A Computer Supported Method to Reveal and Assess Personal Professional

Theories in Vocational Education

Antoine C. M. van den Bogaart[1], Richel Bilderbeek[2], Harmen Schaap[3],

Hans G.K. Hummel[14] & Paul A. Kirschner[4]

[1] NHL University of Applied Science, The Netherlands

[2] University of Groningen, The Netherlands

[3] Radboud University, The Netherlands

[4] Open University of the Netherlands

Author Note

Correspondence concerning this article should be addressed to Antoine C. M. van den

Bogaart, NHL University of Applied Science, Rengerslaan 10, 8917 DD Leeuwarden,

The Netherlands, E-mail: a.c.m.van.den.bogaart@nhl.nl

Abstract

This article introduces a dedicated, computer supported method to construct and formatively assess open, annotated concept maps of Personal Professional Theories (PPTs). These theories are internalised, personal bodies of formal and practical knowledge, values, norms and convictions that professionals use as a reference to interpret and acquire knowledge, and to direct their behaviour, and which vocational students are expected to develop. Monitoring the development of PPTs and assessing their quality is difficult as they are, essentially, mental schemes. Traditional methods, such as semi-structured interviews and concept mapping, are either too labour intensive to be used in an educational setting or are not able to reveal their full quality. The study presents a new method which is valid, reliable and easy to use in education and which reveals the quality in a way that is comparable to or better than interviews.

Introduction

One of the defining characteristics of a profession is its body of knowledge, skills, and attitudes which have to be internalised by professionals and integrated into competences (Baartman & De Bruijn, 2011). In the course of training and praxis, professionals construct and internalise their own personal, profession-specific body of knowledge which they use as a reference framework for acquiring and interpreting new knowledge and for directing their professional behaviour (Schaap, de Bruijn, van der Schaaf, & Kirschner, 2009). This framework not only consists of different kinds of knowledge and skills acquired in formal training (Wilensky, 1964), along with work process and practical knowledge (Verloop, 2009), but also of personal convictions, and the norms and values of the profession. To clarify and emphasise the personal nature of such a framework, a new construct was introduced: Personal Professional Theory (PPT). According to Schaap, et al. (2009, pp. 490-491): "PPTs: (1) consist of declarative and procedural knowledge; (2) are stored in the long-term memory; (3) refer to compiled knowledge; (4) are built upon different types of knowledge, such as formal theories, work process knowledge, practical knowledge and shared knowledge, collective norms and values of a vocational community; and (5) can be specified and applied in different situations". A PPT is, thus, a structured set of elements such as concepts, examples, procedures, heuristics, facts, values and beliefs (Pajares, 1992), along with relations between these elements that might be hierarchical, causal or conditional for example (Toulmin, 1958).

Vocational education lays the foundations of PPTs. In higher vocational education in the Netherlands – the context of this study – this takes place in four-year higher education programmes that include several periods of practical work/internships: often one short, introductory period in the first year and two

extended periods in the third and fourth years. The majority of students enter higher vocational education after completing senior secondary education (i.e. as opposed to having completed vocational secondary education). For them, these periods are their first experience with the praxis of their chosen profession. Though research indicates that integrating school and workplace learning is important for the development of PPTs (Griffiths & Guile, 2003), there is no real consensus as to how best to support students in this (Biemans, Nieuwenhuis, Poell, Mulder, & Wesselink, 2004). One of the major issues discussed is how to evaluate their developing professional knowledge as a whole (Baartman, Bastiaens, Kirschner, & van der Vleuten, 2007). When this evaluation is formative, it provides students with feedback, thereby supporting their further development (Gibbs & Simpson, 2004), and gives teachers a good picture of the students' development so that they can better personalise their teaching. Therefore, it is important to have a valid and reliable method of revealing and assessing PPTs formatively, which is easy to use in practice.

In the literature, many methods for knowledge elicitation have been described, tested and validated (Clariana, 2010; Cooke, 1994; Hoffman & Lintern, 2006). Three frequently encountered methods of revealing and assessing PPTs (Schaap, De Bruijn, Van der Schaaf, Baartman, & Kirschner 2011) are based on: (1) observations and interviews (Scott, Clayton, & Gibson, 1991); (2) process-tracing (Ericsson & Simon, 1984); or (3) elicitation of concepts (McGraw & Harbison-Briggs, 1989). However, these are all too time consuming to be usable in education. This problem led to the aim of this study, namely, to provide a valid and reliable method of formatively revealing and assessing PPTs which is usable in educational practice, easy to learn and to handle, meets the demands of the students and teachers, and is pleasant to work with (Lund, 1997).

Theoretical background

In this section different methods for revealing and assessing PPTs are compared. To make such a comparison, the nature of the content and quality of PPTs has first to be clarified.

The content and quality of a PPT

A person's PPT can often be deduced from how they speak about their profession. However, because of its personal nature and the often implicit or encapsulated character of the knowledge in a PPT (Schmidt & Boshuizen, 1993), it may be difficult to decide if an expression uttered by a professional, such as an English teacher, saying "I read British newspapers", is actually referring to a content element of his/her PPT or expressing his/her opinion on the quality of British newspapers. Schaap et al. (2011) developed a method to overcome this difficulty. Referring to Eraut (1994, 2000) they distinguished *six domains of knowledge*: (1) knowledge of the vocational field, such as general issues and trends in the profession; (2) organisational knowledge, including general work processes, information systems, management and cultural aspects (Rauner, 2007); (3) professional knowledge of the organisation one is working in; (4) target group knowledge, e.g. in education, pupils and parents; (5) technical-instrumental knowledge on how to perform adequately and accountably (Kelchterman & Hamilton, 2004); and (6) knowledge relevant for personal development, such as one's strengths and weaknesses. Together these domains cover the content of personal professional knowledge. An element which could be classified as belonging to at least one of these domains "counts" as member of a PPT. This interpretation is also followed in this study.

The quality of a PPT can be analysed from different viewpoints. Because the aim of this study is to provide a general method usable in vocational education, a definition of quality is needed that is independent of a specific profession. Schaap et al. operationalised such a definition in terms of concreteness, specificity, complexity and richness (Messick, 1990; De Jong & Fergusson-Hessler, 1996). In the present study, these components are referred to as the four "qualities" of a PPT. It is important to state that the meaning of the term "quality" is to be understood as a synonym for "aspect" or "feature" as in philosophy and not as a value judgement (Locke, 1942). For example, a PPT that is more complex than another PPT is not per se a better framework of reference.

The qualities concreteness, specificity and complexity are features of both the individual elements of a PPT and a PPT as a whole; the more concrete, specific or complex the elements that a PPT contains, the more concrete, specific or complex it is considered to be. An element is *concrete* if it refers to knowledge, instruments or means that can be used in action (Gherardi, 2000). Concreteness of knowledge is an aspect of the connection between the owner of this knowledge and the knowledge itself; if the owner knows how to act according to this knowledge, it is concrete (Wilensky, 1991). An element is *specific* if it refers to knowledge applied in a profession or group of related professions which distinguishes that profession from other professions (Smagorinsky & Smith, 1992). For example, a procedure may be general in nature (e.g. plan-do-check-act), but it requires different professional knowledge and tools, depending on the profession in which it is carried out. An element is *complex* if it refers to aspects of professional actions such as consequences, influences, order, causes, exceptions and conditions (Trochim & Cabrera, 2005). Complexity can be expressed by means of relations with other elements (Bakker &

Derry, 2011) or by means of examples and clarifications. *Richness*, the fourth quality, is a global quality of a PPT, which expresses the "species diversity" of domains of knowledge in the PPT (Teigland & Wasko, 2004). Species diversity is an ecological concept and is determined by two factors: the number of species in an ecosystem and the evenness of the distribution of these species (Hurlbert, 1971).

Methods to reveal PPTs and assess their qualities

Each of the three methods discussed has been shown to be valid – via measures of interrater reliability – with researchers triangulating the results, and the three together give a good picture of a PPT (Buitink, 1998; Schaap et al., 2011). However, they all have certain individual and shared drawbacks. Schaap et al. analysed the data generated by the three methods and found that while they were highly interrelated, the semi-structured interviews scored somewhat higher than concept maps for concreteness, specificity and complexity, while concept maps scored higher than traces. With regard to richness, the scores of the concept maps were the highest (Schaap et al. 2011). The reason for these differences is obvious. In semi-structured interviews, the interviewer has the opportunity to ask for clarification and examples. This method should lead to a more concrete, specific and complex representation of a PPT than if the respondent has to generate their thoughts alone. On the other hand, concept maps are well suited to giving an overview of a theory (Cañas et al., 2003). Their spatial character prompts the respondent to fill gaps that might exist, and the completeness of such an overview is what is measured by the quality *richness*.

Finally, log files are a form of process-tracing and are not intended for extracting representations of PPTs – in contrast to semi-structured interviews and concept maps. This is possibly why these two methods scored higher than the log files.

Analysing coded items in documents such as log files and interviews is a specialist and labour intensive task (Robson, 2011), which makes them not really usable in educational practice. Analysing concept maps looks more promising for application in practice, because a concept map can be scanned in one view, and two or more can easily be compared, giving a rough estimation of richness and complexity, while deeper analysis can be supported by means of a computer (Cañas et al., 2003; Liu, 2013). However, then the problem of the lower scores for concreteness, specificity and complexity has to be solved. How this was done is explained in the next section.

A Computer Supported Procedure to Construct and Assess Concept Maps of PPTs

Design considerations: Heightening levels of concreteness, specificity and complexity

To solve the problem of the lower scores for concreteness, specificity and complexity, a tool needs to be designed that expands the traditional concept mapping approach by making it possible to construct an *annotated* concept map with rich descriptions of the nodes and links (Pirnay-Dummer, Ifenthaler, & Spector, 2010). This allows for the adding of examples and clarifications to concepts (Huyts, De Bruijn, & Schaap, 2011) and to relations between concepts (Buitink, 1998), providing concrete, specific and complex information about the concepts.

The problem here is that the annotated concept map is more complex and time consuming to analyse. However, as argued earlier, it might be possible to automate this in a computer program.

Design considerations: Transforming the qualities in programmable algorithms

To operationalise the levels of concreteness, specificity and complexity of a concept map, it has to be divided into clusters, which means into parts containing a concept with all its examples, clarifications and relations to other concepts. To determine the level of concreteness/specificity of a cluster, firstly, each of its examples and clarifications has to be judged as to whether it is concrete/specific or not. Following Schaap et al. (2011), the level of concreteness or specificity of a cluster is then defined as 0 (less), 1 (average) or 2 (more) if it contains less than two, two or three, or more than three examples or clarifications that are concrete or specific. Subsequently, the concreteness and specificity of a concept map of a PPT can be determined as the arithmetical means over all its clusters.

To determine the complexity of a concept map, again the complexity of each of its clusters has to be determined. In the same way as the first two qualities, the complexity of a cluster can be expressed as 0, 1 or 2, depending on the sum of the number of complex examples and clarifications, and the number of relations in the cluster. Again, the level of complexity of a PPT is the arithmetical mean over its clusters.

Finally, the richness of a concept map can be determined by connecting all the examples and clarifications of all its concepts to a domain of knowledge. To express it as a number, first the number of different domains, connected with one or more examples or clarifications has to be counted. To this number another number is added. This second number expresses the level of evenness of the distribution over the different domains. If all elements are evenly distributed over the six domains of knowledge, each domain covers 16.7%, so one could say that a domain contributes to the evenness of the distribution if the number of its elements lies between 12% and

25% of the total number of elements of the PPT (Huyts et al., 2011). This leads to the following formula for the level of richness:

$$\left(\left(\frac{1}{12} \times a + \frac{1}{12} \times b\right) \times 100\right)$$

where a is the number of domains of knowledge, with a positive relative frequency, and b is the number of domains, with a relative frequency between 12% and 25%.

The computer program supporting the construction and analysis of concept maps

An open source computer program – Brainweaver (Bilderbeek, 2013) – was developed and used here. It consists of two modules: one to support the construction of a concept map of a PPT and the other to support the formative assessment of this map.

The construction module gives the opportunity to insert associations raised by the focus statement (Trochim & Cabrera, 2005), as a first step in the construction of a concept map (Figure 1).

INSERT FIGURE 1 ABOUT HERE

In this screenshot, the focus question is "*As a student teacher of English as a foreign language, I have knowledge of: ...*". The inserted associations can be labelled as concepts or examples of concepts. When the association is completed, Brainweaver transfers the concepts to the concept map screen, positioned in a random way around the focus question. Subsequently, it is possible to build the concept map by positioning, adding, changing or deleting concepts or examples, drawing relations and describing them. Most of these functionalities are mimicked from known software in order to construct concept maps such as CmapTools (Cañas et al., 2003). The functionality of inserting examples and explanations into concepts and relations is

new. The result is a concept map of a PPT as shown in Figure 2. In this screen, examples are hidden behind the concepts. By activating a concept, these examples pop up as shown with "Myself as a person".

INSERT FIGURE 2 ABOUT HERE

The assessment module supports the determination of: (1) the distribution of examples of concepts over the domains of knowledge, and (2) four qualities. For the distribution, each example and clarification of each concept has to be classified in one of the domains. This classification is supported as shown in Figure 3. Given a classification, Brainweaver calculates the distribution in the form of the relative frequencies of these classifications, expressed as percentages, and the quality *richness* by means of the formula presented earlier.

INSERT FIGURE 3 ABOUT HERE

To determine the qualities concreteness, specificity and complexity, each example, clarification and relation of each cluster has to be evaluated as either having or not having these qualities. This is shown in Figure 4 for the cluster of the concept "My own professional development". "X" stands for complexity, "C" for concreteness and "S" for specificity.

INSERT FIGURE 4 ABOUT HERE

Brainweaver suggests for each concept and cluster the level of concreteness, specificity and complexity expressed as 0, 1 or 2, based on the rules explained in this section. These levels have to be confirmed to finalise them. After all the concepts and clusters have been evaluated, Brainweaver calculates the concreteness, specificity and complexity of the concept map as the arithmetic means of the scores of the concepts, expressed as percentages.

Research questions

The aims of this study and the development of the construction and assessment procedure supported by Brainweaver, led to two exploratory research questions, namely: (1) Is this procedure usable in educational practice? and (2) How does this procedure compare to semi-structured interviews in relation to the values of the four qualities of PPTs? The following section describes the set-up of a quasi-experimental study to answer both research questions.

Method

The study was conducted during the course of two academic years: 2012/2013 and 2013/2014. Each year, participating students were randomly divided into two groups. One group was asked to produce concept maps of their PPTs supported by the computer program and the other group received a semi-structured interview. Both concept maps and interviews were assessed by two assessors and in the assessment of the concept maps the computer program was used again. By questioning the users of the computer program and comparing the interviews with the concept maps, both research questions were answered.

Participants

All 47 participants in the study were teachers in training at an institution for higher

vocational education in the Netherlands and had a foreign language as their subject

(15 male, 32 female, M_{age} = 22.5 years, *SD = 2.1*). The students were in the third year

of a four-year course and had just finished a period of eight weeks' teaching practice

in secondary education, four days a week. In the first year of our study, all

participants were enrolled for teaching English as a foreign language. In the second

year of our study, as the number of student teachers of English available was smaller,

it was decided to invite the small numbers of student teachers of German and French

to participate. This was possible because their curricula had the same structure as the

English teaching curriculum and half of the courses were attended by all students.

After an introduction of one hour on PPTs and the purpose and design of the

study, all students in the first year of the study agreed to participate. When it came to

the second year, eight students could not participate because of internships abroad. In

2012/2013 this resulted in the participation of 27 student teachers of English (10

male, 17 female, M_{age} = 22.8 years, SD = 2.3), and in 2013/2014 in the participation

of 12 student teachers of English, 6 of German and 2 of French (5 male, 15 female,

M_{age} = 22.1 years, SD = 1.7).

Materials: Examples, protocols and questionnaires

To assess the concept maps and interviews two lists of examples were developed with

the help of two experienced teacher trainers. The first list referred to examples of the

six domains of knowledge relevant to the professional domain, such as the symptoms

of dyslexia as an element of the domain target group and the different roles of

teachers as belonging to the domain technical-instrumental processes. This list was

used to classify examples and clarifications in concept maps and utterances in

interviews as belonging to one of the domains. The second list contained relevant

examples and non-examples of concrete, specific, complex elements, such as "You have to speak English with the pupils" (not concrete) and "In each lesson you may only converse with the pupils in English" (concrete).

The protocol for the semi-structured interviews on PPTs was developed in consultation with a staff member of the institution. After an introduction in which the interviewer had to check whether the student understood what a PPT was, the only questions permitted were questions like: "Could you please further explain your PPT to me?", and "Could you give an example of this or clarify it any further?".

The questionnaire for students on the usability of the student module contained four 5-point Likert-scale questions on different aspects of usability (Lund, 1997), followed by an open invitation to add comment. To get additional information on the way Brainweaver was used and on the content validity of the procedure, six 5-point Likert-scale questions regarding the students' view on the importance of applying the procedure during their education were added, again followed by an open question on this subject. To gather information on the usability of the assessor module, a protocol for a semi-structured group interview with the assessors was developed with similar questions.

Analysis

The analysis of the concept maps was supported by Brainweaver and the two lists of examples. To determine the four qualities from the semi-structured interviews, each interview was transcribed. From each transcription, student utterances about their PPTs were sorted into mutually disjoint sets. Utterances with a common theme, like "You always have to maintain order in the class" and "It is important to arrange the furniture in the classroom before the lesson starts", were put in the same set, in this case within the theme "classroom management". Utterances expressing relations

between two different theme-sets, for instance classroom management and learning, were set aside. The interviews were then analysed, as described in the section on the analysis of concept map, treating the theme-sets as concepts with examples and clarification, and the theme-sets and their relations with other sets as clusters (Kim, 2012).

The four qualities calculated from the concept maps were compared with those calculated from the interviews, using a *t*-test for two independent samples for each quality. Additional information on the content validity of both methods was obtained by comparing the distributions of the domains of knowledge. Content validity of both methods should reveal low and high frequencies for the same domains.

Procedure

In the first year of the study, a group of 14 students (6 male, 8 female, $M_{age} = 22.8$ years, $SD = 2.1$) were asked to construct concept maps of their PPTs supported by Brainweaver in sessions of 75 minutes: 15 minutes' instruction in the use of the computer program, followed by 60 minutes' construction. All sessions took place over a two-day period. Two experienced teacher trainers assessed the concept maps. To test for interrater reliability, two sets of three concept maps were assessed by both assessors. The results of the first set were discussed to sharpen their interpretation, leading to acceptable or almost perfect values of Cohen's Kappa in the second set on the classification of domains ($K = .76$, $N = 153$), and on the assessment of concreteness ($K = .91$, $N = 165$), specificity ($K = .93$, $N = 165$) and complexity ($K = .85$, $N = 165$). The remaining concept maps were distributed between the two assessors. The usability of the student module of Brainweaver was tested in two ways immediately after the construction of the concept maps. The ten students from the group of 14 who produced a concept map were given the questionnaire on usability,

and after analysing the data from the questionnaires, the conclusions were discussed in a group interview with the remaining four students. To complete the exploration of usability, both assessors were interviewed after they had completed the assessment of all the concept maps. The interview took an hour and there were two steps to answering the questions in the interview: (1) giving a personal answer; and (2) discussing these answers to try to reach agreement.

The remaining group of 13 students (4 male, 9 female, M_{age} = 22.8 years, SD = 2.5) received a semi-structured interview on their PPTs. The interviews were held over a three-day period. The researcher and the staff member did all the interviews, the first together and the others separately. The interviews took between 20 and 40 minutes (M = 27.5, SD = 4.0) and were all audiotaped and transcribed. The attribution of the utterances to theme-sets was checked by the two teacher trainers who assessed the concept maps. The teacher trainers assessed each utterance as: (1) belonging to one of the domains (K = .81, N = 98); and (2) according to its concreteness (K = .80, N = 99), specificity (K = .82, N = 99) and complexity (K = .72, N = 99). Based on this assessment, the four qualities were evaluated and the results were given to the assessors for confirmation.

Subsequently, in the second year of the study, 10 students (3 male, 7 female, M_{age} = 21.7 years, SD = 1.5) were asked to produce a concept map and 10 students (2 male, 8 female, M_{age} = 22.5 years, SD = 1.8) were interviewed. Concept maps and interviews were assessed by the same assessors.

Results

The data on usability from questionnaires and interviews

The data on usability from the 10 students who worked with Brainweaver and filled in the questionnaire are displayed in Table 1.

INSERT TABLE 1 ABOUT HERE

Two students commented in the open question that they had wanted a little more time to construct their concept map. Three students commented that they found it difficult to name the relations between concepts. The other remarks related to this question concerned technical suggestions to improve the program. The answers of the students to the questions in the second part of the questionnaire, on the use of the program and on its possible importance for the curriculum, are displayed in Table 2.

INSERT TABLE 2 ABOUT HERE

The answers to the last, open question of the questionnaire supported the results displayed in Table 2. The group interview with the remaining four students on the data from the questionnaires confirmed the results. One of the students remarked that the program helps you to reflect on who you are as a teacher, and the others agreed. Another student said: "The program stimulates to think about my most important issues and the connections between them. This takes time."

The group interview with both assessors on the usability of the assessor module, the assessments and the possible importance of formative assessments of PPTs is reported in Table 3. Agreement on all subjects was reached.

INSERT TABLE 3 ABOUT HERE

The quantitative comparison of concept maps and interviews

The number of clusters in a concept map ($M = 18.2$, $SD = 6.3$) is larger than the average number of groups of expressions referring to a common theme in an interview ($M = 9.9$, $SD = 1.9$). This could reflect the strength of the graphic nature of concept maps as explained earlier. The differences between the number of examples and clarifications in concept maps ($M = 37.2$, $SD = 14.4$) and in interviews ($M = 35.9$, $SD = 11.9$) are small, so the number of examples and clarifications per cluster in concept maps ($M = 2.3$, $SD = 1.0$) is smaller than the number of these in groups with a common theme in interviews ($M = 3.7$, $SD = 1.2$). The number of relations between concept labels in a concept map and its variance ($M = 14$, $SD = 8.2$) are larger than the number of relations between groups of expressions in an interview and its variance ($M = 6.4$, $SD = 1.9$).

Table 4 contains the results on the distribution of the expressions over the domains of knowledge. For each domain, the absolute and relative frequencies taken over all of the concept maps/interviews, and the average number and standard deviations per concept map/interview have been calculated.

INSERT TABLE 4 ABOUT HERE

In Table 5 the results of the assessments on the four qualities are displayed, together with two-tailed *p*-values, calculated from independent-measures *t*-tests for mean differences.

INSERT TABLE 5 ABOUT HERE

Qualitative data on the differences between interviews and concept maps

As the average number of clusters in a concept map is larger than the number of groups with a common theme in an interview, the number of subjects addressed in a concept map is larger than the number of subjects in interviews. However, there is a difference in the modality of the way these subjects are addressed. For instance, concept maps do not show the sense of urgency of a subject in a PPT, while interviews do reveal this. An example of this is the use of English as teaching medium in English as a foreign language lessons. For some students, this is an element of their PPT and it can be found in concept maps as a concept label, "English as teaching medium", or as an example added to a concept such as "pedagogical content knowledge". But in an interview one finds: "I think using English as a teaching medium is very important. I find it disappointing that my coach, who is an experienced teacher, does not think it is possible to use English as a teaching medium on all occasions".

Another difference between concept maps and interviews is the length of examples. For instance, as examples added to the label "learning disabilities", several short lexical items, such as dyslexia and ADHD (attention deficit hyperactivity disorder), can be found in a concept map. In an interview, fewer examples are given, but they are often explained in one or more sentences, like in the student statement: "For children with dyslexia, learning a second language is often very difficult. You have to be very supportive to those children."

Relations in concept maps and between topics in an interview are often expressed in general terms, such as "depends on" and "is influenced by". However, the examples from interviews provided relations between states of affairs that are

sometimes explicated in more specific terms, like in the student statement: "I prefer to have a classroom of my own. Then I'm able to welcome the children when they enter the classroom. This has a calming effect on excited kids."

Conclusion and Discussion

Regarding the first research question, the results show that the computer-supported method to assess PPTs by means of concept maps is usable in educational practice. The questionnaire and interview data show that after some practice, teachers become familiar with the program and are able to analyse a concept map of a PPT in less than half an hour with sufficient interrater reliability. This is a big improvement as just preparing an interview for assessment took at least four hours. Moreover, students and teachers find it useful to include formative assessments of PPTs in the curriculum via the procedure described.

In regard to the second research question, Table 4 shows that both methods produce similar distributions of the domains of knowledge. None of the differences in Table 5 is significant ($p > .05$). The results indicate that concreteness and richness are comparable. The indication that richness using the concept map method, is not significantly better (40.6 versus 41.8), differs from the result reported in Schaap et al. (2011) (65.5 versus 22.0). A possible explanation for this difference is the dominance of technical-instrumental knowledge in the distribution of domains. This could mean that the third-year student teachers were so preoccupied by the technical-instrumental aspects of the profession that there was not much room in their PPTs for other domains. Although not significant, complexity scored better using the concept map method (41.8 versus 32.1). Comparing this with the results Schaap et al. obtained in senior secondary vocational education (1.0 versus 35.5), we have an indication that

the computer program method leads to more complex concept maps for students in higher vocational education. The program stimulates students to relate the main components of their PPT, although they sometimes find it hard to describe these relations. Finally, the scores for specificity were higher in the interviews (56.6 versus 48.3). This is in accordance with the results of Schaap et al. (2011) (54.0 versus 24.5) and can be explained by the encouragement in interviews to explain as much as possible.

It is important to mention that these conclusions are only tentative as the variances are large because of the differences between students. Comparing two groups of students differs from the design used by Schaap et al. (2011), in which the same students produced a concept map and were interviewed. This design was considered not to be suitable for this study because it was expected that concept maps with examples and interviews would influence each other. The large variances also indicate that the assessment procedure discriminates between students. To get more evidence for the power of the computer supported method of eliciting PPTs, as compared with interviews, more experiments need to be done in different professions and with larger number of students. Longitudinal research is necessary to be able to study the development of PPTs and explore the power of the computer supported method for demonstrating development.

The qualitative comparison of concept maps and interviews indicates that both methods have their own strengths: concept maps offer a larger number of different subjects, whereas interviews better reveal students' feelings for different elements of their PPT. This means that in a formative assessment procedure the concept map of a PPT should be discussed between student and assessor, before developmental advice is formulated and given.

Acknowledgements

References

Baartman, L. K. J., Bastiaens, T.J., Kirschner, P. A., & van der Vleuten, C. P. M.

(2007). Evaluation assessment quality in competence-based education: a

qualitative comparison of two frameworks. *Educational Research Review, 2*,

114-129.

Baartman, L. K. J., & de Bruijn, E. (2011). Integrating knowledge, skills and

attitudes: conceptualising learning processes towards vocational competence.

Educational Research Review, 6, 125-134.

Bakker, A., & Derry, J. (2011). Lessons from inferentialism for statistics.

Mathematical Thinking and Learning, 13, 5-26.

Biemans, H., Nieuwenhuis, L., Poell R., Mulder, M., & Wesselink, R. (2004).

Competence-based VET in the Netherlands: background and pitfalls. *Journal of

Vocational Education and Training, 56*(4), 523-538.

Bilderbeek, R. (2013). Brainweaver [C++] (Version 42). Retrieved from

http://www.richelbliderbeek.nl/ProjectBrainweaver.htm.

Buitink, J. (1998). *In-functie opleiden en in-functie leren van aanstaande leraren*

(Unpublished doctoral dissertation). UCLO, Groningen.

Cañas, A. J., Coffey, J. W., Carnot, M. J., Feltovich, P., Hoffman, R. R., Feltovich, J.,

& Novak, J. D. (2003). *A summary of literature pertaining to the use of concept

mapping techniques and technologies for education and performance support.*

Pensacola, FL: The institute for Human and Machine Cognition.

Clariana, R. B. (2010). Multi-decision approaches for eliciting knowledge structure.

In *Computer-based diagnostics and systematic analysis of knowledge (pp. 41-

59).*

Cooke, N. J. (1994). Varieties of knowledge elicitation techniques. *International Journal Human-Computer Studies, 41*, 801-849.

de Jong, T., & Fergusson-Hessler, M. G. M. (1996). Types and qualities of knowledge. *Educational Psychologist, 31*, 105-113.

Eraut, M. (1994). *Developing professional knowledge and competence.* London: Routledge.

Eraut, M. (2009). Transfer of knowledge between education and workplace settings. In H. Daniels, H. Lauder, & J. Porter (Eds.), *Educational policy: a critical perspective*. London: Routledge.

Ericsson, K. A., & Simon, H.A. (1984). *Protocol analysis: verbal reports as data.* Cambridge, MA: MIT press.

Gherardi, S. (2000). Practice-based theorizing on learning and knowing in organizations. *Organisation, 7*, 211-223.

Gibbs, G., & Simpson, C. (2004). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education, 1*, 3-31.

Griffiths, T., & Guile, D. (2003). A connective model of learning: the implications for work process knowledge. *European Educational Research Journal, 2*(1), 56-73.

Hoffman, R. R., & Lintern, G. (2006). Eliciting and representing the knowledge of experts. In Ericsson, K. A., Charness, N., Feltovich, P., & Hofmann, R. (Eds.). *Cambridge handbook of expertise en expert performance* (pp. 203-222) New York: Cambridge University Press.

Hurlbert, S. H. (1971). The nonconcept of species diversity: a critique and alternative parameters. *Ecology, 52*, 577-586.

Huyts, P. M., de Bruijn, E., & Schaap, H. (2011). Revealing personal professional theories. *Quality and Quantity, 45*, 783-800.

Kelchtermans, G., & Hamilton, M. L. (2004). The dialectics of passion and theory: exploring the relationship between self-study and emotion. In J. J. Loughran, M. L. Hamilton, V. Kubler LaBoskey, & T. Russel (Eds.), *International handbook of self-study of teaching and teacher education practice*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Kim, M. G. (2012). Cross-validation study of methods and technologies to assess mental models in a complex problem solving situation. *Computers in Human Behaviour, 28*, 703-717.

Liu, J. (2013). The assessment agent system: design, development, and evaluation. *Educational Technology Research and Development, 61*, 197-215.

Locke, J. (1924). *An essay concerning human understanding.* Abridged and Edited by L. Pringle-Pattison. Oxford: Oxford University Press.

Lund, A. M. (1997). Expert ratings of usability maxims. *Ergonomics in Design: The Quarterly of Human Factors applications, 5*, 15-20.

McGraw, K. L. & Harbison-Briggs, K. (1989). *Knowledge acquisition: principles and guidelines.* Englewood Cliffs, NJ: Prentice Hall.

Messick, S. (1990) *Validity of test interpretation and use* (Research Report No. ETS-RR-90-11). Princeton, NJ: Educational Testing Service.

Pajares, M. F. (1992). Teachers' beliefs and educational research: cleaning up a messy construct. *Review of Educational Research, 62*, 307-332.

Pirnay-Drummer, P., Ifenthaler, D., & Spector, J. M. (2010). Highly integrated model assessment technology and tools. *Educational Technology, Research and Development, 58*, 3-18.

Rauner, F. (2007). Practical knowledge and occupational competence. *European Journal of Vocational Training, 40*(1) 52-66.

Robson, C. (2007). *Real world research.* Chichester: John Wiley & Sons Ltd.

Schaap, H., de Bruijn, E., van der Schaaf, M. F., Baartman, L. K. J., & Kirschner, P. A. (2011). Explicating students' personal professional theories in competence-based vocational education through multi-method triangulation. *Scandinavian Journal of Educational Research, 55*, 567-586.

Schaap, H., de Bruijn, E., van der Schaaf, M. F., & Kirschner, P. A. (2009). Students' personal professional theories in competence-based vocational education: the construction of personal knowledge through internalisation and socialisation. *Journal of Vocational Education and Training, 61*, 481-494.

Schmidt, H. G., & Boshuizen, H. P. A. (1993). On acquiring expertise in medicine. *Educational Psychology Review, 5*, 205-221.

Scott, A. A., Clayton, J. E., & Gibson, E., L. (1991). *A practical guide to knowledge acquisition.* Reading, MA: Addison-Wesley.

Smagorinsky, P., & Smith, M. W. (1992). The nature of knowledge in composition and literary understanding: the question of specificity. *Review of Educational Science, 62*, 279-305.

Teigland, R., & Wasko, M. M. L. (2004). Extending richness with reach: participation and knowledge. In P. Hildreth, & C. Kimble (Eds.), *Knowledge networks: innovation through communities of practice*. Hershey: Idea Group Publishing.

Toulmin, S. E. (1958). *The Uses of Argument.* New York, NY: Cambridge University Press.

Trochim, W. M. K., & Cabrera, D. (2005). The complexity of concept mapping for policy analysis. *E:CO, 7*(1), 11-22.

Verloop, N. (2009). De leraar. In N. Verloop, & L. Lowyck (Eds.), *Onderwijskunde* [Educational Science]. Groningen, The Netherlands: Noordhoff Uitgevers.

Wilensky, H. L. (1964). The professionalization of everyone? *The American Journal of Sociology, 70*(2), 137-158.

Wilensky, U. (1991). Abstract meditations on the concrete and concrete implications for mathematics education. In I. Harel, & S. Papert (Eds.), *Constructionism*. Norwood, NJ: Ablex Publishing Corp.

Figure 1. Association screen

Figure 2. Construction screen

Figure 3. Classification screen

Figure 4. Cluster evaluation screen

Table 1. *Questions on the usability of the student module of Brainweaver*

Questions	--	-	-/+	+	++
1. After the short instruction, the program was easy to use.	0	1	0	3	6
2. I could draw my concept map as I wished.	0	0	2	3	5
3. I did not make many errors.	0	3	1	2	4
4. I find it pleasant to work with the program.	0	0	1	3	6

Table 2. *Questions on the use and possible importance of the student module*

Questions	--	-	-/+	+	++
5. I understand what a personal professional theory is.	0	1	0	4	5
6. Paying attention to the development of your personal professional theory should be included in the curriculum.	0	0	0	5	5
7. Constructing a concept map and discussing this concept map with a teacher trainer is a good way to pay attention to personal professional theories.	0	0	0	7	3
8. The computer program supports the construction of a concept map of my personal professional theory.	0	0	2	3	5

Table 3. *Summary of the semi-structured group interview with the assessors*

1. The usability of the assessor module	No problem. After a short introduction we found our way as we went along. It was easy to use and in the end it took no more than half an hour to assess a concept map.
2. Understanding the concept of a PPT	The image students have built of their future profession from workplace experience and theory.
3. Sequence of assessing elements of a PPT	First, the allocation of examples to domains of knowledge and then the evaluation of the three qualities at each concept.
4. The coding scheme of the domains of knowledge	After using it for a while, it became clear, although borders between different domains were open to question.
5. The operationalisation of concreteness	The problem was the level of concreteness. In a concept map, expressions consist of key words, so you have to interpret in order to grasp the full meaning of the utterance. For instance, an example such as "English literature" within the concept "Knowledge of the English language" could be debated regarding concreteness. We agreed on an open-minded interpretation.
6. The operationalisation of specificity	The problem was how specific: education, secondary education, or teaching a language in secondary education. We agreed on the first interpretation.
7. The operationalisation of complexity	This was the most difficult quality to evaluate. We developed the same view on complexity after a while and adjusted some assessments, but we are still not sure of our assessment of this quality.
8. Qualities that are missing	Richness is also a useful quality, but perhaps profoundness could be added as a fifth quality, although it is not clear how to operationalise this quality.
9. The meaning for a student of a formative assessment of his/her PPT	The student becomes aware of what he/she knows and thinks about his/her future profession in a comprehensive and connected way.
10. The meaning of these formative assessments for an assessor	An assessor can stimulate and help the student to develop facets of him/herself that are not yet well developed. It becomes easier to counsel a student if you have an evaluation of the concept map of his/her PPT, because it gives you a lot of information to discuss with them.
11. Recognition of the students through their PPT.	We know the students whose PPTs we assessed and we recognised them in their PPTs. It was no surprise that the concept maps showed big differences. We discovered facets of students we did not know about. Especially when a student is introvert, this method can reveal thoughts and opinions that are important for teacher trainers to know.

Table 4. *The distribution of the domains of knowledge*

	Concept Maps				Interviews			
	Abs. freq.	Rel. freq.	*M*	*SD*	Abs. freq.	Rel. freq.	*M*	*SD*
Vocational field	22	2.6	0.9	2.0	25	2.7	1.1	1.4
Organisations	45	5.3	1.9	2.0	72	8.0	3.4	2.4
Social environment	4	0.4	0.2	0.6	9.0	0.1	0.5	1.4
Target group	127	15.0	5.3	4.4	129	14.3	5.6	5.7
Technical-instrumental processes	599	70.8	25.1	11.0	603	66.9	26.2	9.5
Personal development	52	6.1	2.2	3.1	64	7.6	2.8	2.9

Table 5. *The values of the four qualities*

	Concept Maps		Interviews		Absolute difference of means	Two-tailed *p*-value
	M	*SD*	*M*	*SD*		
Concreteness	53.8	21.5	57.1	13.9	3.3	0.5
Specificity	48.3	17.6	56.6	15.1	8.3	0.1
Complexity	41.8	23.9	32.1	14.7	9.7	0.1
Richness	40.6	14.5	41.8	1.0	1.2	0.8