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Jonassen, D. H., McAleese, T. M. R., & Duffy, T. M. (1993). Manifesto for a constructivist approach to technology in higher education. In T. M. Duffy, J. Lowyck & D. H. Jonassen (Eds.), *Designing environments for constructive learning: Implications for instructional design and the use of technology*. Heidelburg, FRG: Springer-Verlag.

Jonassen, D. H., Reeves, T., Hong, N., Harvey, D., & Peter, K. (1997). Concept mapping as cognitive learning and assessment tools. *Journal of interactive learning research, 8*(3/4), 289-308.

Jonassen, D.H., & Carr C.S. (2000). Mindtools: affording multiple knowledge representations for learning. In S.P. Lajoie, (Ed.), *Computers as cognitive tools: No more walls*, Vol. 2, (pp.165-196). Mahwah, NJ: Lawrence Erlbaum Associates.

Landauer, T. K., McNamara, D. S., Dennis, S., & Kintsch, W. (2007). *Handbook of Latent Semantic Analysis*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, MA: Cambridge University Press.

Nesbit, J., & Adesope, O. (2006). Learning with concept and knowledge maps: A meta-analysis. *Review of Educational Research, 76*, 413-448.

Nievelstein, F., Van Gog, T., Boshuizen, H. P. A., & Prins, F. J. (2008). Expertise-related differences in ontological and conceptual knowledge development in the legal domain. *European Journal of Cognitive Psychology, 20*(6), 1043-1064.

Nonaka, I., Nayana, R., & Konno, N. (2000). SECI, Ba and Leadership: a unified model of dynamic knowledge creation. *Long Range Planning, 30*(30), 5-34.

Novak, J. D. (1998). *Learning, creating and using knowledge: concept maps as facilitative tools in schools and corporations*. Mahwah, NJ: Erlbaum

Oberholzer, R., Egloff, S., Ryf, S., & Läge, D. (2008). Prodax. Datenauswertung im Bereich der Skalierung. Benutzerhandbuch.

Pirnay-Dummer, P. (2006). *Expertise and model building. MITOCAR.* Unpublished doctoral dissertation. Filburg: University of Freiburg.

Rovai, A. P. (2002). Building a Sense of Community at a Distance. *International Review of Research in Open and Distance Learning, 3*, [Electronic Version] http://www.irrodl.org/index.php/irrodl/article/viewArticle/79/152.

Ruiz-Primo, M. A., Shavelson, R. J., & Schultz, S. E. (1997). On The Validity Of Concept Map-Base Assessment Interpretations: An Experiment Testing The Assumption Of Hierarchical Concept Maps *Science, CSE Technical Report* 455, from Retrieved 14 February 2011 at http://www.stanford.edu/dept/SUSE/SEAL/Reports_Papers/TECH455.pdf

Shute, V. J. (2008). Focus on Formative Feedback. *Review of Educational Research, 78*(1), 153-189.

Shute, V. J., Jeong, A. C., Spector, J. M., Seel, N. M., & Johnson, T. E. (2009). Model-Based Methods for Assessment, Learning, and Instruction: Innovative Educational Technology at Florida State University. In M. Orey (Ed.), *2009 Yearbook Educational Media and Technology*: Greenwood Publishing Group.

Sloep, P. B. (2008). Netwerken voor lerende professionals; hoe leren in netwerken kan bijdragen aan een leven lang leren. Inaugural address. Open Universiteit Nederland. Available at http://dspace.ou.nl/handle/1820/1559.

Sloep, P. B., Boon, J., Cornu, B., Klebl, M., Lefrère, P., Naeve, A., Scott, P., & Tinoca, L. (2011). A European Research Agenda for Lifelong Learning. *International Journal of Technology Enhanced Learning(IJTEL), 2* (3), 204-228.

Smith, A. E., & Humphreys, M. S. (2006). Evaluation of unsupervised semantic mapping of natural language with Leximancer concept mapping. *Behavior Research Methods, 38*(2), 262-279.

Spector, J. M., & Koszalka, T. A. (2004). *The DEEP methodology for assessing learning in complex domains*: Syracuse University.

Stahl, G. (2006). *Group Cognition: Computer Support for Building Collaborative Knowledge*. Cambridge: MIT Press.

Sweller, J., Van Merrienboer, J., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review 10*, 251–296.

Underwood, J. S., & Tregidgo, A.P. (2006). Improving student writing through effective feedback: Best practices and recommendations. *Journal of Teaching Writing*, *22*(2), 73-97.

Van de Wiel, M. W. J., Schmidt, H. G., & Boshuizen, H. P. A. (2000). Knowledge Restructuring in Expertise Development: Evidence From Pathophysiological Representations of Clinical Cases by Students and Physicians. *European Journal of Cognitive Psychology, 12*, 323-355.

Van Rosmalen, P., Sloep, P., Kester, L., Brouns, F., De Croock, M., Pannekeet, K., & Koper, R. (2008). A learner support model based on peer tutor selection. *Journal of Computer Assisted Learning, 24*(1), 74-86.

Vygotsky, L. S. (1978). *Mind in Society*. Cambridge, MA: Harvard University Press.