

Teacher competences for supporting students' reflection

Standards, training, and practice

Agaath Dekker-Groen

Manuscript committee:
Prof.dr. E. de Bruijn
Prof.dr. E.P.J.M Elbers
Prof.dr. J.W.M. Kessels
Prof.dr. W.L.L. Volman
Prof.dr. M. de Winter (chair)

This research was conducted in the context of the Dutch Interuniversity Centre for Educational Research (ICO)

Print: Wöhrmann Print Service

ISBN: 978-90-393-5928-0

© 2013 A.M. Dekker-Groen.

All rights reserved. No parts of this publication may be printed or reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior permission in writing from the author. Citing and referencing for non-commercial or academic use is encouraged, provided that the source is mentioned.

Teacher competences for supporting students' reflection

Standards, training, and practice

Docentcompetenties
om het reflecteren van studenten te ondersteunen
Standaarden, training en praktijk
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht
op gezag van de Rector Magnificus, prof.dr. G.J. van der Zwaan,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

vrijdag 19 april 2013 des ochtends te 10.30 uur

door

Agatha Maria Dekker-Groen

geboren op 15 augustus 1954
te Heerhugowaard

Promotor: Prof.dr. K.M. Stokking
Co-promotor: Dr. M.F. van der Schaaf

The studies in this doctoral thesis were funded by the Dutch Organisation for
Scientific Research NWO-PROO 411-06-311

Contents

Preface	7
1 Introduction	9
1.1 Context, purpose and research questions	9
1.2 Background in theory and research	10
1.3 Overview of studies and chapters	16
2 Teacher competences required for developing reflection skills of nursing students	19
2.1 Introduction	19
2.2 Background	20
2.3 Framework of teacher competences	22
2.4 The study	24
2.5 Results	28
2.6 Discussion	32
2.7 Conclusion	34
Appendix: Task domains, tasks, indicators, and questions in the instrument	35
3 Performance standards for teachers supporting nursing students' reflection skills development	41
3.1 Introduction	41
3.2 Background	42
3.3 Development of the rubrics framework	45
3.4 The study	50
3.5 Results	52
3.6 Discussion	55
3.7 Conclusion	57
4 A teacher competence development programme for supporting students' reflection skills	59
4.1 Introduction	60
4.2 Theoretical background	62
4.3 Methods	63
4.4 Results	70
4.5 Conclusions and discussion	82
Appendix: Example of a teacher competence description and a rubric	86

5 Teachers' questions during feedback dialogues	87
5.1 Introduction	88
5.2 Theoretical framework	89
5.3 Methods	93
5.4 Results	99
5.4.1 Minutes, turns and words in the dialogues	99
5.4.2 Frequencies of teacher and student utterances	100
5.4.3 Sequences of categories in the dialogues	102
5.4.4 Effects of the training on the occurrence of utterances and sequences	106
5.5 Conclusions and discussion	109
Appendix A: Coding scheme question categories	113
Appendix B: Coding scheme response categories	114
Appendix C: Example of coding and sequences (questions and responses)	115
6 Teachers' feedback to support students' reflection	117
6.1 Introduction	118
6.2 Theoretical framework	119
6.3 Methods	122
6.4 Results	129
6.4.1 Duration and numbers of segments	129
6.4.2 Teacher and student utterances in feedback dialogues	130
6.4.3 Sequences of teachers' feedback categories and students' thinking activities	134
6.4.4 Effects of the training on the occurrence of utterances and sequences	140
6.5 Conclusions and discussion	144
Appendix A: Coding scheme teacher utterances	148
Appendix B: Coding scheme student utterances	149
Appendix C: Example of coding and sequences (teacher's feedback and student's reactions)	150
7 Conclusions and discussion	151
References	161
Summary	175
Samenvatting	181
Dankwoord	189
Curriculum vitae	191

Preface

Every word is important...

With Scrabble a long word on the right place brings in many points. In a conversation the right word on the right moment often does wonders. From that we all know examples. In many situations it is an art to convey the message concisely.

Pre-eminently this counts in a doctoral thesis. Every word is weighed up and every sentence is refined until a clear text emerges. A scientific paper has still another difficulty because of the maximum number of words that may be used.

Every word is important: In this doctoral thesis I describe research activities and findings of the expertise-based approach of nursing teachers' professional development activities. In the research project giving feedback is central. In a feedback dialogue one single word can elicit a whole thinking process. Utilising video interaction guidance for teachers who support students' development, fits me very well. Then also non-verbal communication shows to be important.

The front cover shows a portrait of Florence Nightingale, the first nurse, as a symbolic reference to nursing education. Furthermore, it can be seen how video interaction guidance takes place.

Agaath Dekker-Groen,
Houten, april 2013

1 Introduction

1.1 Context, purpose and research questions

In many professions the tasks to be performed are regularly changing. From practitioners in such professions it is expected that they flexibly anticipate and implement new practices. This requires willingness and skills for lifelong learning. A great part of such learning during the career takes place on the job and asks for reflection on work methods and personal experiences. This demands the setting of own learning goals, trying things out and adapting the own behaviour. Reflection skills currently are seen as an important part of such self-regulated learning (Ifenthaler, 2012; Thillmann, Künsting, Wirth & Leutner, 2009).

Institutes for vocational education try to prepare students for the necessity of continuously learning during the career by paying attention to the development of their reflection skills and supporting their learning process. In their turn this requires from teachers that they expand their pedagogical repertoire to include interventions and forms of interaction with their students which might contribute to this. Examples of such interventions and forms of interaction are posing questions and giving feedback (e.g. Chi, Siler, Jeong, Yamauchi & Hausmann, 2001; Chin, 2006; Nicol & Macfarlane-Dick, 2006).

This development is also topical in professions and education in health care. In Dutch nursing education courses many teachers are confronted with the task of giving the development of reflection skills of their students a clear place in the curriculum. They increasingly have conferences with students about their reflections. During such conferences they are expected to ask them questions and give them feedback adaptively and effectively. This partly demands new competences from teachers while in their own professional education they have not been prepared for the learning process which now is expected from them. Teachers can be supported in this process by offering and organising on the job professional development activities during which they receive information and instruction about the new tasks expected from them. This gives them the opportunity to discuss these new tasks, to purposefully exercise, to reflect on it and to receive feedback (Penuel, Fishman, Yamaguchi & Gallagher, 2007).

Offering a comprehensive programme of professional development activities which can be expected to effectively contribute to the development of these new competences of teachers, requires a systematic and well-founded preparation, in

collaboration with the target group and based on state-of-the-art knowledge. Studying the course, appraisal and effects of a programme which is developed this way can contribute to the further development of our knowledge about effective approaches for professional development of teachers and its factors of influence. In this thesis we report on the development and conducting of such a programme. The studies in this thesis aim to gain insight in the possibilities for furthering the development of teacher competences stimulating students' reflection skills. Specifically, the research is about teachers in senior secondary vocational education who are expected to support students' reflection skills development and investigates how this teacher competence development can be furthered by offering an integrative set of knowledge-based activities in a training programme. The research question is: How can teachers' competences in supporting nursing students' reflection skills be developed by a teacher training programme? The sub-questions are: (a) What characteristics should the programme have in view of current insights in the literature on teacher professional development?, (b) How can the programme be designed in view of the knowledge about teachers' interventions and interactions which might contribute to students' reflection skills development?, (c) What are the feasibility and the appreciation of the programme?, (d) What are the effects of the programme?, and (e) Which conclusions can be drawn about the value of a knowledge-based approach for extending teachers' pedagogical repertoire?

The next section describes theoretical and empirical backgrounds about professional development, competence development and learning, seen as a process which can be supported by offering activities which are based on relevant knowledge. The last section of this chapter gives an overview of the studies and the outline of this thesis.

1.2 Background in theory and research

New tasks for teachers require processes of learning and development

Teachers who are confronted during their career with new tasks, such as supporting students in their development of reflection skills, have to develop themselves professionally (Biesma et al., 2008; Mann, Gordon & MacLeod, 2009, Schön, 1991). To support teachers in this, programmes can be offered both individually and to a group of teachers (Bandura, 1977; Kuijpers, Houtveen & Wubbels, 2010). Such programmes can be organised by a school or by external experts, and can include all sorts of activities: giving information, offering

explanations, instruction and demonstrations, letting the participants discuss and exercise, giving feedback, et cetera.

Such programmes can contribute to teachers' learning, as a cognitive and affective process, to their competence development aimed at learning to master a task, and to the professional development during their career. Giving students effectively feedback to support their development of reflection skills can be seen as a new competence which can be developed by suitable activities as part of the continuous professional development. Below we briefly discuss the concepts professional development, competence development, and learning.

Activities for professional development

In former days activities for teacher professional development were often directed by their form, for example that of a class or 'a workshop'. According to Garet, Porter, Desimone, Birman and Suk Yoon (2001), in view of realising effects the exact form of the activities is less important than the content, the kind of learning activities aimed at such as including occasions for active learning, the duration or intensity, and the embeddedness in the teachers' daily work setting by matching current developments in the schools and allowing the teachers to exchange ideas and experiences. Penuel and colleagues (2007) confirm the importance of these factors and add that for implementation of a curriculum innovation also coherence between the different parts of the content and between the activities are important, as well social support. According to Whitcomb, Borko and Liston (2009) traditional forms of in-service staff development workshops currently are being replaced by activities from a new paradigm: situated in practice by focusing on the teachers' concrete tasks, embedded in a social context in which cooperative learning is possible, applicable in daily practice, and focused on the learning of both the teachers themselves and their students.

This shift from traditional organisational forms of professional development activities such as a workshop to the initiating of coherent activities which are embedded in practice and further joint learning, goes together with attention for realising actual changes and improvements in educational practice. For this to occur, it is required that teachers not only develop new knowledge and skills that traditional activities often are aiming at, but also adapt their behaviour (Garet et al., 2001). Furthermore they should develop an attitude by which they themselves recognize that things must go better and see ways for improvement (Evans, 2008).

Teacher competence development: components, processes and factors

Teacher professional development pertains to their development as a professional during their career. With each teacher this development takes place to a certain

extent and in a certain direction. It concerns various aspects such as their practical knowledge, their beliefs about teaching and learning, their personal identity as a professional, and the responsibilities they experience or take on themselves. Compared to professional development, competence development is more marked out in focus and time horizon and can be worked on and supported in a more directed way. It might conceptually be seen as a part of professional development. After all, one of the ways to develop oneself as a professional is by growing in relevant competences.

Dealing with teacher competence development directed to changes and results in practice, four components are often distinguished: cognitions (knowledge, beliefs and attitudes), intentions (intended behaviour or effects), behaviour (of teachers), and effects of that behaviour (on students). The first three components are widely known and are for example also distinguished in the Theory of Planned Behavior (Ajzen, 1985; Azjen & Fishbein, 2005). In the literature about teaching and teacher behaviour to these components the effects on the learning of students are often added as a fourth component (e.g., Desimone, 2009; Roelofs & Sanders, 2007).

Sometimes these four components are described as if they represent a linear causal chain: cognitions → intentions → behaviour → effects. In this vein for example Aguirre and Speer (2000) put that the behaviour of teachers is being influenced by their beliefs and intentions. The process however can also pass off in another way. Following Guskey (1986, 2002) teachers will adapt their beliefs more likely if they see that their (new) behaviour works out as intended. In his model Guskey describes the following chain of events: staff development → change in teachers' practice → change in student outcomes → change in teachers' beliefs and attitudes.

Adjustments in beliefs and attitudes take place not easily at the start of a process of competence development. Teachers can not be forced, for instance by imposing new standards, to look in another way to their own practice (Evans, 2008). Activities are needed by which teachers not only are being informed and get instruction but also can try out new working-methods, exchange experiences with colleagues, get feedback based on observations or comes directly from their students, and reflect on their experiences (Kwakman, 2003).

Also Airasian and Gullickson (1994), Rosenstein (2000), and Ross and Bruce (2007), point toward the importance of feedback and reflection for development and improvement. Both should result in adjustments of behaviour and beliefs of teachers and in better learning results of students. As examples of interventions which might stimulate reflection and professional development with teachers, they suggest the use of video, as a source of feedback and of standards as anchor points for reflection.

The developments sketched above in the thinking about and the practice of teacher professional development also occur in medical education and nursing education. These developments include attaching a growing importance to the directed development of competences, exercising in the own practice, getting and using feedback, and reflection on goals, approaches and results (Cader, Campbell & Watson 2005; Dexter et al., 1997; Steinert et al., 2006).

Competence development and learning activities: methods and approaches

Competence development by teachers requires learning activities (Shuell, 1988; Vermunt & Verloop, 1999). These learning activities preferably are embedded in a social context which allows for forms of cooperative learning and offers possibilities for active learning (Bolhuis & Voeten, 2004; Meirink, Meijer & Verloop, 2007; Whitcomb et al., 2009). Activities should be included of sufficient duration and intensity which are part of the daily work practice (Garet et al., 2001; Penuel et al., 2007).

Along with the shift described above in the approach underlying the activities supporting further teacher development, more attention is given to realising actual changes and improvements in educational practice. This supposes that teachers have room to choose their own position in face of the new knowledge and methods offered to them in relation to the learning goals they want their students to achieve (Fenstermacher, 1987).

In developing a programme of activities intended to support the learning and the development of competences of teachers, several approaches can be distinguished. Coaches and trainers can support teachers by letting them think about their current behaviour and about possible alternative behaviours. This is also known as a 'developmental-reflective approach' (Kuijpers et al., 2010). Teachers can also be supported by the so called 'training-application approach'. This approach implies: informing and instructing teachers, letting them use certain knowledge and exercise skills in their own practice, and exchange and discuss experiences with colleagues (Joyce & Showers, 2002; Tillema & Veenman, 1987). As in theory and research much knowledge has been developed concerning reflection and feedback and assuming that also teachers have to develop their knowledge and skills in these topics, it is obvious to link up with both approaches. Preferably a training programme in these topics connects with the teachers' own experiences and their practical knowledge (bottom-up) and also enables them to connect new knowledge (expert-driven) with knowledge, experiences and beliefs that already have developed, by way of enlargement, enrichment and/or restructuring (Putman & Borko, 1997; Tillema & Imants, 1995). Moreover the developmental-reflective

approach is a very convenient approach to give teachers individual guidance on their own practice.

Learning and competence development by a training programme

To effectively contribute to the development of competences of teachers in giving feedback on students' reflections demands: a) offering information about feedback and reflection (concerning forms and standards) and stimulating discussion and dialogue; b) letting exercise and gather experiences in practice; c) giving the participating teachers feedback on this; d) furthering teachers' reflection.

Such an approach might be seen as a modern form of training in which the traditional components of a training course including instruction, exercise, and feedback, are supplemented by attention for discussion, practical experience and reflection. In such a training programme the trainer's interaction with the participating teachers is based on the same pedagogical repertoire as is presented and discussed during the training programme as part of the repertoire the teachers should develop themselves (Fishman, Marx, Best & Tal, 2003).

To end with, we outline how teachers in such a training course might learn. Trying new tasks in practice gives teachers possibilities to learn, especially as they get information about how to expand good practices and how to overcome less strong aspects (Sadler, 1989, 2010). This might result in changes in teachers' cognitions, intentions, and actions and in outcomes regarding students' learning (Ajzen, 1985; Ajzen & Fishbein, 2005; Roelofs & Sanders, 2007; Desimone, 2009). These four components are brought together in a model of competent teacher performance by Roelofs and Sanders (2007). In this model teachers make considerations based on their personal professional knowledge base and they decide on and plan their actions. These actions can have consequences in terms of teachers' learning and of students' learning activities. Outcomes which teachers might perceive regarding students' learning can be seen and used as feedback on teachers' own actions. Consequently student learning can influence teachers' cognitions (knowledge, beliefs and attitudes) and future actions.

For describing the mutual interconnections between these components also several other models are available (e.g. Guskey, 2002; Clarke & Hollingsworth, 2002). Some of these models also include the influence of feedback, e.g. reciprocal feedback as part of the model of self-regulated learning of Butler and Winne (1995). Their model, which we take as a starting point, represents processes in a cognitive system containing internal feedback paths leading to performance, external feedback, and new tasks. Nicol and Macfarlane-Dick (2006) adapted the model of Butler and Winne by adding interactions between internal and external feedback. A key feature of their model is that the learner, as a feedback receiver, is

assumed to play a central and active role in all feedback processes. In the research reported on in this thesis, external feedback is given in two dialogues, namely between teacher and student and between trainer and teacher. These two dialogues can be added to the model of Nicol and Macfarlane-Dick (2006). Integration to the matters discussed with the model of Roelofs and Sanders (2007) results into a comprehensive model which describes teacher's learning by feedback during a training programme. See Figure 1.1.

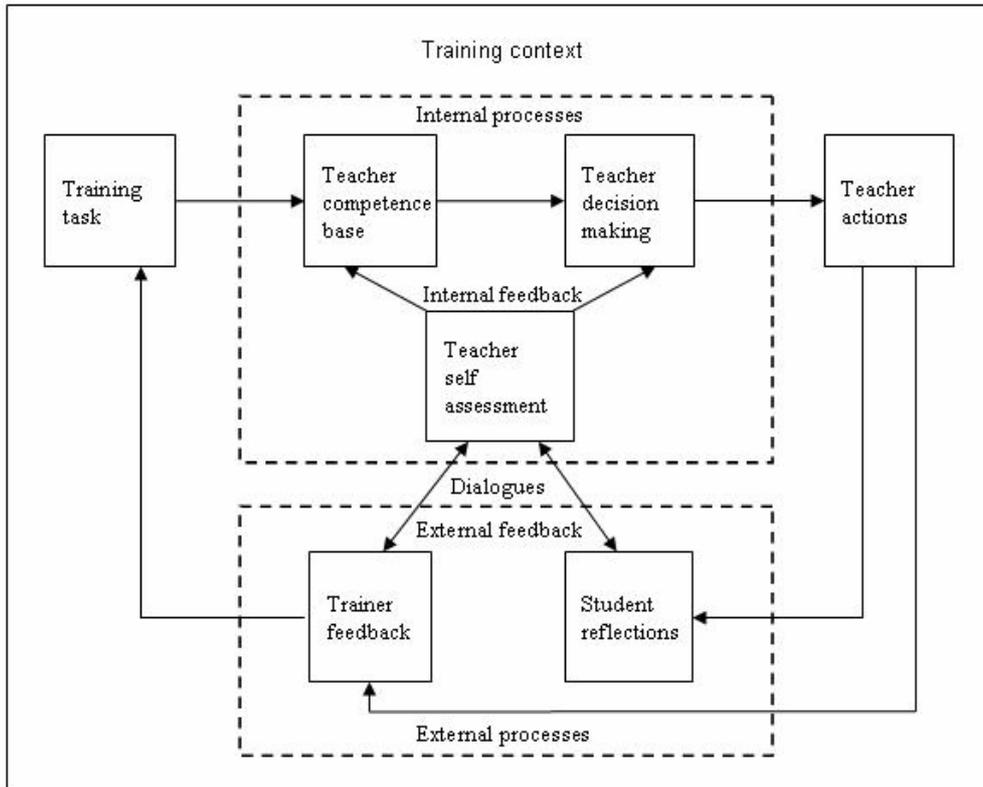


Figure 1.1 A model of teacher learning by feedback during a training programme

In Figure 1.1 five components are distinguished. The first component, training task, includes three core features of effective competence development (Desimone, 2009) (i.e., content focus, active learning, and collective participation). The two other core features (i.e., coherence and duration) are characteristics of the training programme as a whole.

Each training task activates teachers' internal processes. This is depicted in the second component, which refers to teachers' competence base, their decision making and self-assessment. Teachers' competence base represents not only teachers' cognitions about the content of the training in giving feedback to support students' reflection but also teachers' personal characteristics, practical knowledge and professional skills. Next, teachers' (more or less tacitly) decision making is influenced by intentions. Intentions imply for example a teacher's goal setting about how to support students. Further, teachers' self assessment implies a teacher's consideration and judgement of the usefulness of this feedback.

The third component, teacher actions, pertains to teacher's acting in educational practice. In our case a teacher gains experience in giving feedback on students' reflections. Effects of teacher actions can be derived from measuring students' reflection skills, for instance during reflection conferences. Both feedback from a student and feedback from a trainer or a colleague are sources of external feedback for a teacher (the fourth component). The training context, the fifth component, represents characteristics of the training programme, the nursing schools and circumstances in the field of nursing education.

This model has been used as a conceptual source in a tentative way during the design and development of the training programme which occupies a central position in this thesis. See chapter 4 for a description of this programme and chapter 7 for closing evaluative remarks on this training programme as a specimen of a comprehensive set of teacher professional development activities.

1.3 Overview of studies and chapters

For developing and assessing competences of teachers in view of their further professional development it is necessary to settle criteria and standards (Sadler, 1989). Criteria concern the tasks and activities teachers have to (learn to) conduct and standards pertain to the extent or the level of quality to which they minimally have to do this. Because supporting students in their development of reflection skills by giving them feedback is a new task for nursing teachers, at the start of this research no definite and generally accepted criteria and standards were available. The first two studies were conducted to develop these criteria and standards.

Based on a study of relevant documents and the questioning of experts a comprehensive set of criteria was developed and evaluated in a Delphi-study in two rounds by experts and stakeholders (study 1, see chapter 2).

In a standard-setting study using structured questioning of and discussion within teams of teachers working on six schools for nursing education, next to these

criteria standards were developed for the level of execution of relevant tasks in supporting students' reflection skills (study 2, chapter 3).

The criteria and standards were input for the development of a training programme. In addition input was used from a literature study regarding: reflection, posing questions, giving feedback, competence development, and professional development of teachers. The resulting programme included team meetings, the making of video recordings of feedback dialogues between teachers and individual students, meetings with individual teachers to discuss fragments of these video recordings, and periods of independently exercising by teachers in their practice.

The training programme was executed in five school teams, and its course, appreciation by the participating teachers, and effects were investigated in a quasi experimental intervention study including a training group and a control group (study 3, chapter 4).

In order to deepen the insight in the ways teachers interact with students during conversations and to stimulate their reflection by giving prompts, posing questions, and giving feedback, and the insight in the effects of the training programme, in both groups (training and control) per teacher two video registrations of a feedback dialogue with an individual student were made, at the beginning and at the end of the training programme. These videos were observed with a system of categories based on the literature and after that analysed in different ways in search of the teachers' and students' repertoire of utterances and the patterns of sequences in their interactions. These observation studies are reported on in the last two chapters reporting on the empirical studies conducted (studies 4 and 5, chapters 5 and 6).

A final chapter (chapter 7) includes conclusions and a discussion.

2 Teacher competences required for developing reflection skills of nursing students¹

Abstract

This paper reports on a study in senior secondary vocational education designed to develop a framework for teacher competences needed to support nursing students in developing their reflection skills. For healthcare-related professions such as nursing there is a growing attention for developing reflection skills. Little is known about teacher competences required to support the development of reflection skills in nursing students. Developing a framework of teacher competences can contribute to filling up this gap. These competences are specified in 91 indicators distributed over six task domains. A Delphi-study was conducted in the first half of 2008 to get consensus on a framework of teacher competences required for creating the learning environment needed for developing reflection skills of nursing students. Experts judged the teacher competences on a 7-point Likert-type scale.

In the first round the mean scores on the competences were already high and only minor revisions were needed. In the second round mean scores increased whereas standard deviations decreased compared to the first round. These changes were statistically significant. Coaching was seen as the most important task domain. Consensus has been reached on the teacher competences to be used in nursing education to develop students' reflection skills. The framework of competences may be a source for curriculum development concerning reflection skills and for teacher training programmes for learning to coach nursing students' reflections.

2.1 Introduction

In vocational education for healthcare-related professions such as nursing, reflection skills are seen as essential for lifelong learning and development and for taking well-considered decisions and actions in every-day practice. The concept of reflection has been widely adopted in many nursing curricula (Wong, Kember, Chung & Yan, 1995). Reflection in nursing practice has also received considerable

¹ A slightly different version of this chapter has been published under the same title in *Journal of Advance Nursing* 67 (7), 1568-1579.

international attention, e.g. in Australia, England and New Zealand (Teekman, 2000). Unfortunately, various different definitions, models and approaches have been suggested for the meaning of reflection and so far not much consensus has been arrived at (e.g., Boud, Keogh & Walker, 1985; Burton, 2000; Platzer, Blake & Ashford, 2000; Hatton & Smith, 1995; Mackintosh, 1998; Procee, 2006; Rodgers, 2002). Also about how to learn to reflect, no clear-cut conclusions are available and becoming reflective seems to be a tough learning goal (Elshout-Mohr, Oostdam & Overmaat, 2002; Korthagen & Wubbels, 2001; LaBoskey, 1993).

The study is conducted in the Netherlands in the domain of nursing education at the level of senior secondary vocational education (ISCED², level IV). The goal of senior secondary vocational education is to deliver students to the labour market as well as to prepare and stimulate students to continue their educational career in higher professional education. In compliance with the Law on Educational Professions (Ministry of Education, Culture and Science, 2007) vocational institutes have to implement competence-based education. This innovation implies that compared to current practice in nursing education, students' reflection skills will be more prominent and teachers will need to adjust their roles and tasks. As a consequence, teams of nursing teachers will need to alter their curriculum to a more student-centred approach. To many teachers it is unclear which competences are important for developing nursing students' reflection skills and how these competences can be developed. A framework of teacher competences as a guide for teachers' every-day practices can contribute to fulfilling this need. Competences which represent tasks teachers have to carry out can only be formulated in a relative way because it is contestable what can be counted as good education (Goodwin, 2008). A Delphi-study makes it possible to develop formulations of competences which are shared by experts (Linstone & Turoff, 1975).

2.2 Background

To help develop nursing students' reflection skills we need to determine what is meant by reflection. We define reflection as a conscious and well-considered process of thinking about and interpreting situations, events, processes, experiences and emotions in relation to each other and to available knowledge, aiming for better comprehension and for learning (Bradbury-Jones, Hughes, Murphy, Parry & Sutton, 2009; Ertmer & Newby, 1996; Gillespie, 2007; Hatton &

² International Standard Classification of Education, a department of UNESCO.

Smith, 1995; Kelchtermans, 2007). In the process of reflection not only behaviour and competences but also deeper personal layers of beliefs, identity and involvement are included (Korthagen & Vasalos, 2005). This means that in their reflection process students include different objects of reflection, for instance theory (e.g., in their textbooks), their experiences in practice, and their own competences, values and emotions. During the process of reflection students use cognitive, affective, and regulative activities such as described by for instance Shuell (1988), Vermunt and Verloop (1999) and Mansvelder-Longayroux, Beijaard and Verloop (2007). Examples of these activities are analysing (cognitive), attributing (affective), evaluating (regulative). In the sequel we call such activities ‘thinking activities’.

Thinking activities are comparable to activities in well-known taxonomies of learning objectives (e.g., Krathwohl, 2002). Previous research into relevant thinking activities during students’ reflection processes show that at least eight thinking activities are used, namely describing, analysing, structuring, evaluating, attributing, explaining, concluding, and formulating intentions (Oosterbaan, Van der Schaaf, Baartman & Stokking, 2010). For concise descriptions of these thinking activities, see Table 2.1. A reflection process often starts with a description of a situation or an experience and it frequently ends with the formulation of one or more intentions for improvement (Oosterbaan et al., 2010).

Table 2.1
Thinking activities during students’ reflection processes

Thinking activities	Short descriptions
Describing	Precisely, systematically and selectively telling or writing about a situation, experience, reasoning or emotion that gives rise to reflection.
Analysing	Checking and clarifying relevant aspects of the description, for instance what the student and other persons did, thought, wanted or felt.
Structuring	Developing some structure to discover patterns or connections, for instance by making categories or ordering things chronologically.
Explaining	Giving reasons for or causes of what happened or what the student did, thought, wanted or felt, in reaction to the situation, a personal drive or based on a reflex.
Evaluating	Valuing or assessing a situation, experience, reasoning or emotion by using certain criteria and norms or goals.
Concluding	Drawing conclusions about a situation, experience, reasoning or emotion in relation to the students’ own future thinking and/or acting.
Attributing	Attributing aspects or effects to oneself (internal attribution) and/or to things outside oneself (external attribution) and making it meaningful for other situations.
Formulating intentions	Considering what can be learned from the reflection and how this can be used in a new situation.

Teachers need competences for supporting students' reflection processes. Nowadays teacher competence is defined as an integrating concept denoting teachers' knowledge, skills, and attitudes in a certain context while performing professional tasks (Sultana, 2009). In the context of developing students' reflection skills, preferably teachers' support is tailor-made. This means that instruction, assignments, coaching, and feedback link up to students' zone of proximal development (Vygotsky, 1978). This kind of support is part of the cognitive apprenticeship model (Collins, Brown & Newman, 1989). By modeling, scaffolding, and fading the teacher adaptively supports students, steadily offering students more room for regulation by themselves.

Supporting nursing students' development of reflection skills often takes place on the spot during coaching conferences about students' practices, skills, and development. Important elements of coaching conferences are feedback and interaction (Brinko, 1993; Chi, Siler, Jeong, Yamauchi & Hausmann, 2001; Graesser, Person & Magliano, 1995; Mory, 2004; Shute, 2008). For feedback and interaction to be effective, involvement of the student, relevance of the feedback for a certain learning goal and balance between positive and corrective feedback are very important (Blatt, Confessore, Kallenberg & Greenberg, 2008; Gibbs & Simpson, 2005; Ilgen & Davis, 2000; Sadler, 1989).

2.3 Framework of teacher competences

We have developed a framework of teacher competences required for supporting nursing students' grow in reflection skills. This development process was threefold. Firstly, for the design of the proposed framework we carried out a literature study and also studied national policy documents used in Dutch nursing education. We searched educational and nursing journals in the ERIC database from 1980 to 2007 using the following descriptors in successive searches: reflective thinking, critical reflection, teaching models, nursing education, and feedback. We also studied the so-called qualification dossier of Dutch nursing education (Colo & Calibris, 2007), the national experimental competence profile for nursing (Movisie, 2007), and several national curriculum development materials (SCB: Foundation Consortium Vocational Education, 2007). These activities offered insight into the concepts of reflection and feedback, national examination requirements for nursing, and current innovations in Dutch vocational education. We concluded that teachers have to take care of good nursing education and have to create a learning environment in which they use certain competences to support students' reflection skills development. Therefore

a framework of teacher competences was necessary. Secondly, for developing this framework we linked up with the design approach of Morrison, Ross and Kemp (2004). This approach considers the perspective of the learner and focuses on factors that influence learning outcomes, for example: “Which strategies are most appropriate in terms of objectives and learner characteristics? What support is needed for successful learning?”

Morrison et al. distinguished four fundamental components of instructional design: learners, objectives, methods, and evaluation. We used these four components to form the basis of our framework of six task domains. In task domain 1 the teachers themselves are the learners (component 1). The teachers have study meetings to come to a shared vision of what reflection skills education for their students should be all about. They have to learn about the content and goals of reflection education and about how students develop reflection skills. This is prerequisite to preparing a learning environment to support nursing students’ reflection skills. Task domain 2 is about objectives (component 2) to make the development of nursing students’ reflection skills possible. The task domains 3 to 5 are about methods (component 3) to instruct nursing students’ reflection skills; task domain 3: instruction about what reflection is and how thinking activities can be learned; task domain 4: coaching the process of reflection skills development; task domain 5: feedback on task performance to optimise this process. Evaluation (component 4) which forms task domain 6, can be done by assessing reflection conferences and reflection writings. We have named these six task domains: (1) Developing education for learning to reflect; (2) Supporting students’ reflection skills development; (3) Giving students reflection instruction and assignments; (4) Coaching students to develop reflection skills; (5) Supporting students’ development through feedback; (6) Holding reflection conferences and giving assessments.

Thirdly, we filled in the framework by specifying tasks and indicators. For each of the six task domains we established tasks that teachers should perform to create a learning environment which supports the development of nursing students’ reflection skills. This subdivision resulted in fifteen tasks with two to four tasks per task domain. For each task we have drawn up concrete descriptions leading to four to eight indicators per task. We have based their content on insights from the literature studies and on the national documents for nursing education.

This framework of teacher competences was input for a questionnaire used in a pilot study: six task domains including 15 tasks covering 91 indicators. The results of this were used to improve the initial questionnaire. The revised questionnaire was input for the study. All task domains, tasks, and indicators and the questions posed per indicator and task domain are added in the Appendix to this chapter.

2.4 The study

Aim

The aim of the study was to develop a framework of teacher competences required for creating a learning environment for developing nursing students' reflection skills in senior secondary vocational education. The research question of the study was: What are relevant and acceptable descriptions of teacher competences for supporting nursing students' development of reflection skills?

Design

We used a Delphi method to capture experts' judgments of a framework of teacher competences for supporting students' reflection skills development in order to optimise the competence descriptions and to reach support and consensus (Gorney & Ness, 2000; Keeney, Hasson & McKenna, 2006; Linstone & Turoff, 1975; Smith & Simpson, 1995). The number of rounds needed depends on the stability and consistency (indicating consensus) in the responses of the participants. It is good practice to establish the percentage agreement that will be accepted as an appropriate level of consensus about the standards before data collection begins (Keeney et al., 2006). There are no scientific guidelines for a minimum percentage agreement between participants to declare consensus. The definition of consensus as a minimum of 75 percent agreement was used for instance by Murry and Hammons (1995), Keeney et al. (2006) and Williamson (2007), and can be seen as sufficiently robust to determine agreement between participants.

To check the quality of the framework of teacher competences we developed a questionnaire including five questions per indicator and also per task domain. The questions imply the extent to which the indicators and task domains are relevant, unambiguously formulated, important in teaching practice, representative, and complete. Each question had to be answered on a Likert-type scale, from 'not' (1) to 'to a high extent' (5). We additionally asked participants to write down the time in minutes they needed per task domain to score the questions and to judge the difficulty of this on a Likert-type scale, from 'very easy' (1) to 'very difficult' (5).

Participants

We selected a mixed group of experts, recognised in the fields of reflection, feedback, teaching, and nursing education, who are involved and motivated to

participate and who are stakeholders in the implementation of the results of the study (Clayton, 1997; Gorney & Ness, 2000; Keeney et al., 2006). In the pilot study eight experts participated from the same fields as those in the study. In the main study 28 experts participated of whom three teachers only in one round. Participants were eighteen teachers (with a mean of 17 years of work experience) from seven institutes of senior secondary vocational education for nursing, four managers, and six researchers.

Data collection

All data were collected in the first six months of 2008. The participants received an instruction and then they individually and independently filled out the questionnaire. We also asked them to add comments and suggestions to improve the questionnaire. Finally we evaluated the procedure with participants by verifying the time the needed and the difficulty they experienced while filling out the questionnaire.

The framework of teacher competences together with the questionnaire was sent to participants individually, together with an instruction about the procedure and some background information (Keeney et al., 2006). We explained the aim of the procedure and informed the participants that they would receive written feedback between the rounds about the degree of support and consensus among each other. We also gave information about what competences are and about the meaning and goals of reflection.

The suggestions for improvement participants gave in the pilot study focused mainly on phrasing and on lay-out. Participants needed between 2 and 2.5 hours to answer all questions. One participant found it difficult to answer the specific questions per indicator and gave her answers more globally by adding examples from her practical experience. The other participants generally found the questions neither easy nor difficult.

Based on the results we made the following improvements to the questionnaire. First we changed the 5-point Likert-type scale into a 7-point Likert-type scale to increase the possibility for variance. Also an 'I do not know' option was added. Secondly, changes were made by splitting up each of three rather broad indicators into two more focused indicators and moving two indicators from task domain 3 to task domain 5. Thirdly, minor improvements were made by adding one indicator about the teacher as a role model and one about asking students feedback and by making some descriptions more precise, for instance by adding that the intensity of coaching should be phased out in the process. In sum, the

framework of six task domains and fifteen tasks did not change. A summary of the framework including short descriptions of its elements is shown in Table 2.2.

Our assumption was that what participants believe and perceive as important is what they will do in practice. Therefore we asked them to rank the thinking activities underlying reflection skills (in round 1) and to rank task domains (in round 2). At the end of the first round we also asked participants to answer with yes or no three questions concerning the overall relevance, completeness, and feasibility of the competences for teaching nursing students reflection skills.

Ethical considerations

In accordance with Dutch Law, ethical approval is not obligatory for educational studies. No harm to participants could arise from the study. Experts participated voluntarily and received full information about the purpose and procedure of the study together with the questionnaires. Those wishing to participate indicated their consent by completing the questionnaires. We can confirm full confidentiality in the use of questionnaires filled out by participants.

Data analysis

We analysed the participants' scores per round and across rounds. Firstly, we have asked participants to score indicators and task domains on a Likert-type scale, from 'not' (1) to 'optimal' (7) agreement concerning the content relevance, unambiguous formulation, importance in teaching practice, representativeness, and completeness. We computed the means of the 91 indicators and six task domains.

Secondly, we analysed the degree of consensus among participants about the framework, defining consensus in terms of statistical consensus (percentage agreement). Following other Delphi studies (Keeney et al., 2006; Murry & Hammons, 1995; Williamson, 2007), we set beforehand a minimum percentage of agreement of 75% to accept an indicator. We analysed comments and suggestions of participants and made adaptations to optimise indicators for the next round.

Thirdly, we compared the mean scores and standard deviations between rounds. Higher mean scores of the indicators in a following round we took to imply more support. Indicators with mean scores above 5.5 on content relevance, unambiguous formulation, practical importance, representativeness, and completeness were seen as indicators with a high degree of support. Decreasing standard deviations of indicators between rounds were used as indicating improvement of consensus.

Table 2.2
Framework of teacher competences (short descriptions)

91 indicators (I)	15 tasks (T)	6 task domains (D)
I1-5: Develop a vision on reflection and feedback, develop the teacher's own capabilities, practice reflection skills to students, and consciously use materials and methods individually and together with colleagues.	T1 Develop competences for teaching reflection skills	D1 Developing education for learning to reflect
I6-13: Thinking activities are describing, analysing, structuring, explaining, evaluating, attributing, concluding, and formulating intentions. An explanation of these thinking activities is given.	T2 Get competent in distinguishing thinking activities	
I14-18: Create a positive learning climate, demonstrate reflection and feedback, evoke learning processes that support reflection, observe and assess reflections.	T3 Create conditions for learning to reflect	
I19-28: Show thinking activities, stimulate thinking activities by posing questions, give examples of concrete questions per thinking activity.	T4 Stimulate thinking activities	
I29-34: Demonstrate and connect learning questions to thinking activities, formulate learning goals, give reflection tasks, have students reflect on their functioning in groups, support reflection on experiences in nursing practice.	T5 Formulate short-term goals	D2 Supporting students' reflection skills development
I35-40: Teach reflection skills in a secure environment, coach students' practise in thinking activities based on learning questions coherently, show how to reflect on objects, let task complexity increase and teacher control decrease, pay attention to emotions, values and norms in a reflection process, teach students to reflect independently using different methods.	T6 Formulate long-term goals	
I41-46: Know what content information and skills students need so as to be able to reflect, apply instruction and methods to the competence level of students, pay attention to the interaction between theory and practice, give reflection tasks and teach specific skills, show how to reflect, explain and relate reflections to examples of professional development, speak about, inquire and discuss reflections on activities, events and situations.	T7 Give instruction about skills needed for reflection	D3 Giving students reflection instruction and assignments
I47-50: Choose or formulate reflection assignments, discuss them with the students and attune to their level of development, determine concrete intentions (who, when, on what, how, why) and let students perform them.	T8 Give reflection assignments	
I51-56: Know and use suitable coaching strategies, evaluate together with students their acting, be a role model, put in communication skills, pay attention to restraints and resistance of students in the reflection process.	T9 Care for guidance conditions	D4 Coaching students to develop their reflection skills
I57-62: Stimulate reflection of students through their own reflective attitude, promote an active attitude of students through explanations and examples, coach students while learning to reflect, give and organise feedback in the group which leads to growth and improvement, put in self assessment.	T10 Coach students in their learning process	
I63-68: Attune feedback to the intended development, learning goals, experiences and needs of students; give feedback on different aspects of their functioning in different situations, give concrete feedback on behaviour, give as much feedback as convenient and check comprehension, ask feedback from students about one's own functioning.	T11 Give students feedback purposefully	D5 Supporting students' development through feedback
I69-74: Give sincere feedback and also make it noticeable, give constructive feedback, explain, show and have students experience how to receive feedback, support interaction about feedback.	T12 Teach students how to receive and ask for feedback	
I75-79: Give feedback carefully, speak about the use of feedback, stimulate students to give feedback, stimulate conversation about tasks and standards, relate tasks and results to the students' development and the nursing competence profile, stimulate self assessment.	T13 Support students' use of feedback	
I80-86: Use a methodology, work with an open attitude, apply conversation skills and notice non-verbal signals, discuss reflection patterns, compare theory and practice, purposefully put in individual conversations, confront students with contradictions and formulate strategies for the future.	T14 Lead reflection conferences	D6 Holding reflection conferences and giving assessments
I87-91: Formatively assess content and process in relation to goals, challenge students to reflect more deeply, take into account how students worked on reflection tasks, take care of systematic reflection that leads to conclusions for future action.	T15 Assess the students' reflections	

In the fourth place, to gain insight into the most important competences we asked participants to rank both the thinking activities and the task domains in order of importance. Means of these rankings were determined (per thinking activity respectively per task domain).

In the fifth place, the consistency between questions, indicators, and tasks was examined by computing Cronbach's alphas. For every indicator we conducted analyses to determine whether the five questions per indicator formed a reliable scale. If this was the case the five items per indicator could be conjoined meaningfully to one scale (by taking the average) with which further analyses were done to establish whether also scales on task level and successively on task domain level could be formed. Minimum criteria for scalability were item-rest correlations $\geq .35$ and Cronbach's alpha $\geq .70$.

Finally, we tested whether succeeding rounds resulted in increased support and consensus (i.e. increasing mean scores and decreasing standard deviations of scales between rounds). We compared and tested means between rounds by performing pair wise *t*-tests on mean scores of scales. To check whether standard deviations decreased, we performed Levene's tests of homogeneity on standard deviations of scales. In addition, we analysed whether the groups of participants differed in their scores between both rounds. For this, we performed pair wise *t*-tests to compare mean scores of scales per group. To check whether the three groups differed significantly we performed Kruskal-Wallis tests on these scale scores.

At the end of the first round we asked participants to answer with yes or no three questions concerning overall relevance, completeness, and feasibility of the competences for teaching nursing students' reflection skills. We asked them to write down the time in minutes they needed per task domain to score the questions and to judge the effort to do this on a Likert-type scale, from 'very easy' (1) to 'very difficult' (5).

2.5 Results

The study participants (20 women and 8 men) ranged in age from 25 to 60 years. They judged the framework of teacher competences in two rounds.

Scale reliabilities

Both in round 1 and round 2 the scores on the questions per indicator for all 91 indicators met the criteria for scalability. For all 15 tasks the indicators per task also met the criteria for scalability and the same goes for the tasks per task domain. Subsequently the corresponding scales were constructed and their means

and standard deviations were computed. In the first round scale scores of 25-26 participants and in the second round scale scores of 25-27 participants could be computed (the slightly different numbers due to partial non-response). The range of the Cronbach's alphas of the indicators (I), tasks (T), and task domains (D) was round 1: I .75-.98, T .83-.96, D .87-.95 and round 2: I .83-.97, T .81-.97, D .86-.97.

Comparison per indicator

In the first round for 81 indicators all mean scores were $M \geq 5.5$, indicating a high degree of support on the five questions per indicator. For 10 indicators mean scores on the questions about unambiguous formulation and completeness were between $M = 5.0$ and $M = 5.4$. Standard deviations of the 91 indicators were between $SD = 0.65$ and $SD = 2.60$.

Most of the participants (80%) gave written comments and suggestions. In total 200 comments and suggestions were given, mostly about a certain word that was judged not clear or not concrete enough or too specific. This is in line with the lower mean scores on unambiguous formulation and completeness mentioned before. These results indicated that the framework itself (six task domains, 15 tasks and 91 indicators) could be maintained. Only aspects of the content needed some revisions. We decided to clarify nine tasks about which the participants gave most comments and suggestions by adding general context descriptions to these tasks. We made some indicators more explicit by adding text, for instance to "the teacher let students tell about how they reflect" we added "the teacher pays attention to constraints". To "students may modify or reflect on feedback" we added "within the borders of the requirements of the nursing profession". The resulting version of the framework was sent to participants for a second round.

In the second round for all indicators mean scores were $M \geq 5.5$. Consequently, support had improved. As the standard deviations of the 91 indicators were between $SD = .46$ and $SD = 2.22$ the consensus had improved.

Comparison per task domain

We compared the answers for the task domains between rounds 1 and 2 to the same five questions as used for the indicators. For the means, standard deviations and Cronbach's alphas per task domain in the first and the second round, see Table 2.3.

For all but one question about the task domains, the mean score increased from the first to the second round. The mean scores on the questions about unambiguous formulation and completeness were between $M = 5.0$ and $M = 5.4$ in round 1 and $M \geq 5.6$ in round 2. Pair wise *t*-tests revealed that for 11 of these 12 questions, these differences were statistically significant ($p < .05$). For the

questions about content relevance, practical relevance, and representativeness the mean scores were $M \geq 5.8$ in round 1 and $M \geq 6.2$ in round 2. Pair wise T-tests revealed that for two of these 18 questions the mean scores differed significantly between the rounds: for the question about the representativeness of the task domains 1 and 2 ($p < .05$). The standard deviations for 28 of the 30 questions strongly decreased. Levene's tests of homogeneity showed that for four of these 28 variables these differences were significant; namely for the questions about the completeness of task domain 3, 4 and 5 and about the unambiguous formulation of task domain 3 ($p < .05$).

Table 2.3

Means and standard deviations per question per task domain and reliabilities per task domain, in the two Delphi rounds

D ^f	R ^g	Relevant ^a		Unambi- gious ^b		Important for practice ^c		Represen- tative ^d		Complete ^e		α
		M	SD	M	SD	M	SD	M	SD	M	SD	
D1	R1	6.1	0.70	5.0	1.56	6.2	1.76	5.8	1.11	5.2	1.13	.79
	R2	6.2	0.62	5.6	1.28	6.4	0.63	6.2	0.99	6.2	1.06	.84
D2	R1	6.2	1.04	5.2	1.43	5.9	1.17	5.9	1.24	5.2	1.31	.89
	R2	6.2	0.75	5.6	1.26	6.3	0.63	6.2	0.78	5.7	1.22	.85
D3	R1	6.0	1.10	5.3	1.67	6.1	0.99	6.1	1.00	5.4	1.53	.96
	R2	6.2	0.75	5.9	0.70	6.4	0.65	6.3	0.79	5.9	1.13	.91
D4	R1	6.2	0.69	5.3	1.65	6.2	1.12	6.0	1.12	5.4	1.46	.88
	R2	6.3	0.67	5.8	1.10	6.4	0.69	6.3	0.73	5.8	1.04	.90
D5	R1	6.1	0.85	5.3	1.65	6.2	0.76	6.0	0.84	5.0	1.71	.80
	R2	6.2	0.75	5.6	1.20	6.3	0.68	6.2	0.88	5.7	1.23	.91
D6	R1	6.1	0.78	5.1	1.21	6.2	0.80	5.8	0.97	5.2	1.20	.88
	R2	6.3	0.68	5.8	1.37	6.3	0.61	6.2	0.78	5.9	0.69	.92

^{a, b, c, d, e} About each task domain five questions were asked concerning the extent to which the content is: ^a relevant for nursing education, ^b unambiguously formulated, ^c important for the nursing practice, ^d representative for nursing education, ^e completely formulated.

^f D = task domain.

^g R = round.

Rankings

Participants ranked both the thinking activities underlying reflection skills and the task domains. The eight thinking activities had to be ranked from ‘most important’ (1) to ‘least important’ (8). On average the thinking activity analysing took the first place (mean rank $M = 2.4$), followed by describing ($M = 3.3$), and explaining ($M = 3.4$). A position in between was for structuring ($M = 4.4$), concluding ($M = 4.6$), and evaluating ($M = 4.8$). Formulating intentions ($M = 5.4$), and attributing ($M = 6.3$) were judged as least important. Next, participants ranked the importance of the six task domains from ‘most important’ (1) to ‘least important’ (6). On average coaching students was seen as most important ($M = 2.8$), followed by developing education and giving instruction and assignments (both $M = 3.3$), giving feedback and setting goals (both $M = 3.7$), and lastly conducting reflective conferences and assessing reflections ($M = 4.2$).

Comparison of indicators and task domains on scale level

We used the scale scores on the five questions per indicator to establish the percentage of agreement. In the first round 67% of the participants received a score > 5.5 on the indicator scales. In the second round 93% of the participants did so. As consensus (75% agreement) was reached a third round was not necessary.

Comparing the results on the task domain scales between round 1 and round 2, average judgments appeared to have increased and average standard deviations appeared to have decreased. The differences between both rounds in mean scores on task domain scales have been tested for statistical significance using pair wise t -tests on the scores of the experts who participated in both rounds and of whom complete data were available (see Table 2.4).

Table 2.4

T-tests of differences in scale scores per task domain between the two rounds

Task domain	Round 1		Round 2		Pair wise T-test			Levene's test			
	M	SD	M	SD	t	df	p	F	$df1$	$df2$	p
D1	5.7	0.73	6.1	0.62	-2.97	23	.01	.98	1	50	.33
D2	5.6	1.00	6.1	0.73	-2.45	21	.02	1.77	1	48	.19
D3	5.8	0.82	6.2	0.62	-3.29	23	.00	4.52	1	50	.04
D4	5.9	0.75	6.1	0.65	-2.24	23	.04	1.67	1	50	.20
D5	5.7	0.83	6.1	0.62	-3.47	23	.00	.61	1	50	.44
D6	5.7	0.85	6.1	0.67	-3.80	23	.00	1.98	1	50	.17

For all task domains, the mean judgment in the second round was significantly higher than in the first round. Also, for every task domain the variance in the experts' judgments decreased. Only for task domain (D3) this increase was significant.

The average mean scores on the 30 task domain scales of the three groups of participants (teachers, managers and researchers) were in round 1 between $M = 4.9$ and $M = 6.1$ and in round 2 between $M = 5.9$ and $M = 6.6$. The standard deviations were in round 1 between $SD = .55$ and $SD = 1.20$) and in round 2 between $SD = .50$ and $SD = .95$). Kruskal-Wallis tests in both rounds revealed no statistical significant differences between the three groups.

Evaluation

The participants evaluated the framework of competence descriptions holistically in view of the importance for teaching nursing students' reflection skills. Most of the participants thought the descriptions to be complete (78%) and all important competences to be included (88%) and suitable for teaching practice (77%). The Delphi procedure itself was also evaluated. Participants needed 2-3 hours to read the content information and the instruction and to fill out the questionnaire. Answering the questions about the task domains 1 and 2 was judged as of some difficulty (score 4) and doing this about the other task domains as on average neither easy nor difficult (score 3).

2.6 Discussion

Study limitations

In the study statistical consensus has been used as an indication for validity. Although we selected a mixed group of participants from theory and practice (Goodman, 1987), the resulting framework of teacher competences is based on the judgment of a limited group of participants. Therefore, one limitation of our study is that the meaning of the framework of competence descriptions can only be defined relative to the participants in our study (Clayton, 1997; Gorney & Ness, 2000; Keeney et al., 2006). Additionally, we cannot exclude that some participants have gone along with the majority judgment (Greatorex & Dexter, 2000).

Further, the questionnaire used as input for the Delphi procedure will have influenced the results of the study. The participants judged a framework of competences with a certain content and structure. They were not invited to propose a totally different set or structure of competences. Some unintended distortion also may have occurred due to the researcher's own biases while

analysing the results after each round and adjusting the questionnaire (Williams & Webb, 1994). We tried to reduce these influences by a careful selection of competence descriptions based on studies of literature and national documents of nursing education.

Discussion of findings

In the first round of the Delphi procedure all participants strongly supported the indicators and task domains and a rather high mean score of $M \geq 5.0$ was reached, which in the second round increased to $M \geq 5.6$. For each indicator in the second the minimum percentage of agreement of 75% was attained (Murry & Hammons, 1995; Keeney et al., 2006; Williamson, 2007) and on average the standard deviations decreased between rounds; for one task domain this decrease was significant. So, after two rounds a high level of both support and consensus was reached and a third round was not necessary. This result is in line with the findings of Smith and Simpson (1995) and Keeney and colleagues (2006) who concluded that two rounds can be sufficient.

In view of the support of and consensus among participants, we succeeded in developing a framework of descriptions of teacher competences which forms a useful source of information both for nursing education practice and future research. The question arises what might have caused the positive results in the first round and the yet more positive results in the second round. We assume that our approach induced a certain ownership as a result of the information given beforehand about the Delphi procedure (Keeney et al., 2006).

The results on the ranking task for the thinking activities showed that the participants perceived analysing as the most important thinking activity in students' reflection process, followed by describing and explaining. Structuring, concluding and evaluating got a moderate score. Nevertheless, these thinking activities are also important in the reflection process. Students have to put things together, get insights in patterns and determine effects of actions in order to learn from it. The participants judged formulating intentions and attributing as the least important thinking activities in students' reflection process. This is in contrast with the recommendation that a reflection process be finished with decisions about future actions (Eraut, 1994; Schön, 1987) and with the importance of core reflection in which deeper layers of the person are involved (Korthagen & Vasalos, 2005).

Results of the ranking task for the task domains showed that coaching is the most important competence nursing teachers should practise. The importance attached to the role of coaching is in line with the cognitive apprenticeship model (Collins et al., 1989). In the ranking of importance the second task domain was that

teachers should develop and apply teacher competences to distinguish and stimulate thinking activities. This confirms the need of professional development of teachers to enable them to teach reflection skills and also the idea that developing reflection skills is a tough learning goal (Elshout-Mohr et al., 2002; Korthagen & Wubbels, 2001; LaBoskey, 1993). The task domains of feedback and goal setting got a moderate score. The participants seemed to attach less importance to the effect which feedback on goals can exert than is indicated by recent insights in the literature (Brinko, 1993; Chi et al., 2001; Gibbs & Simpson, 2005; Shute, 2008). The lowest scores for conducting reflection conferences and assessing reflections may indicate that these tasks are not common in practice. However, the participants perceived the framework of teacher competences as a whole as being sufficient, necessary, and suitable.

2.7 Conclusion

The framework of teacher competences we developed showed high consistency on three levels: indicators, tasks, and task domains. The framework gives a theoretical input to the description of teacher competences needed to support nursing students' reflection skills development, by detailed descriptions of teacher competences to improve education in supporting students' reflection skills and by concepts to be used to describe reflection skills in the form of thinking activities. Further research will be needed to get deeper insight in each thinking activity and in the patterns of thinking activities in students' reflection processes which especially contribute to their learning.

Practical implications of the study are that the framework of teacher competences to support students' reflection skills development can be a source to use in a teacher training programme for supporting reflection in nursing education. Although formulating intentions was seen as the least important thinking activity in students' reflection processes, it is very important for students to think about future improvements to gain from their reflection process. Teachers should be aware of this, because otherwise a reflection process can become an end in itself instead of a means for learning.

Appendix

Task domains, tasks, indicators, and questions in the instrument

I. Task domains, tasks and indicators

The task domains, tasks and indicators below are based on the final version after the second round of the Delphi-study. The descriptions of the indicators are summaries of the more extended descriptions (mostly 30-50 words) in the instrument. In the instrument used in the study, the thinking activities listed within task 2 (see below) were described once more within task 4, now for each task with a case and some concrete examples of questions which might be asked. There the thinking activities are only indicated with numbers: 21-28.

Before each indicator it can be read: 'The teacher:'

Task domain 1 Developing education for learning to reflect

Task 1 To develop competences for teaching reflection skills

1. develops her beliefs, keeps up with the professional literature, develops a vision on reflection education in collaboration with colleagues;³
2. works with colleagues on education that stimulates reflection skills in accordance with the school's educational policy;
3. develops her professional skills by reflecting, giving and receiving feedback and working on the development of her competences;
4. decides how to use her function as a role model to teach students how to reflect, and uses students' feedback to improve her own reflection skills;
5. checks and considers which reflection assignments and methods are fit for use in the different phases of the nursing education course.

Task 2 To get competent in distinguishing thinking activities

Distinguishes thinking activities which occur during reflection:

6. describing – giving a description of the core of an event or situation with was the main occasion to reflect;
7. analysing – posing oneself questions about aspects that play a role to become aware of what it is all about you reflect;
8. structuring – getting an overall view by discovering patterns and searching for relationships;
9. evaluating – determining effects and appraising these effects;
10. attributing – attributing certain aspects to oneself and/or others and giving them meaning for other situations;

³ Here and elsewhere for 'her' also 'his' might be read.

11. explaining – indicating reasons why and naming causes in connection with the own behaviour, thoughts and feelings;
12. concluding – drawing conclusions about thinking, behaviour and oneself in the future;
13. formulating intentions – setting new goals based upon what is learned by reflecting and how to apply what is learned.

Task 3 To create conditions for learning to reflect

14. creates conditions for a socially supporting learning environment and coaches students to their needs;
15. creates a climate for collaboration between students and stimulates students to think critically about oneself;
16. is a role model in reflecting and giving and receiving feedback;
17. creates a learning environment which triggers reflection and uses learning materials and methods which have been developed or selected by the team;
18. develops, in collaboration with colleagues, a system to observe and record students' progress.

Task 4 To stimulate thinking activities

19. models in front of the class how she reflects;
20. poses questions to stimulate students' thinking activities and lets students collaborate in exchanging such questions.
- 21-28: in the instrument used in the study for each thinking activity a short case was described which might be brought up by a student. Also several questions were added which a teacher could pose on occasion of such a case.

Task domain 2 Supporting students' reflection skills development

Task 5 To formulate short-term goals

29. demonstrates reflection as a conscious process to learn from events;
30. searches reflection assignments matching with students' learning needs to teach using the thinking activities;
31. supports students to formulate personal learning goals while they work on reflection assignments and participate in reflection conversations;
32. lets students reflect on their functioning in the learning group;
33. supports students in reflecting together on their education and internship;
34. reflects on her role in the group in relation to the students' learning goal.

Task 6 To formulate long-term learning goals

35. initiates and maintains a relationship of trust in which students gradually use reflection in an integrated way to discover new points of view;
36. supports students in view of their learning needs when they exercise the thinking activities;

37. demonstrates how she reflects on her own opinions and competences;
38. uses reflection goals for each phase in the nursing education course or specific learning path for reflection;
39. teaches students to think consciously about their behaviour, emotions, norms, and dilemmas in professional situations;
40. teaches students different methods and how to choose an appropriate method in the situation at hand.

Task domain 3 Giving students reflection instruction and assignments

Task 7 To give instruction about skills needed for reflection

41. knows which knowledge and skills students need for using the thinking activities;
42. knows and uses instruction strategies and pedagogical methods to teach students how to reflect;
43. shows the interplay between the knowledge, skills, values, norms, and opinions of the student, the group, the school, and the social environment;
44. clarifies reflection assignments, provides useful material and teaches skills;
45. explains what reflection is and exemplifies how to do it;
46. supports students by discussing the thinking activities and reflection objects.

Task 8 To give reflection assignments

47. gives a reflection assignment which fits the educational phase and students' level;
48. decides in consultation with the students on what and how they are going to reflect;
49. informs the students about what is expected with regard to the form and goal of their reflections and the learning environment;
50. coaches students individually and in the group in view of their learning questions.

Task domain 4 Coaching students to develop their reflection skills

Task 9 To care for guidance conditions

51. uses coaching strategies and pedagogical methods which fit the students' needs;
52. evaluates with the students whether their behaviour has been adequate and teaches them to evaluate their learning process;
53. encourages students by her reflective attitude (as a role model) to pose reflective questions;
54. supports students in going through the reflection process;
55. uses communicative skills and shows an accepting and empathic attitude;

56. lets students tell how they perform a reflection assignment and lets them verbalise their learning process. Gives attention to interferences and impediments.

Task 10 To coach students in their learning process

57. realises that her attitude is important, shows comprehension, has patience, and confronts students with uneasy aspects;
58. gives explanations, examples, and suggestions which might lead to improvement of students;
59. coaches students, aiming at diminishing the teacher's directive influence;
60. gives feedback, lets students give feedback to each other, and lets them learn from successes;
61. supports students' thinking about ways in which professional behaviour might be improved and afterwards to improve their own behaviour;
62. lets students assess themselves the execution of a reflection assignment with help of an assessment form, and uses this in a coaching conversation.

Task domain 5 Supporting students' development through feedback

Task 11 To give students feedback purposefully

63. feels how she gives feedback in a dialogue to strengthen the students' learning process;
64. gives feedback which is tuned to the learning goals, the competences to be developed, and the students;
65. gives students feedback on their functioning individually, in the learning group, and during their internship;
66. gives feedback on behaviour, which is specific, concrete and focused;
67. gives purposefully feedback, preferably when asked, and does not overfeed students;
68. asks students feedback on her teaching of reflection and uses this feedback.

Task 12 To teach students how to receive and ask for feedback

69. gives feedback that is sincere and is supported in non-verbal ways;
70. gives constructive critique and describes which behaviour is preferable;
71. explains how students might receive feedback;
72. shows how to take in feedback and lets students experience how to do this;
73. teaches students to ask for feedback and to react by asking questions and discussing.

Task 13 To support students' use of feedback

74. gives feedback preferably shortly after the event or if necessary at a later moment when the feedback has been considered;

75. gives students information on their functioning and gives them a certain degree of freedom to accept, modify or reject the feedback;
76. stimulates students to give her and each other feedback and lets them exercise;
77. advances a discussion on the task, the feedback and the standards in order to let them develop insight;
78. supports students in developing a mental picture of the desired quality by relating their results to principles of nursing and to the competence profile their education is aiming for;
79. stimulates students to monitor and evaluate their own performances.

Task domain 6 Holding reflection conferences and giving assessments

Task 14 To lead reflection conferences

80. uses a method for holding reflection conferences with students who participate actively in this;
81. has an open attitude towards students and supports them on a basis of equality as a person;
82. masters conversation skills, notes non-verbal signals, keeps posing questions, and stimulates thinking;
83. observes patterns in reflections and reports and brings these up for discussion without being denouncing;
84. discusses similarities and differences between theory and practice of professional behaviour;
85. holds an individual reflection conference in case students do not show professional behaviour;
86. teaches students to look while reflecting for their own role, and makes agreements.

Task 15 To assess students' reflections

87. gives students a view on the quality of reflections, both formatively and summatively;
88. assesses both the content and the process of reflection and gives feedback in an appropriate quantity;
89. observes the reflection on objects and thinking activities and gives feedback which challenges students;
90. determines whether reflection questions have been sufficiently answered; if not, she firstly poses questions about the content and secondly about the process;
91. determines whether students reflect systematically and execute their intentions; she gives them feedback on these issues and criticises them if necessary.

II. Questions in the instrument

In the Delphi-study an instrument was used in which questions were posed (a) per indicator, (b) per task domain, and (c) about all task domains together. Pertaining to (a) and (b) always the same five questions were posed, to be answered on a 7-point scale from 1=not to 7=optimally: “To what extent the indicator respectively task domain is: (1) relevant pertaining to the content? (2) unambiguously formulated? (3) important in teaching practice? (4) representative? (5) completely formulated?” For the task domains always an open question was added: “What can be changed and how?”, as well as some space to write comments and additions. With respect to (c) the six task domains were listed and participants were asked to rank them in order of importance for supporting the development of students’ reflection skills, by using the letters A (most important) to F (least important).

3 Performance standards for teachers supporting nursing students' reflection skills development⁴

Abstract

In this study performance standards were developed and validated for nursing teachers who have to support students' reflection skills development. A ten-step procedure was followed to ensure procedural and internal validity. National competence descriptions and specific content standards for supporting nursing students' reflection skills were used in a preliminary rubrics framework which was piloted. Next, forty participants from six nursing schools judged this framework of eight competences covering thirty rubric attributes. They also discussed the prerequisite minimum performance level and the judgmental model to be preferred. These judgments and discussions resulted in consensus on the rubrics, a general cut-off score, and a conjunctive judgmental model convenient for assessing nursing teachers' competences. The rubrics might be used in a teacher training program. Also institutes of nursing education can employ the rubrics as a tool for teaching and formatively assessing students' reflection skills.

3.1 Introduction

Health care professionals and teachers are continuously facing changes in working practices and technologies. To respond to these changes they are given new roles and need to develop new competences (Biesma et al., 2008; Harden & Crosby, 2000). Consequently, students as future professionals have to be equipped with skills for lifelong learning, of which reflection is an important one (Govaerts, 2008; Mann, Gordon & MacLeod, 2009; Platzer & Snelling, 1997; Schön, 1991). Reflection can be seen as a meta-cognitive and affective skill to be employed as a means for developing complex competences, both in school and in the workplace. It implies a process of thinking about and interpreting situations, events, experiences and emotions, aiming at critically analysing decisions, actions and effects in order to learn from them (Ertmer & Newby, 1996; Mann et al., 2009; Tanner, 2006).

⁴ A shorter version of this chapter has been published under the same title in *Journal of Nursing Education and Practice* 2 (1), 9-19.

Reflection skills are difficult to develop (Korthagen & Wubbels, 2001) and students need support for acquiring and improving these skills (O'Donovan, 2007). This support is possible in a powerful learning environment where teachers flexibly adapt their instruction to students' individual needs and possibilities (De Corte, 2000). Giving students tailor-made support is part of the cognitive apprenticeship model (Collins, Brown, & Newman, 1989). In this model, by articulating, scaffolding, modeling, coaching and fading teachers stimulate the students steadily in steering their own learning process to become competent.

For acquiring reflection skills coaching and constructive feedback turn out to be essential (Ericsson, 2004). Feedback is information communicated to students for providing changes in thinking or performance (Shute, 2008). It should be used for closing the gap between current performance and certain standards and therefore it should be based on those standards (Sadler, 1989, 2010). Similarly to students their teachers also need feedback to acquire and develop competences for supporting the development of their students' reflection skills. This feedback has to inform the teachers about the performance desired from them. International attention for monitoring teacher performances and improving teacher quality has led to the development of teacher standards (European Commission Education & Training, 2011; NBPTS, 2011). These standards, however, are not yet specified regarding supporting students' reflection.

The purpose of the study was to develop and validate performance standards which are relevant and acceptable for nursing teachers to work with in everyday practice in which they have to support students' reflection skills development.

The research questions are: What are relevant and acceptable performance standards for teachers who have to support nursing students' development of reflection skills? Which judgmental model is, according to nursing teacher teams, convenient for assessing the teachers' competences?

3.2 Background

Standard-setting

Teacher competences pertain to the integrated knowledge, skills and attitudes needed to perform their professional tasks. Consequently, the feedback to be given to teachers should aim at the competences to be developed and should inform them about the extent to which they are expected to meet certain standards. Standards for teacher competences include content standards and performance standards. Content standards specify what teachers are expected to know and be able to do (Hambleton & Pitoniak, 2006). Performance standards

specify norms for the frequency, intensity or quality on which these performance aspects will be scored while assessing teachers (Boursicot, Roberts & Pell, 2006). For the development of performance standards many methods are available. Most methods are either norm-referenced or criterion-referenced. In norm-referenced methods the performance is scored and compared to the performance of a norm group, setting the standards so that a certain percentage of the examinees pass (Hambleton & Pitoniak, 2006). In criterion-referenced methods the performance is compared with a certain quality standard specified beforehand, and obviously this is the kind of performance standards we need. Norm-referenced methods are not feasible because in that case scores are interpreted with respect to being better or worse than other teachers instead of in terms of the necessary level of competence as used in criterion-referenced methods (Sireci, 2005). Criterion-referenced methods for standard setting are traditionally placed on a continuum from test-centered to examinee-centered. An example of a test-centered method is the Angoff method in which a panel estimates whether a just-qualified person would answer an item correctly or not (Hambleton & Pitoniak, 2006). An example of an examinee-centered method is Judgement Policy Capturing in which a panel reviews score profiles across exercises and assigns each score profile to a proficiency category (Hambleton & Pitoniak, 2006).

Currently performance standards are increasingly based on judgments of a group of selected experts about performance levels and cut-off scores. A cut-off score indicates a minimally required or 'just sufficient' performance level (Berk, 1996; Hambleton & Pitoniak, 2006; Kane, 1994). A convenient standard-setting method for nursing teachers supporting students' reflection skills is using rubrics to describe and judge standards. A rubric consists of parts or aspects of the work that constitute its quality together with several performance level descriptions regarding each particular task (Arter & Chappuis, 2007; Hafner & Hafner, 2003).

Judgmental model

In setting performance standards not only a standard-setting method should be chosen but also a judgmental model. A judgmental model is needed for making a final judgment about the performance based on the individual scores for each task on each standard. To combine these scores, several judgmental models can be applied: a compensatory model, a conjunctive model, or a disjunctive model (Berk, 1996; Hambleton & Pitoniak, 2006). In a compensatory model the total score obtained on the different standards determines whether the overall performance is sufficient. In this model high scores on certain standards can compensate for low scores on other standards. In a conjunctive model a certain minimal score has to be reached on each standard. In a disjunctive model for certain standards a

minimal score is mandatory and for other standards low scores are allowed if high scores on other standards compensate enough to reach the minimum total score.

Validation strategies

Rubrics have to meet certain quality requirements, of which the most important is validity. In strategies for validating assessment instruments and procedures different kinds of evidence play a role: procedural, internal, and external validity evidence (Kane, 1994). Procedural evidence consists of demonstrating the soundness of a procedure to develop the rubrics. To warrant procedural validity the standard-setting procedure has to be clearly and explicitly defined and documented. Its implementation should be easy and systematic. Moreover judges should know the purpose for which the standards are being set and understand and feel comfortable with the standard setting procedure they are using (Hambleton & Pitoniak, 2006). Internal evidence focuses on the consistency of the judges in translating ratings on performance standards in a cut-off score. To warrant internal validity, the rubrics have to be representative of the target domain. Also, the rubrics have to be clear to the judges (Kane, 1994) who rate the rubrics independently (Berk, 1996) and reach consensus about the judgments (Hambleton & Pitoniak, 2006). External evidence can be provided by comparing the resulting performance standards with other sources of information, such as decisions based on performance assessments in authentic situations or in simulations.

As teaching and assessing reflection skills is rather new in nursing education and external information is not yet available, this study is restricted to procedural and internal validation strategies. Developing rubrics as well as the standard-setting itself are commonly based on standard-setting procedures that provide procedural and internal evidence.

Development of a comprehensive standard-setting procedure

In this study we combine three approaches for standard-setting. Berk (1996) recommended a generic method for performance standard-setting. Hambleton and Pitoniak (2006) explicitly considered the selection of a performance standard setting process and discussed additional steps which may be essential depending on the method chosen. Kane (1994) distinguished criteria for the implementation of standard-setting procedures. While the terminology these authors used differ the activities they describe show a considerable overlap. The activities they have in common are activities that according to Hambleton and Pitoniak (2006), when given careful attention, can substantially increase the defensibility of the standards: the selection of the method, the selection and training of the judges, the

consecutive activities from defining standards to computing cut-off scores, and the validation and documentation of the total process.

Striving for a procedure as complete and unambiguous as possible we combined these three ways of standard setting while sometimes adapting the terminology for clearness. Based on this a standard-setting procedure including ten activities was developed (see Table 3.1). This study is composed according to these ten activities.

Table 3.1

Procedure for performance standard-setting

Three ways of performance standard-setting (Berk, 1996 (B); Hambleton & Pitoniak, 2006 (H&P); Kane, 1994 (K)) were combined to the standard-setting procedure used in the study. In this procedure ten activities (see below) are distinguished to provide procedural and internal validity evidence. The sources used are noted within the parentheses.
<ol style="list-style-type: none">1. Select a method for the standard-setting procedure (H&P)2. Define goals for the standard-setting procedure (K)3. Prepare performance descriptions: specify types of performance that are expected (B, H&P, K) and the achievement levels (B)4. Select judges who represent the target group (B, H&P, K)5. Train the judges: inform them and get them acquainted to the procedure (B, H&P, K)6. Process of independent judgment: let the participants judge (B, K) meanwhile providing feedback to the judges and encourage discussions (H&P, K) to get consensus (B) on appropriate and clearly stated definitions (K)7. Compute the cut-off scores based on the judges' ratings and decide about these scores (B, H&P)8. Select a judgmental model (B)9. Evaluate the standard-setting procedure with judges (H&P)10. Check validity (H&P, K) and document the findings (H&P)

3.3 Development of the rubrics framework

For the method to be used in the standard-setting procedure (activity 1 in Table 3.1) we selected rubrics as instrument because rubrics have proven to be convenient for developing performance standards that are feasible for teachers in professional practice who have to support students' reflections (activity 2) (Lane & Stone, 2006). Firstly, a rubrics framework was developed by integrating Dutch national teacher competence descriptions and specific content standards for teachers supporting nursing students' reflection skills development.

The national teacher competence descriptions were chosen because of their legal status in the Netherlands (Ministry of Education, Culture and Science, 2007).

These competence descriptions represent seven general competence domains: (a) interpersonal domain with a focus on effective communication, (b) pedagogical domain of caring for a positive learning climate, (c) subject matter and pedagogics domain concerning instruction, guidance, feedback, and assessment, in this case of reflection skills, (d) organisational domain including the selection of goals and tasks and the use of a system for observation, assessment and registration, (e) cooperation with the colleagues within the team, (f) cooperation with the environment, for example about the students' internship or national curriculum development materials, (g) the teacher's own professional development. In this study the competence domains e, f and g were combined into one competence domain called professional development because these three domains together can be seen as a means for this.

Specific content standards for nursing teachers who have to support students' reflection skills development have been developed and validated in an earlier Delphi study (Dekker-Groen, Van der Schaaf & Stokking, 2011; see chapter 2). These standards have been arranged into six task domains. Analyses showed that the first task domain, preparation of reflection instruction, partly corresponds with both the organisational (d) and the professional development (e, f, g) competence domain. The second task domain, learning goals, links to both the interpersonal (a) and pedagogical (b) competence domain. The other four task domains: instruction, coaching, feedback, and assessment, correspond with the subject matter and pedagogics (c) competence domain in which they represent four separate competences. As a result, for this study eight competences were established within five competence domains, namely: (1) Interpersonal, (2) Pedagogical, (3) Subject matter and pedagogics (four competences), (4) Organisational and (5) Professional development competence domain.

The six task domains in the earlier Delphi study covered 91 different content standards (see the Appendix to chapter 2). Two researchers independently allocated these 91 content standards to one of the eight competences mentioned before. Both researchers could successfully allocate 72 content standards. After discussing the remaining 19 content standards, six content standards were left out because they were specifications of one or more of these 72 content standards. Due to some overlap the other 13 content standards could be combined with the already allocated content standards.

Consequently, 72 separate content standards from the content standards study could be used for performance descriptions in the rubrics framework, which include the eight competences within five competence domains. Per competence two to five rubric attributes could be distinguished, adding up to a total of 30. For all these 30 rubric attributes teacher performance descriptions on three

performance levels were developed (activity 3). See Table 3.2 for the rubrics framework.

Table 3.2
Rubrics framework for setting performance standards

Five competence domains	Eight competences	30 rubric attributes ^a	Content standards ^b 72 of 91 included	α ^c
Interpersonal	1. Effective communication with individual and group	4	7	0.88
Pedagogical	2. Care for a positive learning climate	3	6	0.76
Subject matter and pedagogics	3. Give information and instruction how to reflect	5	26	0.95
	4. Coach the learning process of how to reflect	4	6	0.90
	5. Give feedback, and teach how to receive and ask for feedback	5	11	0.91
	6. Use criteria to assess reflections	4	6	0.90
Organisational	7. Design and organise the learning environment	2	3	0.87
Professional development	8. Work individually and in cooperation with colleagues	3	7	0.86

^a Each rubric attribute got four performance level descriptions. ^b See Dekker-Groen, Van der Schaaf & Stokking (2011). ^c Reliability of allocation of the content standards per competence.

The quality of allocating 72 content standards to eight competences was checked by analysing whether for each competence the questions asked about the content standards in the previous Delphi study (see chapter 2) which were allocated now to that competence do form together a reliable scale. See Table 3.2 for results. Minimum criteria for scalability were item-rest correlations ≥ 0.35 and Cronbach's alpha ≥ 0.70 . For all competences the accompanying content standards except one (item-rest correlation = 0.34) met these criteria of scalability. So, allocation of content standards to a preliminary rubrics framework of performance standards has been successful.

Thirdly, a pilot study was done. Participants were two researchers, two teacher educators, two team managers and two workplace coaches, all with experience in teaching nursing students reflecting skills in a traditional as well as in a competence-based nursing curriculum. We asked these experts: whether the preliminary rubrics framework sufficiently covered the content standards, whether the rubrics are unambiguously formulated and feasible in practice, and whether three performance levels are adequate or four levels are preferable. They judged the rubrics to be clear, representative of the content standards and feasible in practice. They preferred a rubrics framework with four in stead of three performance levels because it makes more differentiated assessments possible.

As including four performance levels would result in a framework covering a broader continuum for teacher development, rubrics with four performance levels were developed. Also to each of the eight competences a general description was added. See Table 3.3 for an example of a rubric for one of the competences. These adjustments resulted in the first version of the rubrics framework which was input for the empirical performance standard-setting study.

Table 3.3
Rubric of competence 3 (first version)

Give information and instruction how to reflect				
General description: The teacher distinguishes between the objects and thinking activities of reflection which can be applied. Objects of reflection are: theory, practice and person (oneself). Thinking activities are: describing, analysing, structuring, evaluating, attributing, explaining, concluding, formulating intentions. She ^a supports students by discussing, explaining, illustrating and demonstrating objects and thinking activities of reflection. She stimulates discussions between students and concludes together with them. In addition she pays attention to similarities and differences between theory and practice and the interplay between knowledge, skills and beliefs of students and the learning environment. She teaches students to apply thinking activities. She gives students reflection tasks and suitable material and asks questions to stimulate the use of thinking activities. She teaches students to ask each other and themselves such questions.				
Rubric attributes ↓	Four performance levels			
	Beginning	Developing	Proficient	Competent
Objects of reflection	Does not distinguish objects	Distinguishes theory, practice, self	Teaches students to distinguish	Lets students distinguish objects
Thinking activities	Does not distinguish thinking activities	Distinguishes (eight) thinking activities	Teaches students to distinguish thinking	Lets students apply thinking activities
Examples, demonstration how to reflect	Does not explain and does not give examples	Explains and gives examples	Demonstrates, stimulates discussion	Lets students lead discussions, concludes
Questions to stimulate reflection	Does not ask questions	Asks stimulating questions	Teaches students to ask each other questions	Teaches students to ask themselves questions
Reflection tasks	Does not give tasks	Gives stimulating tasks	Elucidates tasks	Suggests study tasks and material

^a Where 'she' is used also 'he' can be read

3.4 The study

Participants

Six nursing teams voluntarily participated in the rubric procedure (activity 4). These teams cooperate with the Dutch Foundation Consortium Vocational Education, which develops methods, materials and instruments for teachers to support students' growth in competence (SCB, 2007). The participants were 33 teachers, three workplace coaches and four team managers (10 men and 30 women) who ranged in age from 26 to 63 years. The teachers had on average 11 years of work experience in nursing education.

Design and procedure

Per team three meetings of 90 minutes were conducted within one year. These meetings were audio taped to be able to gather procedural and internal evidence for the validity of the rubric procedure (activity 10). Procedural validity evidence is concerned with the question whether participants understood the rubric procedure and perceived this procedure as feasible in the available time, and whether they set performance standards. Internal validity evidence pertains to the consensus reached between the participants on the content of the rubrics framework and on the judgmental model.

In the first meeting with each team the participants were informed about the development process of the rubrics. Also the goal and the method of the standard-setting procedure were explained (activity 5). At the end of the first meeting the teachers decided about their voluntary participation in the study. Next, each participant received the first version of the rubrics framework by email. In the second team meeting the rubrics framework was explained and the performance levels descriptions were illustrated. The participants could freely ask questions about the framework (activity 5). Then, to get acquainted with the rubrics, they used the performance descriptions to self-assess their competences. Next, the participants discussed the performance descriptions and the labels for the performance levels in view of their relevance and acceptability. Further, suggestions and revisions deemed necessary by the teams already visited before, were brought in if they were not mentioned by the present team. The purpose was to obtain clear and acceptable rubrics and labels on which consensus could be reached within and between teams (activity 9). This approach was expected to contribute to the internal validity of the procedure (activity 10). Finally, participants determined a cut-off score per rubric attribute independently from

other participants by scoring the performance level they would consider as a “just sufficient” for a starting teacher (activity 6). After this meeting the mean cut-off scores were computed (activity 7).

One team was unable to participate in the third meeting due to unexpected heavy workload in their organisation. Therefore the third meeting was held in five teams. Each team was informed about the different suggestions and revisions which had been contributed to the second version of the rubrics. Then the team was asked whether they could agree with this second version. Further, each team was informed about their own mean cut-off scores per rubric attribute and these scores were compared with the mean cut-off scores of all teams together. In addition, the team was informed about the different judgmental models they could apply and their essential differences (activity 4). After that, each team discussed which judgmental model the team members assessed as most suitable (activity 8). This discussion lasted until the team reached consensus. Finally, the consensus on the preferred cut-off scores and on the judgmental model was member-checked by asking whether each participant supported the choices made (activity 9). All meetings were audio-taped.

Data analysis

The audio tapes were used to analyze afterwards whether participants indicated that the rubric procedure was clear and that the time schedule was attainable for each team. Also a check was done whether each participant determined a cut-off score per attribute on paper. The audio tapes were also used to check the suggestions and revisions made by the teams about the rubric descriptions and the labels and whether consensus had been reached (activity 10).

The participants’ judgments of the cut-off scores were analysed for the degree of consensus on the performance standards. Per rubric attribute the means and standard deviations were computed of the cut-off scores of all participants. Then the degree of consensus was determined between the participants on minimal acceptable performance levels (activity 7). In several Delphi studies for statistical consensus (percentage of agreement) a minimum percentage of agreement of 75% between experts has been judged to be adequate (Murry & Hammons, 1995; Williamson, 2000; Keeney et al., 2006), and it was decided to use this criterion also in this study.

Next, before starting the discussion about cut-off scores within the teams, it was checked whether significant differences existed between the teams in their judgments on the cut-off scores. Per rubric attribute the mean cut-off scores and standard deviations were computed of all participants per team. Univariate

analyses of variance were used to determine whether the teams differed significantly in their mean cut-off scores and $p = .05$ was considered to indicate a statistically significant difference. Consistency between and within the teams is expected to contribute to the internal validity evidence for the performance standards setting procedure (activity 10).

Finally, reliability analyses were conducted to determine whether the participants' cut-off scores per rubric attribute formed a reliable scale per competence. The consistency between the rubric attributes per competence was examined for all participants together by computing item-rest correlations and Cronbach's alphas. Minimum criteria for scalability were item-rest correlations ≥ 0.35 and Cronbach's alpha ≥ 0.70 . Consistency between rubric attributes per competence also contributes to the internal validity evidence (activity 10).

3.5 Results

Procedural and internal validation

Procedural validity evidence was obtained as participants had no questions about or comments on the procedure and this procedure was judged as attainable within the three meetings arranged within one school year. Also, each participant judged the performance standards and determined cut-off scores.

In the second and third meetings per team the rubrics framework was discussed. In the second meeting the first version of the rubric descriptions appeared to be fully clear to the six teams. They agreed to a high extent with the proposed descriptions of the performance levels of the rubrics. The teams had no comments on the descriptions of the competences 'coaching' (competence 4) and 'feedback' (competence 5). For the other six competences they suggested minor changes. All suggestions were used to develop the second version of the rubrics framework.

Two of the four labels of the performance levels generated some confusion. All teams thought that the performance label 'Developing' for the second level could not be used as a level because in vocational education the focus on lifelong learning implies that development is a permanent process. The teams chose 'Basic' as the new label. In addition, in every team some participants perceived the label 'Competent' of the fourth performance level as incorrect because competent cannot be the highest possible performance level. After a discussion about possible alternatives, agreement was reached in all teams about 'Expert' as the new label for the highest performance level. In each of the six teams it was unanimously decided that the labels for the first and the third performance levels needed no changes. A discussion was held about how performance levels should

be established. Teachers stated that their approach to the students depends on the grade in which they are. The general opinion in the teams was that a teacher who is able to perform on a certain performance level can also be expected to have the knowledge and skills for performing in accordance with the descriptions on a lower level.

In the third meetings each participant agreed with the second version of the framework. The teams concluded that this version contained unambiguously formulated rubrics which are relevant and acceptable for teachers who have to support nursing students' reflection skills. Internal validity evidence was obtained because consensus was reached between and within the teams.

Cut-off scores

Scoring on the performance levels of the rubrics appeared to be feasible for the participants. They individually scored for all 30 rubric attributes which performance level they would consider as 'just sufficient'. For all 30 rubric attributes most of the individual cut-off scores were on performance level 2. In four teams the mean cut-off scores were ≤ 2.5 for most attributes (87% till 93%). In the fifth team this was 50% and in the sixth (workplace) team 33%, due to several level 4 scores of one participant in the team. The mean cut-off scores of all participants together on all rubric attributes varied between $M = 2.1$ to $M = 2.6$. For an overview of the results, see Table 3.4.

The lowest mean score ($M = 2.1$) was obtained on four attributes within three of the competences about Subject matter and pedagogics: 'distinguish and discuss reflection objects' (competence 3: Instruction), 'discuss patterns' (competence 5: Feedback), and 'support broad reflection' respectively 'support deep reflection' (competence 6: Assessment). The highest mean score ($M = 2.6$) was obtained on the rubric attributes about applying verbal respectively non-verbal skills (competence 1: Effective communication).

Five teams discussed the mean cut-off scores of their own team and compared these with the mean cut-off scores of the other teams. For most of the rubric attributes the mean cut-off scores per team were between $M = 2.0$ and $M = 2.7$. On three of the 30 attributes the mean cut-off scores differed significantly between the teams. This was the case with the attribute 'apply verbal skills' of competence 1: $F(5) = 3.2, p = .02$, the attribute 'distinguish thinking activities' of competence 3, $F(5) = 3.6, p = .01$, and the attribute 'discuss patterns' of competence 5, $F(5) = 2.8, p = .03$. Effect sizes, calculated using η^2 , were respectively .05, .03 and .09. Post-hoc comparison using the Bonferroni test indicated that the mean scores were only significantly different on the attribute 'distinguish thinking activities' ($p < .05, p = .03, p = .01$).

Table 3.4

Competence scale reliabilities, Means and SD's of the cut-off scores per attribute

Competence scales		Cut-off scores per rubric attribute		
Competence	<i>a</i>	Rubric attribute	<i>M</i>	<i>SD</i>
1 Effective communication with individual and group	0.84	1 apply verbal skills	2.6	0.6
		2 apply nonverbal skills	2.6	0.7
		3 discuss students' functioning	2.4	0.6
		4 evaluate students' functioning	2.3	0.5
2 Care for a positive learning climate	0.70	5 build a pedagogical relation	2.4	0.5
		6 attune the target group	2.5	0.5
		7 coach empathically	2.5	0.8
3 Give information and instruction how to reflect	0.86	8 distinguish and discuss reflection objects	2.1	0.5
		9 distinguish and apply thinking activities	2.2	0.7
		10 explain examples, demonstrate how to reflect	2.3	0.5
		11 ask questions to stimulate reflection	2.4	0.7
		12 give reflection tasks	2.2	0.7
4 Coach the learning process of how to reflect	0.88	13 give guidance	2.4	0.7
		14 use methodology	2.4	0.7
		15 teach how to develop an attitude	2.5	0.6
		16 coach students	2.3	0.6
5 Give feedback, and teach how to receive and ask for feedback	0.90	17 give feedback	2.3	0.7
		18 receive feedback	2.4	0.8
		19 ask feedback	2.4	0.6
		20 discuss patterns	2.1	0.5
		21 make relations	2.2	0.5
6 Use criteria to assess reflections	0.86	22 support broad reflection	2.1	0.5
		23 support deep reflection	2.1	0.6
		24 support systematic reflection	2.5	0.7
		25 discuss constraints	2.2	0.6
7 Design and organise the learning environment	0.94	26 care for materials	2.3	0.6
		27 work systematically	2.3	0.6
8 Work individually and in cooperation with colleagues	0.87	28 cooperate with colleagues	2.3	0.5
		29 professionalise	2.4	0.6
		30 be a role model	2.3	0.6

Choice of a judgmental model

The five teams searched for and discussed rubric attributes for which mutual compensating within the same competence might be possible. No rubric attribute could be found for which the teams thought the performance level 1 was enough. The discussions per team led to consensus on level 2 as the 'just sufficient' performance level for each rubric attribute within each of the eight competences. The teams decided to apply a conjunctive judgmental model for each competence.

Scale reliabilities

Using the individually chosen cut-off scores of the participants, scale analyses showed that for all eight competences all rubric attributes per competence had item-rest correlations ≥ 0.35 and that all Cronbach's alphas were ≥ 0.70 , mostly between 0.84 and 0.88. The lowest alpha was 0.70 (competence 2) and the highest 0.94 (competence 7). See Table 3.4.

3.6 Discussion

All 10 activities of the developed standard setting procedure could be executed as planned. The choice of using rubrics (activity 1) has shown to be effective because it stimulated discussions within the teams about the relevance and acceptability of the competence descriptions and the rubric attributes. The developed rubrics framework of performance descriptions (activity 3) appeared to be relevant and acceptable for all six teams. The participants formed a panel of representative judges of senior secondary nursing education (activity 4). These judges received information about the rubrics, the standard-setting procedure and the judgmental models and were trained in executing the activities (activity 5).

The participants independently of each other decided about cut-off scores on each of the 30 rubric attributes (activity 6). The mean cut-off scores per team were mostly between 2.0 and 2.5 (activity 7). More extreme scores of four participants resulted in a significant difference on only one attribute within one competence (competence 3). These findings show a high amount of consistency between the cut-off scores of the teams. This consistency in judgments contributes to the internal validity evidence (activity 10).

Discussions were held until consensus was reached on the judgmental model (activity 8). The teams decided for all eight competences that the minimum performance level would be level 2 for each rubric attribute. So the teams chose a conjunctive model for each competence. Moreover, between all participating teams full agreement was obtained.

During the standard-setting procedure activities were evaluated and documented (activities 9 and 10). The procedure developed for setting performance standards in ten activities appeared to be useful according to the participants and has led to the intended outcomes: performance standards and a judgmental model convenient for assessing teachers' competences to support nursing students' development of reflection skills. These findings contribute to the procedural validity evidence (activity 10).

Additional indications for the quality of the content of the rubrics framework were that about 80% of the comprehensive set of content standards developed in an earlier study was represented in the framework, and that, after these standards had been allocated to the eight competences distinguished for the present study, the participants' scores on the rubric attributes per competence for each competence did form a reliable scale.

The teams discussed about how to establish the right performance level because a teacher adapts his or her approach depending on the students. Arguments were that in the lower grades a teacher structures a lot for the students while in higher grades a teacher preferably takes action only if necessary and let students monitor their own learning process. These arguments refer to teachers providing for cognitive apprenticeship (Collins et al., 1989) and a powerful learning environment (De Corte, 2000). Using these approaches a teacher gives tailor made support and makes deliberate choices which are seen as adequate for the target group or in a certain situation.

The conjunctive judgmental model chosen by the participants asks for analytical scoring in which specific attention is given to each rubric attribute. It also implies that teacher development and teacher competence assessment should focus on each separate attribute (Morrison et al., 2004). The decision to use the same cut-off score for all 30 rubric attributes might be an indication that holistic scoring will also be possible. Analytic scoring compared to holistic scoring enables more authentic situations to be assessed (Cadman Slater & Boulet, 2001). Also, feedback on the performance with respect to each rubric attribute offers a 'profile' of specific strengths and weaknesses (Mertler, 2001).

This study has been executed in agreement with the three design principles of Morrison and colleagues (2004). The first principle is to develop a concise instrument. We developed a framework of eight competences with 30 attributes, making it possible for teachers to quickly examine each competence. The second principle is to make attention possible for separate observable tasks. As the teams did not want to compensate between separate scores they preferred analytical scoring per task (that is, per attribute). When analytic and task specific descriptions and also exemplars are given and/or training is provided the performance per task

can be reliably scored ((Jonsson & Svingby, 2007). The third principle is that an instrument should be relevant and acceptable for use in practice and all participating teams have concluded this to be the case.

These three design principles also legitimate our choice for an instrument in the form of a set of rubrics (Hambleton & Pitoniak, 2006). Rubrics are suitable for qualitative judgements of authentic or complex performances and for assessment for learning (Jonsson & Svingby, 2007). This makes it possible to give feedback that can be used in further competence development.

A limitation in the study was that the meetings for the six teams were held subsequently. Changes deemed necessary by a team were discussed in the other teams, to take their bearings into account and come to performance levels that are appropriate, unambiguously formulated, and clearly distinguishable. Although we gave all teams the same information in the same sequence, the specific sequence used might have influenced the comments of other teams later on.

3.7 Conclusion

The research questions were: What are relevant and acceptable performance standards for teachers who have to support nursing students' development of reflection skills? Which judgmental model is, according to nursing teacher teams, convenient for assessing the teachers' competences?

Comparing the general Dutch national teacher competence descriptions and a set of more specific content standards for nursing teachers developed and validated in an earlier Delphi-study (see chapter 2), a framework of eight competences was chosen for which rubrics were developed. Within the eight rubrics a total of 30 attributes could be distinguished, covering almost all content standards from the Delphi-study. The teams participating in the current study judged the final version of these rubrics to be relevant and acceptable for teachers who have to support nursing students' reflection skills. The teams preferred using four performance levels (defining a scale from 1 to 4) for each rubric: beginning, basic, proficient, expert, because using four levels covers a broad continuum for teacher development. After all participants had scored for all 30 attributes which performance level they would consider as 'just sufficient' (the cut-off score), on all attributes the mean cut-off score was between 'beginning' and half-way 'beginning' and 'basic'. The teams searched for rubric attributes for which mutual compensating within the same competence might be possible. They found no attribute for which they thought performance level 1 would be enough. The discussions per team in all teams led to consensus on level 2 as the 'just sufficient'

performance level for each rubric attribute within each of the eight competences. Therefore, the teams decided to apply a conjunctive judgmental model for each competence.

In this study, rubrics have shown to be convenient for developing and validating performance standards. Procedural and internal validity evidence both justify the conclusion that the procedure used in this study including the choice for an instrument in the form of a set of rubrics has contributed to consensus within and between the participating teacher teams in nursing education.

This study generates practical implications, since teacher teams perceived the rubrics as relevant and acceptable for supporting nursing students' reflection skills development. So these rubrics might be used in teacher training on giving nursing students feedback to support their reflection skills development.

Further research can be done to explore whether the rubrics can be employed to prepare and formatively assess reflection conversations, and to monitor and assess how teachers develop their competences for supporting students' reflection skills development. Research on the use of rubrics in nursing education may give insights into whether analytical scoring is really suitable to teachers. Also, research might focus on the generalization of the rubrics to other domains, e.g. social work and teacher education. Future studies might also provide external validity evidence for the rubrics.

4 A teacher competence development programme for supporting students' reflection skills⁵

Abstract

This study aimed to evaluate a training programme for Dutch teachers in nursing education to support students' reflection skills development. The research question was: What are the feasibility, quality and effects of the programme? The training programme focussed on teachers' competences regarding instructing, guiding, giving feedback on, and assessing students' reflection skills. The programme development was informed by research-based features of effective teacher professional development. Concerning the content presented to and discussed with the teachers and the feedback given to them, specific expertise and knowledge was brought in by the trainer.

Twenty-seven teachers of six Dutch institutes for nursing education participated in the study. Nineteen teachers took part in the training programme and eight teachers formed a control group. Before the programme data were gathered concerning teachers' initial knowledge, experience, beliefs, and practices, and their self-assessment on four key competences. During the programme data were gathered about participants' appraisal of the content and activities of the programme and their learning concerns, activities, and perceived results of the feedback received. After the programme additional data on the quality and effects of the programme were gathered by way of teachers' appraisals of a number of aspects of the programme, their assessment of their own development as a result of the programme, their current teaching activities and the importance they attach to these activities, and repeated measurement of certain practices and beliefs and self-assessment of their competences.

The study showed that teachers' appraisal of the training programme was fairly positive. They reported to have developed their knowledge, skills and attitudes for supporting students' reflection skills development as well as different aspects of their teaching practice. Comparisons between training group and control group revealed that, according to their teachers, the students in the training group reflected on more diverse objects of reflection after the training than before and

⁵ A shorter version of this chapter is accepted for publication in *Teachers and Teaching: Theory and Practice*, in April, 2013. The iFirst version has been made available on the website of the publisher in November, 2012.

also more in comparison to the students in the control group. With regard to their self-reported activities after the programme, the teachers in the training group did score higher than those in the control group, meaning that they on average thought to be more successful in stimulating students' reflection. The results indicate that teachers can develop relevant specific competences in supporting students' reflection skills and improve their practice in a feasible way and in a relatively modest amount of time by participating in a purposeful and coherent programme including relevant content, active learning, and expert feedback.

4.1 Introduction

During their education, nursing students have to develop competences for continuous learning and development during their career. An important skill for this is reflection (Biesma et al., 2008; Schön, 1991). Reflection implies a process of thinking about and interpreting situations, events, experiences, and emotions, aimed at critically analyzing decisions, actions, and effects in order to learn from them (Ertmer & Newby, 1996; Hatton & Smith, 1995; Mann, Gordon & MacLeod, 2009). Supporting the development of students' reflection skills is new for most teachers and demands new competences.

Developing teachers' competences in supporting students' reflection demands professional development. In the past fifty years several approaches for professional development can be distinguished. First, in classical training approaches teachers apply knowledge and practise skills (Tillema & Veenman, 1987). Second, in developmental-reflective approaches teachers are stimulated to think about their current and possible alternative ways of acting (Mena Marcos & Tillema, 2006). Third, in approaches based on principles of social learning activities such as discussing and exchanging experiences with colleagues are important (Bandura, 1977). Currently, evidence-based or expertise-based research into these approaches of professional development emphasise that teacher learning should focus on interactions between teachers and students (Timperley, Wilson, Barrar & Fung, 2007).

Nowadays consensus is growing about the core features of effective professional development of teachers. These features imply: content focus, active learning, coherence, duration, and collective participation (Desimone, 2009; Garet, Porter, Desimone, Birman & Suk Yoon, 2001; Penuel, Fishman, Yamaguchi & Gallagher, 2007). These features have parallels with the three approaches of professional development mentioned above: content knowledge is deemed important in classical training approaches; active learning can be linked to developmental-

reflective approaches; and collective participation makes social learning possible. The feature 'content' refers to the knowledge teachers should develop, including both knowledge about what the students should learn (in this study: reflection) and knowledge about the teaching activities needed to support students' learning. The latter is also known as pedagogical content knowledge; in this study for instance knowledge about giving feedback (Ball, Thames & Phelps, 2008; Bransford, Derry, Berliner & Hammerness, 2005; Shulman, 1987).

Feedback is a key element in formative assessment, both as information about how successful one's functioning is and for suggesting how to alter the gap between current and desired functioning (Sadler, 1989, 2010). The purpose of formative assessment is furthering a learner's development, which assumes a feedback process in which a learner (e.g., a student or a teacher) is actively involved. In the literature 'formative assessment' and 'feedback' are often used interchangeably although these terms do not have the same meaning. In this study we will use the term feedback in the broad sense of formative assessment: as a process of communicating information to support both students' and teachers' competence development.

The aim of the study is to evaluate a training programme for teachers in senior secondary nursing education to support students' reflection skills development. The research question is: What are the feasibility, quality and effects of the programme?

This study is part of a project of the Dutch national research cooperation in vocational education and training in which three formative assessment approaches for teacher professional development are studied: expertise-based feedback as part of a training programme, self assessment combined with peer-coaching, and negotiated assessment. This study represents the expertise-based approach. In two earlier studies we have developed and validated content standards and performance standards in cooperation with experts and teachers in the field of senior secondary nursing education (Dekker-Groen, Van der Schaaf & Stokking, 2011, 2012). For this study we have developed a training programme, based on the core features of effective professional development mentioned above. Special attention is given to available knowledge and expertise from the literature and experts in the field about reflection (the skill nursing students have to develop) and feedback (an important part of the competences teachers need to support students' reflection skills development).

4.2 Theoretical background

Teachers' competence development should result in improvements of teachers' actions and in students' achievement of learning goals. To support this, expertise should be used both about effective teachers' competence development and students' learning goals (Garet et al., 2001; Kuijpers, Houtveen & Wubbels, 2010; Timperley et al., 2007). Consensus is growing that for teacher professional development activities to be effective the following features are important: content focus, active learning, coherence, duration, and collective participation (Desimone, 2009; Garet et al., 2001; Penuel et al., 2007).

The word 'content' in 'content focus' refers to the knowledge teachers should develop. This implies knowledge about what the students should learn (content knowledge) as well as knowledge about teaching actions needed to support students' learning (pedagogical content knowledge) (Ball et al., 2008; Bransford et al., 2005; Shulman, 1987). The word 'focus' means that a teacher development programme should concentrate on what the teachers have to develop, the core of which in this study is: giving feedback to support students' reflection. Both on reflection (e.g., Hatton & Smith, 1995; Mann et al., 2009) and on feedback (e.g., Kluger & DeNisi, 1996; Shute, 2008) much important knowledge has been developed. This knowledge can be used to develop standards for desired actions, both for giving feedback to *students* developing their reflection skills (teachers' actions) and for giving feedback to *teachers* participating in a training programme (trainer's actions).

Active learning can be stimulated by giving teachers feedback to reflect upon and using standards as anchor points (Airasian & Gullickson, 1994; Rosenstein, 2002; Ross & Bruce, 2007). The importance of reflection on experiences for teachers' competence development is broadly recognised (Korthagen, Kessels, Koster, Lagerwerf & Wubbels, 2001; Kwakman, 2003; Timperley et al., 2007). Coherence can be furthered through making connections between the content and the activities of the programme (also known as: teach as you preach) (Penuel et al., 2007). The content and activities also have to build upon and link up with the teachers' existing cognitions, knowledge, skills, and needs (Tillema & Imants, 1995).

Duration concerns the time that has to be spent in competence development activities and also the intensity of these activities. Offering time to get used to and practise activities improves teachers' acceptance of and their commitment to the programme (Evans, 2008). Embedding activities in the teachers' daily work situation is much more important than the specific form of the activities (Garet et

al., 2001). Collective participation makes social learning and support possible (Penuel et al., 2007).

The growing consensus about the importance of these features is supported by the observation of Whitcomb, Borko and Liston (2009), that nowadays the traditional in-service staff workshops are often replaced by interconnected activities that are situated in practice and are focussed on concrete teacher tasks, embedded in a social context, applicable in daily practice, focussed on both teachers' and students' learning process and aiming for the realisation of improvements in practice. It can be concluded that a contemporary training programme should include the features mentioned (cf. Opfer & Pedder, 2011).

Comparable insights about teachers' competence development pertaining to the importance of their experience in practice and the use of feedback have also been developed in medical education and nursing education research (Cader, Campbell & Watson, 2005; Dexter et al., 1997; Steinert et al., 2006).

The growing consensus in the literature also pertains to general insights into processes of human learning and development. These insights include that these processes are highly the same for students and teachers, that for both while learning their knowledge, skills, beliefs, expectations, attitudes and behaviour are in constant interplay, that for learning environments to be effective interactions with others are important, and that teachers not only influence their students but students' behaviour and learning also function as feedback for teachers (Ajzen & Fishbein, 2005; Bransford et al., 2005; Eraut, 1994; Fishman, Marx, Best & Tal, 2003; Roelofs & Sanders, 2007). This implies that for evaluating a teacher training programme a fairly broad selection of aspects should be covered of relevant knowledge, skills, beliefs, expectations, attitudes and behaviour (teacher practices, learning activities, interactions).

4.3 Methods

Design

A quasi experimental design was used, with the training programme as intervention, including pre-measurements, measurements during the training, and post-measurements within the training group and a selection of pre-measurements and post-measurements within the control group. In the classification of Borko (2004) the research concerning the training programme can be classified as a phase 2 study: a well-specified programme that was carried out on several schools, but with the same trainer instead of multiple trainers.

Context and participants

During the study the situation in Dutch nursing education was turbulent, with many governmental initiatives, organisational changes, cutbacks, high workload and changing management. In these circumstances random selection of schools or teachers and matching of groups for the quasi experimental design were not feasible. The researcher approached seven schools, consulted the teams and gathered information about needs and circumstances, and when the manager agreed to take part in the study the manager and the researcher together decided which teachers would be asked to participate in the training group. Selection criteria were: being interested in developing own competences, being responsible for teaching and guiding students for a minimum number of hours per week, not being close to retirement, and having enough time to participate in the programme. The remaining teachers became a comparison group for the purpose of analyzing the representativeness of the participating teachers. Eight teachers from other teams were willing to participate in a control group. These teachers did not receive information about the research questions nor the training programme and did not have any contact with the teachers from the training group.

The training group included 14 female and five male teachers on five schools, with on average 13 years of teaching experience. All teachers of the control group were female and they had on average 16 years of teaching experience. The comparison group consisted of 24 female and four male teachers with on average 15 years of teaching experience. The difference in teaching experience between the groups was not significant.

Intervention

The programme consisted of three cycles each including the same sequence of activities: a group meeting, a period of independent practice, an individual meeting per participating teacher, and a second period of independent practice. See Figure 4.1 for an overview of the structure and content of the programme. Because of the workload in nursing education the duration of the meetings was adjusted to the time schedules of the participating schools. Group meetings lasted 90 minutes and individual meetings 45 minutes.

Content of the programme. Based on the literature described before, the programme was given a strong focus on content, using available knowledge about reflection and feedback, and information from the target group on current needs and circumstances. Also used were the content standards and performance

standards for nursing teachers supporting students' reflection skills development as developed and validated in two earlier studies in cooperation with experts and teachers in the field (Dekker-Groen et al., 2011, 2012). An example of a rubric of a competence is included in the Appendix at the end of this chapter.

Cycle	Activities → Main topics ↓	Group meeting: information, instruction, discussion and exercises using daily practice, cases and video fragments	Eight-week period of individual practice and videotaping a teacher- student reflection dialogue	Individual meeting: feedback, discussion and teacher reflection based on video fragments of the dialogue	Eight-week period of independent individual practice
1	Giving students instruction and asking questions on reflection	Start →	→	→	→
2	Giving feedback on and discussing students' reflections	→	→	→	→
3	Coaching students in depth and assessing students' reflections	→	→	→	→ Finish

Figure 4.1 Structure and content of the programme

The content of the first group meeting was the competence 'Giving information and instruction how to reflect'. Central in this competence were 'asking questions to stimulate reflection', 'objects of reflection' and 'thinking activities of reflection'. A distinction was made between general and specific questions and between short answer and deep reasoning questions (e.g., Graesser & Person, 1994). During a period of individual practice teachers gave feedback on students' reflections and the effect and impact of this feedback could be derived from students' reactions (Chin, 2006). In the individual meetings the teachers received feedback about their

way of questioning and the amount of short answer questions and deep reasoning questions. Also a connection was made with the theory and practice of effective questioning with the intention to support the students' reflection.

In the second group meeting the focus was on the competence 'Development through teacher feedback'. Central in this competence were 'giving feedback', 'receiving feedback', 'asking feedback', 'discussing patterns' and 'making relations', focused on students' reflections. The feedback on content was based on categories distinguished by Chi (1996; Chi, Siler, Jeong, Yamauchi & Hausmann, 2001): 'suggestive feedback', 'didactic explanations' and 'corrective feedback' and specific categories from Chin (2006), and Shute (2008). Also the classification into feedback and feed forward was used (Hattie & Timperley, 2007). Two feedback categories focusing on the interaction between the teacher and the individual student were question categories from Graesser and Person (1994) and reaction categories from Blatt, Confessore, Kallenberg and Greenberg (2008). In the individual meetings the teachers received feedback about their way of asking questions and giving feedback. Also attention was paid to preparing and guiding reflection conferences and to developing alternative feedback repertoire.

In the third group meeting the focus was on two competences. First, attention was paid to the competence 'guiding the learning process of how to reflect' including 'giving guidance', 'teaching how to develop an attitude' and 'guiding students who learn to reflect'. Second, the focus was on the competence 'Use criteria to assess reflection' including 'broad reflection', 'deep reflection' 'systematic reflection' and 'discussing constraints of students to reflect'. Questions which stimulated students to use the thinking activities 'explaining' and 'concluding' leading to 'formulating intention' can support students to broader, deeper and more systematic reflection (Mezirow, 1991; Oosterbaan, Van der Schaaf, Baartman & Stokking, 2010; Wallman, Kettis Lindblad, Hall, Lundmark & Ring, 2008). In the individual meetings the teachers received feedback about the impact of their questions and feedback on students' reflections.

Procedure of the programme. Active learning is not only stimulated by informing and instructing the participating teachers but also letting them try out new skills and practices and giving them feedback on videotaped conferences with individual students and stimulating reflection and discussion on their experiences (Borko, Jacobs, Eiteljorg & Pittman, 2008; Caris-Verhallen, Kerkstra, Bensing & Grypdonck, 2000; Van Es & Sherin, 2008; Roter et al., 2004). Coherence was given attention throughout by frequently connecting the content and the learning activities and through interactions and discussions in the meetings. Duration (enough time and intensity) was furthered by alternating group meetings, periods

of individual practice and individual meetings. Collective participation was enhanced by group meetings including discussions related to the teachers' own work contexts and experiences.

In each group meeting, the trainer distributed information, expounded the theme and goals, gave instruction, and demonstrated and discussed theory in interaction with the participants. The content resembled the teachers' intentions and actions in their daily practice. Teachers discussed video fragments of representative teacher-student conferences with their colleagues and exchanged experiences. Teachers were encouraged to ask questions about the theory presented and the video fragments, to bring in own experiences and to discuss the content within the group. Then they applied theory and skills in short cases and video fragments, followed by a discussion. Subsequently conclusions were drawn about what was learned and could be practised. The meeting ended with an example of a teacher action that can be expected to support students to reflect. Also an assignment was given for future reflection conferences.

Each group meeting was followed by a period of independent practice. Possibilities of focussed practising were offered by giving teachers a form with questions supporting the preparation of reflection conferences with a student. Moreover teachers could determine what the result was of their practise by filling the form for reflection afterwards. About two months after a group meeting the teachers videotaped themselves a reflection conference with an individual student. The trainer used the forms which were filled out as content in the individual meetings with the teachers.

In the individual meetings the trainer gave specific feedback and guidance and discussed with each teacher three video fragments of the videotaped reflection conference with a student. The teacher reflected on teacher's own actions in these conferences in relation to the content of the training programme. Together with the trainer the teacher determined the impact of the teacher's feedback on the student in relation to the teacher's own goals. Moreover the teacher reflected about what is successful and what could be improved and how. After each individual meeting another period of independent practise was planned for teachers to gain more experience.

Instrumentation and data collection

Data were collected before, during and after the programme. In total nine instruments were used, developed by the researcher. See Figure 4.2 for an overview of measurement periods, instruments, constructs, and question formats.

Measurement period	Instrument	Construct	Question format	Also in control group
Pre-measurement	1 Teacher's knowledge, experience, beliefs and practices	Knowledge about reflection and feedback	Scale ^a	X
		Experience in guiding students' reflection	Scale	X
		Extent of students' reflecting on the reflection objects and importance of the thinking activities	Scales	X
		Meaning and importance of reflection and feedback	Open	X
		Frequency students' reflecting and teacher's feedback	Single	X
		Expectations of participating in the programme	Tick	
	2 Self-assessment of competences	Teacher's self-assessment of competences with regard to instruction, guidance, feedback, reflection	Open	
Measurement during the programme	3 Evaluation of group meeting (3x)	Teacher's appraisal of content and activities (at the end of each group meeting)	Scale	
		Teacher's goals concerning themselves and students	Open	
	4 Practice and beliefs about preparing conferences	Frequency and perceived difficulty of and attached importance to preparing a teacher-student reflection conference (once, before the second cycle of the programme)	Scales	
	5 Evaluation of individual meeting (3x)	Teacher's appraisal of content and activities (at the end of each individual meeting)	Scale	
		Teacher's goals concerning themselves and students	Open	
	6 Learner report (3x)	Concerns and activities with regard to developing competence in guiding students' reflection (at the end of each cycle of the programme)	Tick	
Perceived results of feedback got during the meetings (at the end of each cycle of the programme)		Tick		
Post-measurement	7 Evaluation of the quality of the programme	Teacher's appraisal of a number of organizational, material and pedagogical features and aspects of the programme as a learning environment	Scale	
		Teacher's appraisal of the content and purposes of the meetings, the quality and relevance of the knowledge brought in and the feedback given by the trainer	Scale	
		Perceived quality of the competence descriptions used	Scale	
		Teacher's perception of the extent to which own needs and expectations are fulfilled and results are useful	Scale	
	8 Self-assessment of the effect of the programme	Teacher's self-assessment of own development of specific knowledge, skills and attitudes during and as an effect of the programme	Scale	
	9 Self-estimated development, current actions, and attached importance to these actions	The extent to which the teacher estimates to have developed own competences in activities in instructing, giving feedback, and guiding students' reflection	Scale	
		The extent to which the teacher practices these activities, and attaches importance to these activities	Scales	X
	1 Teacher's beliefs	Extent of students' reflecting on the reflection objects and importance of the thinking activities (same questions as in pre-measurement)	Scales	X
		Meaning and importance of reflection and feedback	Open	X
	2 Assessment of competences	Teacher's self-assessment of competences with regard to instruction, guidance, feedback, reflection (same questions as in pre-measurement)	Scales	

^a Scale = several items in Likert format which together form a reliable scale (in terms of Cronbach's alpha); Single = single questions, pre-structured answers; Tick = items to be ticked; Open = single open questions.

Figure 4.2 Data collection: measurement periods, instruments, constructs, question formats

Most of the data consisted of answers on questions measuring constructs using Likert-type items. Measuring constructs in this way (in this study using three, four, five or seven answer possibilities and three to fifteen items per construct) resulted in measures in terms of amount or extent: the amount of teachers' knowledge and experience and the extent to which students reflect on certain objects (1 = not or few, 2 = moderate, 3 = strong), perceived difficulty or usefulness (e.g., 1 = very difficult, 5 = very easy), the level on which they self-assess their competences (1 = beginner, 2 = basic, 3 = advanced, 4 = expert), the relevance of the content and activities of the meetings (1 = not, 2 = somewhat, 3 = sufficient, 4 = considerable, 5 = fully), the perceived quality of aspects of the programme (1 = not, 5 = optimal), the frequency of certain activities (1 = never, 5 = always).

Before the start of the programme the teachers filled out two instruments concerning their relevant knowledge and experience, the extent to which their students reflected on a number of reflection objects, the importance they attached to a number of thinking activities, the frequency of students' reflecting and the teacher's giving feedback, the expectations of the programme, and their self-assessment of their competences. During the programme the participating teachers at the end of each meeting completed instruments in which they appraised the content and activities of the meeting and before the second group meeting they also answered questions about their preparation of teacher-student reflection conferences; in addition, they filled out three times a learner report with questions about their learning concerns and activities and the results they perceived of the feedback they received from the trainer.

After the programme they completed instruments in which they evaluated features and aspects of the programme and the extent to which their needs had been fulfilled and the extent to which they thought as a result of the programme to have developed their knowledge, skills and attitudes; in addition they answered questions about their current teaching activities and the importance they attach to these activities and they indicated the extent to which their students currently reflect on a number of reflection objects and the importance they currently attach to a number of thinking activities, and they again self-assessed their own competences.

Data analyses

Quantitative analyses included preliminary analyses of the reliability of each scale developed to measure a certain construct, and the comparability of training group and control group, and descriptive analyses, analyses of differences, and analyses of relations. Minimum criteria for scale reliability were item-rest correlations at

least $\geq .20$ and preferable $\geq .35$ and Cronbach's alpha at least $\geq .60$ and preferable $\geq .70$. Analyses on differences between groups of teachers and between moments of measurements were done using *t*-tests and Ancova. Additionally relationships were checked with gender (using *t*-tests) and age, experience as a teacher, and number of years working at the present school (using correlations). Criterion for significance was $p < .05$. Effect sizes were calculated.

4.4 Results

Results of preliminary analyses

First, the quality of the scales was analysed. Table 4.1 gives per instrument per construct an example item and the relevant psychometric information. Where the same scale was used for several groups, assessment objects and occasions, the analysis was repeated each time and the range of the alphas is given.

Next, it was analysed whether the teachers selected for participation in the programme can be judged to be representative for all teachers from the participating schools. Data from 12 scales and single questions were available from all three groups. On two scales there was a significant difference between the training group ($n = 19$) and the other teachers (comparison group and control group together, $n = 36$). The teachers in the training group on average self-reported less knowledge about reflection and feedback (training group $M = 2.1$, $SD = 0.38$, remaining teachers $M = 2.4$, $SD = 0.48$), and less experience with reflection and feedback ($M = 2.0$, $SD = 0.43$ respectively $M = 2.4$, $SD = 0.40$). On all other questions the groups did not differ significantly. These findings indicate that the teachers in the training group can be considered to be sufficiently representative of the teachers of these nursing schools.

Finally, it was checked to what extent the training group and the control group before the start of the programme were comparable. Teachers in the training group had on average less knowledge ($M = 2.1$, $SD = 0.38$ respectively $M = 2.5$, $SD = 0.50$) and experience ($M = 2.0$, $SD = 0.43$ respectively $M = 2.6$, $SD = 0.37$), but their students reflected more often (1-2 times per week respectively 1-2 times per month). These differences will be controlled for using covariance analyses. The frequency of giving students feedback on their reflection in both groups was on average twice a week.

Table 4.1

Scales: instruments, constructs, item examples, and psychometric information

Instrument	Construct	Item example	Psychometric information		
			K	Type	α
1	Knowledge about reflection and feedback	Current national qualifications of nurses	5	1-3	.65 - .86
	Experience in guiding students' reflection	Giving feedback on a students' reflection	7	1-3	.77 - .80
	Extent of students' reflecting on the reflection objects (2 times) (in pre- and post-measurement)	Purposeful professional action	7	1-3	.64 - .76
	Importance of the thinking activities (2 times) (in pre- and post-measurement)	Analysing experiences and events	9	1-3	.77 - .92
	Expectations of participation in programme	Receive information on recent insights in giving feedback	5	0/1	.84
2	Teacher's self-assessment of four competences (2 times) (in pre- and post-measurement)	Teaches students to receive feedback	3-5	1-4	.61 - .88
3	Appraisal of content and activities of group meeting (3 times)	Did the explanation give insight in how to ask questions?	15	1-5	.94 - .97
4	Frequency of preparing a reflection conference	I prepare questions for a reflection conference	5	1-5	.92
	Difficulty of preparing a reflection conference	I think formulating feedback goals is easy	5	1-5	.85
	Importance of preparing a reflection conference	Goal achievement concerning students' reflection	3	1-5	.81
5	Appraisal of content and activities of individual meeting (3 times)	Were fragments chosen representative for your practice?	5	1-5	.83 - .92
6	Perceived results of feedback got during the meetings	I was made to think about how I guide students' reflection	9	0/1	.66 - .72
7	Appraisal of organizational, material and pedagogical features and aspects of the programme as a learning environment	It gave me freedom to decide about results I want to attain	15	1-5	.96
	Appraisal of the content of the meetings, quality and relevance of the knowledge brought in and feedback given by trainer	The feedback from the trainer	15	1-5	.94
	Perceived quality of the competence descriptions used	Relevant content for nursing education	5	1-7	.94
	Teacher's perception of extent to which needs and expectations are fulfilled	Did the training approach comply with your needs?	7	1-5	.94 - .97
8	Teacher's self-assessment of own development of specific knowledge, skills and attitude during and as an effect of the programme	I have become more skilled in guiding students' reflection	5	1-5	.92
9	Self-estimated amount of development of competences in a number of activities	Asking the student about reasons for acting	5-9	1-5	.71 - .91
	Current frequency in practice of these activities	Continue to ask questions on what the student says	5-9	1-5	.51 ^a - .87
	Importance attached to these activities	Active listening to what student brings in	5-9	1-5	.72 - .91

^a For one of the scales the alpha was lower than .60.

Pre-measurements before the start of the training programme

Teachers' initial knowledge, experience, beliefs and practices. The teachers who participated in the programme initially varied in their self-assessed knowledge of currently desired teacher competences in Dutch nursing education, including the guidance of and provision of feedback on students' reflections ($M = 2.0$, range 1.5 – 2.8, $SD = 0.38$). They also varied considerably in their experience with teaching reflection and guiding of, giving feedback on, and assessing students' reflections ($M = 2.1$, range 1.4 – 3.0, $SD = 0.43$).

Concerning seven objects on which students may reflect, on the 3-point scale used (1 = not or little, 2 = moderate, 3 = strong) the teachers indicated that their students mostly reflected on their professional attitude and professional development (68% strong, 32% moderate) and their personal development (68% strong, 26% moderate). They also reflected on study progress and future perspectives (53% strong, 47% moderate), the relationship between theory and practice (53% strong, 37% moderate), purposeful professional action (37% strong, 53% moderate), and professional knowledge and skills (32% strong, 53% moderate). According to the teachers, their students reflected the least on theory in nursing education (16% strong, 47% moderate). The average on these objects was relatively high ($M = 2.4$, range 2.1 – 2.9, $SD = 0.36$). See also Table 4.2.

The teachers considered eight thinking activities students might use during reflection as specified in the questionnaire on average to be of relatively high importance for reflection ($M = 2.6$, range 2.2 – 3.0, $SD = 0.35$).

Answering open questions, all teachers mentioned that reflection is important in nursing education because it contributes to their students' insights and learning process, while according to half of the teachers ($n=9$) feedback is important because it contributes to students' learning process and development and according to the other half ($n=10$) this is the case because it gives students' insight into their functioning. Asked to what extent their students already reflect, all teachers answered that this differs among students and that the quality of their reflection often is mediocre.

Table 4.2

*Participants' scores about students' reflection on objects before and after the training
(Part of instrument 1: Beliefs)*

Object	Pre-measurement				Post-measurement				T-test of difference		
	%	%	M	SD	%	%	M	SD	t	df	p
	Strong	Mod ^b			Strong	Mod ^b					
Theory in education	16	47	1.79	0.71	17	61	1.94	0.64	-1.84	17	.08
Relation theory and practice	53	37	2.42	0.69	81	19	2.83	0.38	-1.83	17	.09
Professional knowledge and skills	32	53	2.16	0.69	44	44	2.33	0.69	-.90	17	.38
Purposeful professional acting	37	53	2.28	0.67	67	33	2.67	0.49	-2.38	17	.03 ^a
Personal development	68	26	2.63	0.60	72	22	2.67	0.59	.00	17	1.0
Professional attitude and development	68	32	2.65	0.49	78	22	2.78	0.43	-.44	17	.67
Study progress and future perspective	53	47	2.53	0.51	56	44	2.56	0.51	.00	17	1.0

^a $p < .05$, ^b Mod=Moderate.

Half of the teachers indicated that it is difficult for students to look with enough depth to themselves and the other half reported that students often don't see the purpose of reflection. Asked about the problems encountered, half of the teachers said to experience difficulties in teaching students to reflect and the other half that students' behaviour or resistance is a problem. For half of the teachers breaking through students' resistance and motivating students is a problem in teaching reflection skills. The other half mentioned the age period of the students (puberty) as a cause for their troubles in opening themselves and going into depth.

The teachers indicated that they are used to give feedback in various ways: using reflection models and learning goals, on content and form, on strong points and improvement points, on fixed moments and on initiative of the teacher, verbally and in written form. Asked on what they give feedback and to what they refer, two-thirds of the teachers said to give feedback on students' behaviour, attitude or learning process and one third mentioned the content to be learned. The teachers referred to agreements made with the students as well as to rules and criteria valid for certain situations and/or products. Two-thirds of the teachers said that when they give feedback they pay attention to the adequacy of the situation (an individual meeting or a group meeting) and whether there is enough time, and one third of the teachers mentioned the atmosphere or the student's feelings.

On the question to what they pay attention in formulating their feedback, well over half of the teachers said it is important to be clear and understandable. A quarter of the teachers mentioned the importance of 'first-person messages' and the others mentioned the importance of being positive and not being judging.

Asked what they think to be important once the feedback is given, something more than half of the teachers answered that it is important that the student has comprehended what has been discussed and knows how to go forward. Finally, asked which problems they experience in giving feedback, one-third of the teachers mentioned their own skills, particularly in how to make critical remarks, and some teachers said being afraid to hurt students' feelings. Two-thirds of the teachers mentioned students' personal characteristics such as taking feedback personally and becoming defensive.

The frequency in which students reflected according to the teachers varied from each day to fortnightly and was on average between once and twice per week, while the frequency in which they give the students feedback on reflection varied from each day to once per month and was on average twice per week.

Teachers' expectations, needs, and self-assessed initial competences (pre-measurement). The teachers expected that the training would give them: more insight into the way they give students feedback (n = 13), information about other ways to do this (17), and about recent insights about it (9), insight into the way of giving feedback in accordance to school's policy (8), and possibilities for exchanging thoughts with colleagues and researchers (7). Asked in which aspects they wanted to grow concerning reflection and feedback, half of the teachers liked to grow in asking questions (n = 10) and most other teachers in guiding students' reflection (6). On the main competences on which the programme was focused (see Table 4.3), the teachers assessed themselves on average on level 3 (advanced), but individually they differed considerably in their assessments.

Table 4.3

*Participants' scores on four competences before and after the training
(Instrument 2: Self-assessment)*

Competence (scale)	Pre-measurement		Post-measurement		T-test of difference	
	Range	M (SD)	Range	M (SD)	T	p
Giving information and instruction about reflection	1.8-4.0	3.0 (0.63)	2.0-3.8	3.0 (0.55)	.46	.65
Guiding the process of learning to reflect	1.8-3.5	3.0 (0.53)	2.0-4.0	3.2 (0.59)	-1.86	.08
Giving and teaching asking and receiving feedback	2.0-4.0	3.0 (0.52)	2.2-3.8	3.1 (0.49)	-.78	.45
Assessing reflections	2.0-3.8	2.9 (0.46)	2.5-3.8	3.1 (0.39)	-2.39	.03

Measurements during the programme

Teachers' appraisals of meetings and activities. After each meeting the participants filled out a questionnaire with a 5-point Likert-scale about the relevance (1 = not, 2 = somewhat, 3 = sufficient, 4 = considerable, 5 = fully) of the content and the usefulness of the instruction, the exercises, and the discussion. Participants rather appreciated the three group meetings ($M = 3.8$, $SD = 0.68$; $M = 3.9$, $SD = 0.67$; $M = 3.5$, $SD = 0.90$) and were very satisfied with the individual meetings ($M = 4.5$, $SD = 0.52$; $M = 4.5$, $SD = 0.59$; $M = 4.3$, $SD = 0.65$). Pairwise *t*-tests showed that these mean evaluations do not significantly differ between meetings.

At the end of each group meeting, the teachers were stimulated to prepare reflection conferences with individual students and to reflect on it afterwards. As a preparation they answered written open questions about their own goals and the goals they wanted the student to attain, how they planned to achieve this, and which questions they wanted to pose the student. Afterwards they answered written open questions about how they had handled the conference and what the impact was on themselves and on the student, what they learned from it and which intentions they had for next time. The answers received were used by the trainer in the individual meetings with the teacher. In the first cycle of the programme the teachers' goals focused on posing conscious and open questions and giving the students more room to think by introducing a pause after each question. Further, they aimed to avoid posing two or more questions at the same time. They really intended the students themselves to think and grow. In the second cycle teachers' goals regarded giving the students more consciously feedback. They strove for the students to be more actively involved in the dialogues. In the third cycle the teachers wanted to stimulate students' reflective thinking activities and to monitor their reflection level. They wanted the students to reflect at a higher level.

Learner report. On three moments during the programme the teachers were asked to complete a learner report with questions about their learning concerns and activities and about the results they perceived from the feedback they had received from the trainer (see Table 4.4).

Table 4.4

Concerns, activities and results per moment of measurement and in total (percentages)
(Instrument 6: Learner report)

	Total	Moment 1	Moment 2	Moment 3
<i>1. In my development of guiding of students' reflection skills development I was concerned with</i>				
1a The content of reflections (themes within professional practice)	62	47	72	67
1b The process of reflection (before, during and after an activity)	55	68	44	50
1c The results in terms of insights in situations in the professional practice	47	42	50	50
1d The results in terms of a critical view on the (own role in the) professional practice	67	58	67	78
1e Safety: caring that the student feels comprehended and opens him/herself	58	63	51	61
1f Self-regulation: caring that the student takes the next step in the own learning process	58	63	56	56
<i>2. For this I have carried out the following activities</i>				
2a I tried all sorts of things	24	16	17	39
2b I consulted a colleague	22	26	17	22
2c I looked for information in manuals, journals, books and on internet	29	21	28	39
2d I put my faith strongly in my own feelings and intuition	71	79	67	67
2e I thought critically about what could be a good approach	75	79	67	78
<i>3. The feedback I received about my guiding of students' reflection skills had the following results for me</i>				
3a I did know what to do to go ahead with guiding students on reflection	33	32	39	28
3b I was forced to think about my way of guiding students on reflection	93	89	100	89
3c I was stimulated to choose another approach for guiding reflection	62	63	56	67
3d I got more self-confidence about my guiding of reflection	25	16	22	39
3e I got more insight in different situations during the guiding of reflection	75	89	61	78
3f I got an other view on my guiding of reflection	24	16	17	39
3g I comprehended better the way others think about guiding reflection	16	16	17	17
3h I felt supported in what I want to achieve with guiding reflection	27	21	39	22
3i I was better able to judge my own way of guiding reflection	69	68	72	67

In the first cycle of the programme the teachers' concerns were especially on the process of reflection and on students' safety and self-regulation. In the second cycle they showed an increased focus on the content of students' reflections and on results of guidance in terms of a critical look at their own role as a teacher. The latter also got more attention in the third cycle. Learning activities the teachers undertook during the programme were based on their own feelings and intuition, and on critical self-examination of what a good approach would be. Two activities became pursued more during the training: trying all sorts of things and looking for information. Consulting a colleague was the least-mentioned activity. The results the teachers mentioned of the feedback they received were that they were made to think and given more insight in guiding reflection. They indicated they were stimulated to carry out their guidance in a new way and could better judge their own way of guiding reflection. Results in terms of knowing what to do and feeling supported in what to achieve with guiding reflection were less often reported. Getting more self-confidence and getting another view on guiding reflection were results they more often mentioned in the third cycle of the training. The nine results did form a reliable scale (see instrument 6 in Table 4.1), on which the teachers on average ticked half of the items (47%).

Post-measurements after the programme

Teachers' overall appraisals of the programme. After the programme the teachers were asked about their overall appraisal of the meetings, both on **organisational**, material and pedagogical aspects (form) and on the content, the quality and relevance of the knowledge brought in by the trainer, and the feedback she had given (content). On average they appraised both fairly positive ($M = 3.8$, $SD = 0.69$) and ($M = 3.6$, $SD = 0.67$). The teachers had spent, according to themselves, on average 26 hours on the programme (including the meetings, the individual learning activities, and the periods of the independent practice). The quality of the descriptions and rubrics concerning the four main competences the programme was focused on, evaluated by the teachers on five criteria (content relevance, unambiguousness and completeness of formulation, relevance and representativeness for nursing education) on a 7-point scale, resulted in an average score over the four competences of 5.3 ($SD = 0.90$). Asked whether the content and activities in the meetings had met their own needs and expectations, were feasible for their teaching practice and reflection conferences, and had supplied handles to stimulate students' reflection, the group meetings on average were appraised with 3.4 ($SD = 0.76$) and the individual meetings with 3.7 ($SD = 0.86$).

Teachers' self-assessed development as result of the programme. The teachers were asked to assess themselves on the extent to which they as a result of the programme had developed their knowledge, skills and attitude in guiding reflection and the extent to which they had become more conscious of their own behaviour while guiding and stimulating students' reflection. On the 5-point Likert-scale used their assessments varied considerably. Scores were on average 3.8 (range 2.8 – 5.0, $SD = 0.69$). In a more specified way the teachers were also asked to assess their own development in conducting the following activities: attending to relations between theory, practice, actions and experiences; giving feedback on possibilities for improvement; giving feedback stimulating thinking in general; supporting thinking activities which are specific for reflection; and guiding students' personal and professional development. The average on these five 5-point Likert-scales was 3.7 (range 3.0 – 4.5, $SD = 0.44$).

Teachers' current activities and attached importance. In addition to their development in conducting the five groups of activities mentioned above, the teachers also indicated the extent to which they currently conduct these activities and the extent to which they think these activities are important, both again on 5-point Likert-scales. The average extent to which they currently conducted these activities was 3.7 (range 3.2 – 4.0, $SD = 0.37$), the average importance score was 4.3 (range 3.8 – 4.6, $SD = 0.45$). See Table 4.5.

Students' reflection and teachers' feedback and their self-assessed competences (repeated measurements). Concerning the same objects and thinking activities and the same 3-point Likert-scale used (1 = not or few, 2 = moderate, 3 = strong) as in the pre-measurement the teachers indicated in the post-measurement an increase on all seven objects on which students may reflect (see Table 4.2). The teachers now indicated that their students mostly reflect on their professional attitude and professional development (78% strong, 22% moderate) and their personal development (72% strong, 22% moderate). Further, they reflect on relationship between theory and practice (81% strong, 19% moderate), purposeful professional actions (67% strong, 33% moderate), study progress and future perspectives (56% strong, 44% moderate), and professional knowledge and skills (44% strong, 44% moderate). Least reflected on was the theory in nursing education (17% strong, 61% moderate). The average score of all objects was 2.6 (range 2.1 – 2.9, $SD = 0.21$). The teachers now scored the importance of the eight thinking activities specified in the questionnaire students might use during reflection on average with 2.6 (range 2.0 – 3.0, $SD = 0.27$).

Table 4.5

*Current actions and attached importance concerning aspects of guiding students' reflection, post-measurement in training group and control group
(Instrument 9: Current actions and beliefs)*

Scale	Training group			Control group			T-test		
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>df</i>	<i>p</i>
<i>1 Current actions</i>									
1a Attending to relations between theory, practice, actions and experiences	17	3.8	0.43	8	3.4	0.68	1.9	23	.07
1b Feedback on possibilities for improvement	15	3.6	0.57	8	3.4	0.43	.80	21	.43
1c Feedback stimulating thinking in general	17	3.7	0.39	8	3.6	0.47	.85	23	.40
1d Supporting thinking activities which are specific for reflection	17	3.7	0.43	7	3.5	0.34	1.4	22	.18
1e Guiding students' personal and professional development.	15	3.6	0.40	8	3.3	0.40	1.9	21	.07
Current actions total (scale of five scales)	14	3.7	0.37	7	3.4	0.38	2.0	19	.06
<i>2 Importance</i>									
2a Attending to relations between theory, practice, actions and experiences	17	4.2	0.61	8	4.1	0.69	.41	23	.69
2b Feedback on possibilities for improvement	16	4.1	0.60	8	4.2	0.51	-.40	22	.70
2c Feedback stimulating thinking in general	17	4.4	0.40	8	4.4	0.51	.11	23	.91
2d Supporting thinking activities which are specific for reflection	17	4.3	0.42	7	4.2	0.51	.60	22	.56
2e Guiding students' personal and professional development.	15	4.2	0.42	8	4.2	0.39	.20	21	.85
Importance total (scale of five scales)	15	4.3	0.45	7	4.2	0.50	.51	20	.62

Also repeated were the open questions about the meaning of reflection and feedback mean and why these activities are important. All teachers ($n = 19$) said that for them reflection is looking (back) to own behaviour or functioning in certain situations. Eight teachers also mentioned looking at own thinking, motives, experiences, choices, knowledge, skills and/or attitude. Most teachers ($n = 15$) specified that reflection should include looking or thinking critically or with some

depth, by analysing causes and effects, judging what went well and what not so well and why, and the like. Also most teachers ($n = 15$) indicated on which reflection should be directed, namely: learning, improvement, development, transfer or application in new situations. Asked about the importance of reflection, teachers added: becoming aware of and developing insight in choices, effects and possible improvements, development and growth.

For two-thirds of the teachers ($n = 12$) feedback meant reacting on, making explicit, giving information, and drawing attention to how people interpret something or what the effect of something is, and also two-thirds said that feedback means judging in terms of good or less good, giving constructive criticism, or mentioning possibilities for improvement (some teachers answered in both ways). One teacher interpreted feedback only in terms of feeding back what happened, without any further interpretation of qualification. Four teachers described feedback in terms of asking questions, discussing together, or counselling, so more in terms of interaction. Asked about the importance of feedback, most teachers said that it contributes to learning, insight, improvement and development. Half of the teachers ($n = 9$) said feedback to have the function of helping students to recognize or becoming aware of something (which would otherwise probably stay a blind spot) and/or judging something in terms of good or less good. One teacher said that feedback simulates reflection, another teacher that feedback furthers collaboration within a team.

On the main competences on which the programme was focused (see Table 4.3), now the teachers assessed themselves on average slightly above level 3 (advanced), but individually they still differed considerably in their assessments.

Comparison between pre- and post-measurement

Concerning reflection objects (see Table 4.2) *t*-tests of difference in average score before and after the programme concerning reflection objects (see Table 2) showed a significant increase for ‘purposeful professional acting’ ($t(17) = -2.4, p = .03, r = .50$). On the scale of all objects on which the students may reflect, the average score after the programme is significantly higher than before ($t(17) = -2.5, p = .02, r = .52$). See Table 4.2 for the extent to which the students, according to the teachers, reflected on the separate objects after the programme.

The teachers’ scores on the importance of two of nine thinking activities during reflection were on average significantly higher after the programme than before, namely ‘drawing conclusions’ ($t(13) = -2.3, p = .04, r = .53$) and ‘formulating intentions’ ($t(14) = -3.5, p = .01, r = .68$).

Participants of the training group were in the post-measurements questioned again on what they understood by reflecting and feedback and why, according to them,

these are important. Answers on average included more specifications and nuances than in the pre-measurements, and more different aspects were mentioned. In the post-measurements they paid somewhat more attention to feelings and perceptions of students and to what students had learned of their reflections. As to feedback, in the post-measurements more teachers mentioned the importance of reduction of their students' blind spots.

Concerning teachers' self-assessment of their competences (see Table 4.3), their average judgment of competence 6 (Assessing reflections) significantly increased ($t(15) = -2.4, p = .03, r = .53$).

Comparison between training group and control group

The questions as part of the measurements after the programme concerning the teachers' current activities and the importance they attach to these activities, on which we reported above, were also asked from the teachers in the control group (see Table 4.5). The scale means concerning teachers' current activities (their self-reported behaviour), are consistently a little higher in the training group than in the control group, while three of these five differences tend to significance. The scale means on the five scales measuring the attached importance are essentially the same for both groups.

The questions about the extent to which the teachers' students reflect on a number of reflection objects and the importance the teachers attach to a number of thinking activities during reflection, which questions were asked both before and after the programme and the results of which in the training group we described above, were also asked from the teachers in the control group, also both before and after the programme. In contrast to the training group, in which the means of one reflection object and the scale of all reflection objects after the programme were significantly higher than before, in the control group no significant difference was found between pre- and post-measurement.

For both questions, students' reflection objects (as reported by their teachers) and the importance the teachers attach to the thinking activities, we conducted analyses of covariance with 'group' as factor and the respective pre-measurement as covariate. On the scales of three of the reflection objects and on the overall scale of all reflection objects in the training group the means after the programme, corrected for the scores before, were significantly higher than those in the control group. For 'relationship between theory and practice' the results were $F(1, 23) = 4.3, p = .05, \text{partial } \eta^2 = .16$; for 'purposeful professional actions' $F(1, 23) = 5.3, p = .03, \text{partial } \eta^2 = .19$; for 'professional attitude and professional development' $F(1, 23) = 5.4, p = .03, \text{partial } \eta^2 = .19$; and for the overall scale $F(1, 23) = 4.9, p = .04, \text{partial } \eta^2 = .18$. The means on the importance the teachers attach to the

thinking activities, also corrected for the scores before, were not significantly different between both groups of teachers.

The open questions about what was understood by reflection and feedback, and why these concepts are important, were in the post-measurements not only asked in the training group (see earlier) but also in the control group. In the control group, in respect of reflection, only one of the eight teachers mentioned that reflection is about looking back by students at their own feelings and perceptions and about analysing of how they did something, and that reflection is important for their insight into their own actions. That is in contrast to the training group of which seven of 19 teachers mentioned this. Also the meaning of feedback as holding up of a mirror to students to reduce blind spots was stated clearly less in the control group, namely two of the eight teachers of this group against 11 of 19 teachers of the training group.

The quantitative outcomes reported above concerning the (perceived) results and effects of the training, do not significantly correlate with the teachers' age or their experience as a teacher, nor with the number of years they have worked at their present school.

4.5 Conclusions and discussion

In this study a training programme has been developed for supporting teachers' development of competences they need for supporting students' reflection skills. The programme consisted of three cycles each including the same sequence of activities: a group meeting (lasting 90 minutes), a period of independent practice, an individual meeting per participating teacher (lasting 45 minutes), and a second period of independent practice. The development of the programme has been based on available knowledge (from the scientific literature and experts) about reflection skills (including thinking activities) and teacher activities expected to be relevant for supporting students' reflection skills (including asking questions, coaching and giving feedback), content standards and performance standards for teacher competences in this domain (developed in two earlier studies), information from the target group about needs and circumstances, and literature about teacher learning and professional development activities (including training). In the study 19 teachers from five Dutch nursing schools participated in the training group; eight teachers on other locations formed a control group. The research question was: What are the feasibility, quality and effects of the programme?

With regard to the *feasibility* of the programme it can be concluded that three of the five core features for effective professional development recently described in the literature on which the programme was based, have been realized.

The programme had a strong *content focus*, using knowledge from scientific literature about reflection and thinking activities and about asking questions, coaching and giving feedback, and content standards and performance standards in this domain validated in two earlier studies in cooperation with experts and teachers in the field (Dekker-Groen, Van der Schaaf & Stokking, 2011; 2012). This knowledge was brought in by the trainer as input during the meetings and as part of the feedback she gave the teachers. *Active learning* was realized as teachers actively participated in group meetings, prepared and practised conferences with individual students about student's reflections, and recorded them on video. Afterwards and during individual meetings with the trainer who guided the participants they reflected on their actions. Feedback on video fragments was discussed which provided teachers insights in their current actions and into possibilities for future actions (Joyce & Showers, 2002; Rosenstein, 2002). Activities which teachers increasingly performed were trying all sorts of things and searching for information. *Coherence* was provided by the information and instruction given by the trainer, the exchange of experiences and discussions by the teachers during the group meetings, the connections between the subsequent activities (group meetings, periods of practice in the teachers' own work setting, individual meetings with feedback and reflection on a videotaped conference, the feedback also being based on the standards in the rubrics), and the three cycles building upon each other.

The *time* the teachers could devote to the programme remained modest, however: according to the teachers on average 26 hours (during the 12-15 months the programme lasted per school). One third of these hours were used for attending the meetings, so two thirds may have been used for preparation of and reflection on conferences with students and other activities during periods of independent practise. Possibly these hours provided participants enough time to get used to the content and the standards of the training (Evans, 2008; Fenstermacher, 1987; Ross & Bruce, 2007). Finally, *collective participation* did not happen as team. Due to the turbulent circumstances in nursing education only three to five teachers per team participated in the training. Exchange between these participants during group meetings was restricted, as only limited time for these meetings was available. Consulting a colleague has been the least done learning activity, so this activity will have hardly contributed to the results of the training.

Concerning the *quality* of the programme, according to the participants' evaluations of each meeting directly afterwards and their appraisals after the

programme of a number of aspects regarding its form and content, the quality of the competence descriptions and rubrics used, and the extent to which their own needs and expectations had been fulfilled and the results had been useful, it can be concluded that the participating teachers on average were fairly satisfied with the programme.

Indications for *effects* of the programme in this study have been gathered in five ways. First through the learner reports about the teachers' concerns and activities and the results they perceived of the feedback they received. Second the self-assessed development of knowledge, skills and attitudes during the programme and as a result of the programme. Third the self-estimated development of competences in instructing, giving feedback, and guiding students' reflection, represented by the description of a number of specific activities. Fourth the self-reported extent to which they currently (after the programme) conducted these activities was measured as well as the importance they attached to these activities, to be compared with the teachers in the control group. Fifth the extent to which the students reflect on a number of objects for reflection and the frequency of the teachers giving feedback, and the self-assessment of four key competences, the latter two to be compared with the same measurements both before the programme and in the control group.

Each of these ways of information gathering resulted in some indications that the programme had effects as intended. According to the learner reports the participating teachers on average increased in their focus on the content of students' reflections and the results of their guidance, in trying things and looking for more information, and in self confidence and getting another view on guiding reflection; after the programme the teachers indicated they had developed their knowledge, skills and attitudes during the programme and also that they had developed in conducting a number of specific activities. This increase was on average between 'sufficiently' and 'substantially' ('optimally' being the highest score possible) and the averages on these specific activities were higher and the differences between teachers smaller than on the more global 'knowledge, skills, and attitudes' mentioned before. The means on the scales concerning the teachers' current activities in the training group were consistently, but not significantly, higher than in the control group. The extent to which their students reflect on a number of possible objects of reflection according to the teachers who participated in the programme was after the programme significantly higher than before and, controlled for their respective pre-measurements, also significantly higher than in the control group. These indications of effects obviously do not represent overwhelming evidence, but all effects found were in the intended

direction. In other circumstances, with less turbulence and more time, we estimate that more substantial effects might be expected.

The results fit with the conclusion of Desimone, Porter, Garet, Suk Yoon, and Birman (2002) that competence development focussed on specific instructional practices increases teachers' use of those practices and that a specific feature such as active learning increases the effect. The overall conclusion is that teachers who want actively develop their competences in supporting students' reflection skills and are offered a coherent, purposeful programme including instruction, exercise, and feedback, can improve their relevant practice with a fairly modest investment of their time.

Limitations of the study are that, as the training programme was developed as a comprehensive one deliberately combining a number of factors known in the literature as effective, it is difficult to attribute the results to specific parts or aspects of the programme. Moreover results have to be regarded with some caution because only a small number of teachers participated in the programme due to the turbulent situation in the schools.

Important questions for further study are in what way and to what degree teachers might develop their competences and practices when investing more time and more team collaboration. Also more research is needed into factors which can specifically account for teachers' competence development during training. Finally, further study is desirable concerning the ultimate impact of teachers' improvements in giving feedback and guiding students' reflection on students' development of reflection skills.

Appendix

Example of a teacher competence description and a rubric

Competence 4 Coach the learning process of how to reflect

General description: The teacher knows several potential guiding strategies to stimulate students' reflection skills. Depending on the specific situation and the target group she^a picks the most suitable guiding strategy. She guides students to formulate learning goals for reflection tasks and reflection conferences, based on students' learning needs. She stimulates students to practice reflective thinking activities in coherence. She also teaches students several reflection methods and how to pick a suitable method. She discusses ethical en political dilemmas and pays attention to students' constraints. She teaches students to take a position about events and to evaluate their learning process. She guides students' development in dealing with feedback. Her guiding is contingent upon what students know and are being able to and she fades her guidance if possible.

Rubric descriptions on four performance levels

Levels →	Beginning	Basic	Advanced	Expert
Rubric attributes ↓				
Guidance of reflection	Does not know a strategy	Knows a potential guiding strategy	Applies a guiding strategy	Applies several guiding strategies
Reflection methodology	Does not use a method	Teaches students to use a method	Teaches students to use several methods	Teaches students to choose a method
Development of attitude	Does not pay attention to how people treat each other	Teaches students to consider their own actions and norms	Teaches students to consider personal and professional norms	Discusses ethical and political dilemmas
Guiding reflection	Does not structure students' learning or structures students' learning too much	Structures and controls students' learning if necessary	Models and scaffolds students about what and how to control their learning process	Fades the guidance of students to let them monitor their learning process

^a Where 'she' is used also 'he' can be read.

5 Teachers' questions and responses during teacher-student feedback dialogues⁶

Abstract

In senior secondary vocational education students have to develop competences for reflection to self-regulate their learning and development during their career. Students' reflection and self-regulation can be supported by teachers interacting with students and giving them prompts in the form of questions and responses. To improve students' reflection and self-regulation, research into the relation between teacher's questions and student's responses is relevant.

In this study 46 videotaped feedback dialogues between teachers and individual students were analysed from teachers who participated in a training programme to further their competences in asking questions and giving feedback ($n = 16$) and teachers of a control group ($n = 7$). All teachers' and students' utterances were coded using two coding schemes based on the literature, one for questions (including seven teacher and four student categories) and one for responses (four teacher categories and three student categories).

The study aimed to get more insight into: (1) teachers' repertoire of questions and responses and its relation to students' reactions and (2) the effects of the training programme. The data were analysed using Multi Episode Protocol Analysis and Ancova, focusing on single teacher (T) and student (S) utterances and sequences of two (T-S) and three (T-S-T) utterances.

Teachers did use all question categories studied and students used two out of four question categories. Both teachers and students did use all response categories. On the post-test the repertoire of the teachers in the training group was broader than on the pre-test and also broader than the repertoire of the teachers in the control group when controlled for the pre-test. The training group teachers on average used more hybrid, hybrid prompt and deep reasoning questions and neutral and prompt responses of which prompts and reasoning questions potentially stimulate students to elaborate and reflect. They used less recall and surface reasoning questions. No differences between both groups were found regarding teachers' responses.

⁶ A slightly shorter version of this chapter is submitted as manuscript to an international peer-reviewed journal.

5.1 Introduction

In vocational education students have to develop competences for continuous learning and development during their career. A competence is a capacity to enact specific combinations of knowledge, skills and attitudes needed to adequately perform tasks in jobs and job contexts (Lizzio & Wilson, 2004). Competences for continuous learning and development demand self-regulation. This concerns setting goals, using learning strategies to achieve the goals, monitoring one's performance for progress, restructuring the context and managing one's time, attributing causation to results, and adapting future strategies (Zimmerman, 2002). An essential skill for students to become self-regulated learners is reflection, which is a focused way of thinking. The development of both vocational competences and self-regulating skills, such as reflection, can be stimulated by giving feedback. Feedback is information about how successful students' functioning is and how to bridge the gap between their actual performance and certain learning goals aiming to improve their competences (Hattie & Timperley, 2007; Nicol & Macfarlane-Dick, 2006; Sadler, 1989, 2010). The effectiveness of feedback depends on whether students use the information to become more competent. Prerequisite is that students understand the feedback and interpret it accurately (Van der Schaaf, Baartman, Prins, Oosterbaan & Schaap, 2011).

In interactive learning environments and in problem solving situations students' self-regulation appears to be supported by teachers giving prompts (Davis, 2003; Ifenthaler, 2012; Thillmann, Künsting, Wirth & Leutner, 2009). Prompts are questions and responses which teachers pose and give. Prompts can induce learning strategies for self-regulation, including reflection (Gé & Land, 2004; Wirth, 2009). Little is known about how teachers pose questions and give responses in feedback dialogues to support a student's reflections. Getting more insight into the effects of questions and responses in feedback dialogues is relevant because of two reasons. Firstly, feedback dialogues between a teacher and an individual student are an important pedagogical tool in vocational education. Secondly, results about effective teacher behaviour can be used to train teachers in posing questions to students in feedback dialogues which can support their reflection skills (Aronson, 2011; Branch & Paranjape, 2002).

This study focused on exploring teachers' and students' questions and responses within feedback dialogues and on the effects of a training programme on teachers' repertoire of questions and responses. These dialogues were videotaped as part of a study concerning a teacher training programme in nursing education in the Netherlands (Dekker-Groen, Van der Schaaf & Stokking, 2013). The goal of Dutch senior secondary vocational education is to deliver students to the labour

market as well as to prepare and stimulate them to continue their educational career in higher professional education.

The study aimed to answer two research questions: Which questions and responses are given in feedback dialogues between nursing teachers and individual students? What are the effects of the training programme on the questions and responses of teachers and individual students in feedback dialogues?

5.2 Theoretical framework

Self-regulation and prompting

In feedback dialogues that aim to support the development of students' reflection skills, teachers and students exchange information and teachers support students when necessary. Such dialogues are prone to the so-called 'assistance dilemma' in teacher - student interaction, pertaining to the degree of external support in relation to the student's degree of self-regulation in learning how to succeed (Koedinger & Aleven, 2007; Ifenthaler, 2012). If students experience difficulties in their learning the teacher's additional information and support can be helpful but should be in balance with the students' needs to figure out insights and solutions themselves (Shuell, 1996; Thompson, 2009).

Self-regulation processes consist of three cyclical phases: forethought, performance, and reflection (Zimmerman, 2002). The study concerns students' reflection, the third phase in the model of Zimmerman. In each phase of the model teachers can support students by giving prompts in the form of questions and responses which can be rather generic or specific (Davis, 2003; Ifenthaler, 2012). Generic prompts (e.g., "Which step can help to solve this problem?"), are effective in focusing students' attention to content and strategies they already possess but do not adequately use (production deficiency: Thillmann et al., 2009; Wirth, 2009). A generic prompt guides students' self-regulation and learning. Specific prompts are given in case of a lack in knowledge or skills for self-regulation (availability deficiency: Thillmann et al., 2009; Wirth, 2009). A specific prompt contains contextualised directed information, for example "You may use the protocol". Such prompts can include useful information but it might be necessary to give additional instruction and exercises as well.

The content of prompts should be related to the processes which should be activated (Wirth, 2009). Prompts can be very effective by focusing students on certain cognitive processes and evoking specific reflective activities (Butler, 1998; Chi, De Leeuw, Chiu & LaVancher, 1994; Nückles, Hübner & Renkl, 2009; Van den Boom, Paas & Van Merriënboer, 2007). In feedback dialogues prompts are

often given to direct students' thinking. Examples are: allowing students to focus on one aspect of a task followed by an open-ended question "How does this affect your functioning?", prompting students to think about alternatives "What else can you do in such a situation?" and giving hints in the form of relevant information or by questioning "Why is it important?".

Questions and responses in feedback dialogues

In feedback dialogues teachers' questions and responses constitute teachers' support for students' reflection. Chi, Siler, Jeong, Yamauchi, & Hausmann (2001) did call a dialogue interactive and constructive when questions elicit a response from the student and the student responds on a hint by elaborating on that hint. Acknowledgement responses and continuers such as 'uh-huh' and 'okay' are examples of interactive but non-constructive responses, and lengthy explanations by a tutor are called non-interactive and non-constructive. The latter is in line with Chin (2006) who stated that students are challenged the most by teacher-led conversations but not by teacher-dominated ones. In this study the focus is on the support teachers give to students during reflection. It is assumed that teachers optimally support students' reflection when the students use reflective formulations they would not use without the support (scaffolding, Collins, Brown & Newman, 1989; zone of proximal development, Vygotsky, 1978).

To enhance teacher-student interaction teachers should give students in the dialogue enough room to talk (Blatt, Confessore, Kallenberg & Greenberg, 2008). Also it is important that teachers pose questions. The function of a question is to elicit a verbal response from the person to whom the question is addressed (Kearsley, 1976). This response will depend on the viewpoint of the receiver. As a consequence there is not a one-to-one relation between a question and response. The question receiver determines how a question is conceived and which response will be given. The response depends on the social context and the foregoing questions and responses in the dialogue (Thomas Farrar, 1986).

In a feedback dialogue a response can be an answer to a question, a statement, or a comment (Chin, 2006). Each of these three response types can be a positive, neutral, or negative utterance or a corrective response (Blatt et al., 2008). In dialogues both the teacher and the student can react with an acceptance (positive), a denial (negative), or a neutral utterance. The fourth option, a corrective response, corresponds with corrective feedback (i.e. Chi, 1996) and has a function with respect to the content of the feedback. In addition, teachers can give prompt responses, for example a hint to activate students (e.g. Wirth, 2009).

Questions in feedback dialogues can be explicitly stated as a question, for example "Did you consider that?" or as a statement which is posed as if it was a question,

for example “You did not think about it by then?” It can be assumed that in dialogues to support nursing students’ reflections the teachers’ questions are effective (Branch & Paranjape, 2002). Question categories that are often used in dialogues in other educational contexts might be useful as question categories in this study. Prompt questions, discussed above, are such a question category.

Based on their research about Computer Stimulated Collaborative Learning, Erkens and Janssen (2008) distinguished three question categories which are often used within dialogues: open, closed, and set questions. An open question can be answered in one or more sentences, for example “How would you tackle this next time?”. Closed questions can be answered in few words, for example by means of a number, a name, or verification: “How many tasks do you still have to do?”, “Where is your internship?”, “Is Anthony your supervisor?”. A set question means that two questions are posed in one utterance, for example “Does it concern activities during your internship or in lessons at school?”. A set question contains two questions at the same time and mostly only one of them will be answered.

Chin (2006) used a categorisation of open and closed questions based on the function of the questions. According to Chin questions are closed if they focus at information gathering, while open questions are meant to stimulate a dialogue or a discussion or to get the receiver’s view. The categorisation of Chin links two categories that are often used in a learning situation. One category concerns questions aimed at the recall of information (lower order questions) and the other concerns questions that focus on the development of a critical thinking process (higher order questions). More differentiated than Chin’s dichotomy lower-higher are the categories in the taxonomy of Bloom, elaborated by Krathwohl (2002): remember, understand, apply, analyse, evaluate, and create.

Graesser and Person (1994) used the Graesser-Person-Huber question scheme in which the names of question categories refer to the kind of responses a teacher expects. Besides requests and hybrid questions, they distinguished short and long answer questions. An example of a request is “Please tell me about that”. A hybrid question combines two or more question categories. For example the question “Was it difficult for you?” is a closed question which can be answered with ‘Yes’ or ‘No’ (short answer) but can also be intended to stimulate reasoning as a result of which a student may be encouraged to give a longer answer. Questions that intend a short answer (the so called short answer questions) are for example questions for verification (verification questions), e.g., “Did you do the assignment?”, questions about quantitative data, e.g., “How many days did you miss of your internship?” and questions about feature specifications, e.g., “What does the protocol says about it?” To answer such ‘short answer questions’ it is intended that students have to rely on their memory; i.e. recall of information

(Chin, 2006). As this need for recall clarifies the function of the category, in stead of the name 'short answer question' we from now on will use 'recall question'. Greasser and Person (1994) spoke of 'long answer questions' when an answer longer than one word or one sentence is expected, for instance questions which focus on how something happened or what made that something went well or wrong. Within long answer questions they distinguished two types. The first type concerns questions posed to get a definition, an example, a comparison, an interpretation, or a judgment. The function of this question type seems to be to diagnose students' knowledge and to get their opinions in cases that few reasoning is necessary, e.g., "How are you going?" Therefore we will use the label 'surface reasoning' for the first long answer type. The second long answer question type concerns questions asked to get information about cause and effect, purposefulness, or expectations, and instrumental or procedural questions. These questions are meant to let a student elucidate patterns of deep reasoning. From now on we will use the label 'deep reasoning questions' for the second type long answer questions. We suppose that deep reasoning questions, e.g., "How do you think to handle this next time?" aim to evoke reflection because they intend to stimulate thinking about experiences and processes in order to learn from them. Considering the potential benefit of prompting for self-regulated learning in interactive learning environments (Ifenthaler, 2012; Thillmann et al., 2009) a combination of prompt questions and certain of the question categories of Greasser and Pearson (1994) might be fruitful to induce students to use learning strategies which they do not spontaneously use (Berthold, Nückles & Renkl, 2007; Gé & Land, 2004; Wirth, 2009). The use of such combinations of different question categories, namely a hybrid question, e.g., "Are we the doctors' support?" and a long answer question including a prompt which we call 'Reasoning prompt', e.g., "How do you expect others think of you?" seem desirable to match the complexity of questions and their meanings (Thomas Farrar, 1986).

Sequences of teacher and student questions and responses

In the literature about dialogues different types of sequences are distinguished concerning the function that questions and responses can fulfil. An example is the sequence: teacher Initiation, student Response, teacher Evaluation (IRE sequence, Van Zee & Minstrell, 1997). In this sequence teachers' information-seeking questions are often meant to understand students (Flammer, 1981). Teachers can also start a dialogue to give feedback. In that case the sequence is: teacher Initiation, student Response, teacher Feedback/Follow-up (IRF sequence, Mercer, 2004; Chin, 2006). When this sequence starts with a question of the teacher followed by an answer of a student the IRF sequence is: teacher Question, student

Answer, teacher Feedback (QAF pattern). A succession of five utterances can be formed by using a dialogue frame of five steps in which for example a tutor poses deep-reasoning questions, gives short feedback, and asks questions which becomes more specific (Chi et al., 2001).

In the studies referred to before the focus was on the teachers' or tutors' behaviour in a lesson situation and on questions related to the course content and not to the students' reflection on their development. In our study the focus is on teachers' behaviour in dialogues with individual students. Both the teacher and the student may pose questions and give responses. The dialogue is a process in which teacher utterances and student utterances follow each other continuously. A certain sequence of utterances is called a pattern if certain utterances follow each other significantly often in the same order (Jeong, Clark, Sampson & Menekse, 2011). According to Flammer (1981) a teacher's utterances can be seen as interventions because the teacher uses the student's information to pose questions, for example: "Which alternative solutions might be successful in such circumstances?" Moreover teachers give responses to support students' reflection. To get insight in the role of teachers in feedback dialogues we will analyse two types of sequences, all starting with a teacher utterance: two-sequences (Teacher utterance - Student utterance, to be abbreviated as T-S) and three-sequences (Teacher utterance - Student utterance - Teacher utterance, T-S-T). For both the teacher's and the student's utterances question categories and response categories will be distinguished.

5.3 Methods

Design

We used a quasi experimental pre-post test design including an experimental group and a control group to explore teachers' questions and responses during feedback dialogues and to examine the effects of a teacher training in using prompts, asking deep reasoning questions and giving feedback (see Figure 5.1). Both the pre-test and the post-test consisted of the registration of a feedback dialogue of a teacher with an individual student. The time lag between the pre-test (t1) and post-test (t3) was between 12 and 15 months.

	t1	t2	t3
Experimental group	O1	X	O2
Control group	O3		O4

Figure 5.1 Quasi experimental pre-post test control group design

The experimental group participated in a training programme (t2). From now on we will call this group the 'training group'.

Intervention

The intervention was a teacher training programme for supporting nursing students' reflection skills development (Dekker-Groen, Van der Schaaf & Stokking, 2013). In three group meetings of 90 minutes each, information and instruction were given about the function of questioning, general and specific questions, question types, and their relation with feedback and reflection. Moreover video fragments of representative feedback dialogues between teachers and students were used as a means for the teachers to discuss with colleagues and to exchange experiences (Borko, Jacobs, Eiteljorg & Pittman, 2008; Van Es & Sherin, 2008). Each group meeting was followed by a period of independent practice of new behaviour in interaction with students. In individual meetings of 45 minutes each, video interaction analysis was used to give teachers insight into their actions (Caris-Verhallen, Kerkstra, Bensing, & Grypdonck, 2000; Roter et al., 2004). The researcher, also a video interaction coach, discussed with each teacher video fragments of their own feedback dialogues with a student and guided the teacher to reflect on experiences and to formulate new goals. After each individual meeting there was another period of independent practice.

Participants

Six schools of senior secondary nursing education participated in the study. These schools were spread over the country. A purposeful sample of teachers for the training group was composed based on the selection criteria of being interested in developing their own competences and having enough time to participate in the training. For both the training group and the control group further selection criteria were: being responsible for teaching and guiding students for a minimum number of 20 hours per week and not being close to retirement. As a result 19 teachers joined the training programme and eight teachers formed the control group. Three teachers of the training group and one teacher of the control group could not participate in the post-test measurements due to departure to another job and/or personal reasons. Therefore two feedback dialogues (pre and post) of 16 teachers of the training group and seven teachers of the control group could be analysed.

The training group consisted of 12 female and four male teachers, on average 51 years old and with 13 years of teaching experience. The teachers in the control

group were all female, on average 47 years old and with 16 years of teaching experience. The mean differences in age and experience between the two groups were not significant.

The teachers of the control group worked on other locations of the same institutes as the teachers of the training group. The control group however did not receive any information about the research questions nor about the training programme and did not have contact with the teachers from the training group. In the schools with different locations the education did not differ between locations because all teachers of the same school used the same methods and same materials. All teachers were asked to choose a student to participate in an individual feedback dialogue. Students who participated were in their second year of nursing education, mostly female (per group 79% or more), and on average 20 years old. As during the training programme, which lasted on average one school year, the students were assigned to other teachers as their mentor, the participating teachers had to ask other students, but comparable in age and experience, to participate in the post-test dialogue. The students on the post-test were other students than in the pre-test, but not significantly different on gender, age and experience.

Instrumentation

Coding schemes

Based on the literature described before, two schemes were developed for coding activities concerning teachers' and students' questions and responses in the feedback dialogues. Also a coding map was developed in which we included a description of and an example for each code and the coding rules. If for example the same utterance contained both a question and a response only the question was coded. Also the context of questions and responses had to be taken in account in order to code the purpose of the person asking the question. Due to this coding map it was clear what was understood by a code and when a code had to be used (Chi, 1997).

Question categories in both the teacher and the student coding scheme were based on the Graesser-Person-Huber scheme described before (Graesser & Person, 1994). In the teacher scheme these question categories were combined with prompt questions (see e.g. Berthold et al., 2007; Davis, 2003; Gé & Land, 2004; Ifenthaler, 2012). Seven categories were distinguished: (1) Recall, (2) Hybrid, (3) Hybrid prompt, (4) Request, (5) Surface reasoning, (6) Deep reasoning, (7) Reasoning prompt. The choice of coding quite a number of different teacher question categories is in line with Thomas Farrar (1986) who states that it seems

desirable to match the complexity of questions and their meanings in the social context and their relations with the kind of dialogue.

As response categories the following three coding categories were used for both teachers' and students' utterances: (1) accept (positive), (2) deny (negative), and (3) neutral response (Blatt et al., 2008; Erkens & Janssen, 2008). The neutral response was also used when an affirmation or a critical remark was made less strong (e.g., "I agree but ..."). In the teacher coding scheme the category (4) prompt (prompt response) was added (see e.g. Davis, 2003; Gé & Land, 2004; Nückles et al., 2009). An example of this response category is "So you will have the propensity to take responsibilities but people around you also will".

See Appendix A for the coding scheme of question categories and Appendix B for the coding scheme of response categories.

Segmentation and interrater agreement

All videotaped feedback dialogues were transcribed verbatim (Mercer, 2004). Next, the transcripts were imported in the program Multi Episode Protocol Analysis (Erkens, 2005). Speech turn-taking was used as the segmentation criterion in each transcript, because this fits both the natural course of the conversation and the research questions (Chi, 1997). Moreover, turn-taking is a unit of segmentation which can be objectively applied. All words spoken in one turn will be called an utterance, here to be called a segment. Each segment of a teacher and of a student was coded with a question category or a response category.

Two research assistants independently coded 100-102 segments randomly chosen of three conversations (in total 302 allotted codes; 150 to teachers' segments and 152 to students' segments) each of a different feedback dialogue between a teacher and a student who did not participate in the study but were from the same schools as in the sample for the study.

Interrater agreement and percentage agreement were computed. Cohen's Kappa for the teachers' segments was 'fair' (.56) and for the students' segments 'good' (.74) (Banerjee, Capozzoli, McSweeney, & Sinha, 1999; De Wever, Schellens, Valcke & Van Keer, 2006; Strijbos & Stahl, 2007). Percentage agreement of the allotted codes for the teachers was 63% and for the students 85%. Next the research assistants independently coded all 46 feedback dialogues. This concerned in total 10185 segments.

Data collection

In nursing education teachers regularly have feedback dialogues with individual students. Data was gathered in the form of videos of feedback dialogues of the

teachers to support individual students' reflection skills. Two feedback dialogues per teacher were videotaped, one before and one after the intervention. In the training group this took place in the summer of 2009 respectively in the summer and autumn of 2010. In the control group for each teacher the length of the period between the two dialogues was similar. All teachers got the same written request to videotape a feedback dialogue with a student who participated in their mentor group. The teachers asked the students' consent about the use of the videotape for research purposes.

In both teacher groups (training and control) only one feedback dialogue in the pre-test concerned feedback on how successful a practical task in the curriculum was done, all the other dialogues were about the student's progress in the study and during the internship. All teachers used the student's portfolio or a written reflective report as input for the feedback dialogue. In the post-test the feedback dialogues of all teachers were about the progress in the study and the internship. In this post-measurement one-third of the teachers in the training group ($n = 12$) used their student's written answers on a reflection task as input. They had given their students this reflection task after the training programme was done. Two-thirds of the teachers in the training group used their student's written report just as in the first videotaped dialogue. In the control group all teachers used their student's portfolio or a written reflective report as input for the feedback dialogue, just as in the first videotaped dialogue. The videotaped feedback dialogues in the training group lasted on average 18 minutes in the pre-test and 24 minutes in the post-test. For the control group the duration was on average 23 minutes in the pre-test and 27 minutes in the post-test.

Data analysis

Means and standard deviations of the number and duration of utterances in the videotaped feedback dialogues were computed. Differences in number and duration of utterances and in words between the training group and the control group were computed and tested for statistical significance using independent two-tailed *t*-tests, and within each condition between pre-test and post-test by using two-tailed paired *t*-tests. In addition, for the significant results the effect sizes were computed.

In MEPA all utterances per feedback dialogue were coded. Utterances could be questions or responses. Next the frequencies of the teacher utterances and the student utterances were computed separately for the training group and the control group. To control for differences in number of utterances in the dialogues

between the groups and within both groups between pre-test and post-test also percentages were calculated.

Sequences of two utterances were analysed to determine how often a certain teacher utterance was followed by a certain student utterance (T-S sequence). We chose a teacher utterance as starting point because this study focused on the teachers' activities. In MEPA four different scores of these T-S sequences were computed: Fr-score (succession frequency), E-score (expected succession frequency), Z-score and K-score (Kappa score). The Z-score was computed to determine whether the difference between the frequency and the expected value of a certain student utterance which followed on a certain teacher utterance was statistically significant (Jeong, 2005). K-scores are corrected Z-scores. In the computation of the K-score a correction is made for differences in the number of segments within and between groups, resulting in scores between -1 and 1 (Wampold & Margolin, 1982). Negative K-scores indicate that the sequence occurs less often than expected, positive scores indicate that the sequence occurs more often than expected.

Additionally analyses of sequences of three utterances were computed: Teacher utterance-Student utterance-Teacher utterance (T-S-T sequence) (Chi et al., 2001; Chin, 2006). See Appendix C for an example of how 17 successive utterances were divided in sequences and then coded.

A large number of possible sequences occurred (very) infrequently. Therefore we decided to use a threshold for further analysing and reporting the results. For analysing and reporting about utterances and sequences of utterances we used a minimum total frequency of 16 in the training group and of seven in the control group to ensure that in both groups on average the utterance or sequence occurred at least once per dialogue. This threshold should be reached on at least one of the two moments of measurement (pre-test and post-test). In case the threshold is reached for questions or responses which were seen as important in literature or which were explicitly trained, short descriptive results will be added.

To analyse the effects of the training Ancova was used to control for differences in pre-test percentages of teacher utterances between the training group and the control group (Huck & McLean, 1975). Ancova was also used for analysing differences in percentages of teacher - student sequences (T-S) and teacher-student-teacher sequences (T-S-T). Because the sample was small and the data were not normally distributed these analyses on the differences between the two groups were preceded by a bootstrap procedure (Field, 2009). Bootstrapping was carried out by creating 2000 bootstrap samples with replacement. Besides the F-value of the Ancova both the p -value and partial η^2 of the bootstrap were computed.

5.4 Results

5.4.1 Minutes, turns and words in the dialogues

The duration of the videotaped feedback dialogues in the training group in the pre-test was between 6 and 31 minutes and in the post-test 13-38 minutes. For the control group the duration of the dialogues in the pre-test was 8-36 minutes, and in the post-test 15-43 minutes. See Table 5.1 for the means and SDs. Two-tailed paired *t*-tests revealed a significant difference between pre-test and the post-test of the training group ($t(15) = -2.87, p = .012, \eta^2 = .36$).

Due to the choice for turn-taking as the segmentation criterion the numbers of teacher and student turns are (almost) the same. The teachers always initiated the dialogues. The number of teacher turns in the dialogues in the training group in the pre-test was between 38 and 300 and in the post-test 75-517. For the control group the number of turns in the pre-test was 44-562 and in the post-test 207-539. See Table 5.1 for the means and SDs. Two-tailed paired *t*-tests on the number of turns between the pre-test and the post-test of the training group revealed a significant difference ($t(15) = -2.77, p = .014, \eta^2 = .34$).

Table 5.1

Mean and Standard Deviation of duration, # turns, and # words in the feedback dialogues

	Pre-test				Post-test			
	Training group		Control group		Training group		Control group	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Duration in minutes	17.4	7.5	22.6	10.6	24.1	8.5	26.6	8.9
# Turns	155.4	72.2	254.6	171.9	249.6	131.1	274.7	120.3
# Turns per minute	9.2	3.4	10.8	4.6	10.1	4.1	10.5	2.5
# Teacher words per turn	19.2	22.1	15.3	19.5	14.4	16.2	17.0	22.2
# Student words per turn	17.1	22.8	14.2	21.5	18.2	26.2	15.1	23.2

The average number of turns per minute in the training group in the post-test was on average one turn larger than in the pre-test (9.2 respectively 10.1). In the control group the average number of turns in the post-test was one third turn less than in the pre-test (10.8 respectively 10.5). These differences were not statistically significant. The number of words per turn of teachers of the training group was on average five words less in the post-test, while their students used one word more in the post-test. The number of words per turn of teachers of the control group was on average two words more in the post-test, and their students used one word more in the post-test. Two-tailed paired *t*-tests on the average number of words the teachers in the training group used per turn between pre-test and post-test revealed a significant difference ($t(15) = 2.13, p = .04, \eta^2 = .23$).

The number of words per minute used of teachers of the training group in the post-test (not shown in Table 5.1) was nine words less than in the pre-test (on average 85 respectively 76 words), for their students in the post-test it was 21 words more than in the pre-test (on average 75 respectively 96 words). The teachers and students of the control group used about the same number of words in the pre-test and in the post-test, the teachers on average 86 respectively 88 words per minute and their students on average 80 respectively 78.

5.4.2 Frequencies of teacher and student utterances

Teacher utterances

For the coding of the teacher utterances in total 11 categories were distinguished (see Appendices A and B). In both groups the teachers used a Request on average less than one time per feedback dialogue. See Table 5.2 for the frequencies and percentages of all other teacher categories for which on average the teachers used the category at least one time per dialogue.

In the pre-test the number of questions the teachers of the training group used was about the same as the number of responses (utterances other than questions), of which they used Recall questions and Accept responses the most. In the post-test 48% of their utterances were questions and 52% were responses. In the post-test Hybrid prompt, Deep reasoning, and Reasoning prompt questions and Accept and Prompt responses were used more, while Recall, Hybrid, and Surface reasoning questions and Deny and Neutral responses were used less (see the percentages in Table 5.2).

In the pre-test in the control group 25% of teachers' utterances were questions, mostly Recall questions, and 75% were responses, the most of which were Neutral. In the post-test 31% of their utterances were questions and 69% were responses. In the post-test Recall and Surface reasoning questions were used more and the other question categories were used less, while Prompt responses were used less and the other response categories were used more (see Table 5.2).

Analyses of responses of one word utterances showed that the teachers used four single words on average at least one time in a feedback dialogue in the pre-test and/or post-test, see 'One word utterance' in Table 5.2. The teachers used 'Yes' far the most. The teachers of the training group used 'Yes' four times as often in the post-test compared to the pre-test. The teachers hardly used 'No' and 'Okay' responses. The teachers of the training group used '(Uh)mhm' and 'Okay' only in the post-test (25 respectively 39 times), while the teachers of the control group did not use '(Uh)mhm' at all and 'Okay' on average about two to three times per dialogue.

Table 5.2
Frequencies and percentages of teachers' utterances^a

	Pre-test				Post-test			
	Training		Control		Training		Control	
	Freq	Perc	Freq	Perc	Freq	Perc	Freq	Perc
Teacher question								
Recall	230	18.4	131	14.7	269	13.4	209	21.7
Hybrid	124	9.9	37	4.2	191	9.6	33	3.4
Hybrid prompt	63	5.0	11	1.2	114	5.7	8	0.8
Request	(7)	(0.6)	(1)	(0.1)	(13)	(0.7)	(0)	(0.0)
Surface reasoning	122	9.8	32	3.6	135	6.8	48	5.0
Deep reasoning	71	5.7	8	0.9	173	8.7	(4)	(0.4)
Reasoning prompt	(11)	(0.9)	(5)	(0.6)	69	3.5	(0)	(0.0)
Teacher response ^b								
Neutral	176	14.1	371	41.6	463	23.1	389	40.4
Deny	(2)	(0.2)	9	1.1	(2)	(0.1)	10	1.0
Accept	356	28.5	261	29.3	412	20.6	251	26.0
Prompt	87	7.0	25	2.8	160	8.0	12	1.2
Total of questions and responses ^c	1229	98.4	885	99.4	1986	99.4	960	99.5
One word utterance								
Yes	47	83.9	106	82.2	198	71.2	71	63.4
No	(3)	(5.4)	8	6.2	16	5.8	18	16.1
(Uh)mhm	(0)	(0.0)	(3)	(2.3)	39	14.0	(4)	(3.6)
Okay	(6)	(10.7)	12	9.3	25	9.0	19	17.0
Total of one word utterances ^c	47	83.9	126	97.7	278	100.0	108	96.5

^a Frequencies which did not reach the threshold (for training group at least 16 and for the control group at least 7) and their percentages are between parentheses.

^b Teacher utterance other than a question.

^c These totals are without the numbers given above between parentheses.

Student utterances

For the coding of the students' utterances in total seven categories were distinguished (see Appendices A and B). In both the pre-test and the post-test the students of both groups in the experiment hardly asked questions. Only 2.2-5.3% of students' utterances were questions, three to seven questions per dialogue. These questions were Recall questions (82% or more) and Surface reasoning questions. Students especially used Accept and Neutral responses (both 40-50%) and some Deny responses. Analyses of the utterances 'Yes', 'No', '(Uh)mhm' and 'Okay' showed that students did use about twice as much such responses than their teachers. They used the utterance 'Yes' far the most. Students of the training group used 'Yes' 164 times in the pre-test and 411 times in the post-test, students in the control group did this 187 respectively 158 times. Students used 'No' on average one time per dialogue in the pre-test and three times in the post-test. Students of the training group used 'Okay' on average less than one time per dialogue and students of the control group on average about three times.

5.4.3 Sequences of categories in the dialogues

Teacher - Student (T-S) Sequence

Each T-S sequence consists of a pair of two utterances (questions or responses) starting with a teacher utterance followed by a student utterance. Eleven different teacher categories and seven different student categories allowed for 77 different sequences. About 20% of these sequences occurred on average at least one time per dialogue in the training group, of which 15 sequences in both the pre-test and the post-test and another sequence only in the post-test. For the control group this was about 14% of the sequences, namely 11 sequences both in the pre-test and the post-test, all 11 belonging to the group of 15 sequences mentioned before with regard to the training group. These 16 (training group) and 11 (control group) different T-S sequences in the training group and the control group consisted of combinations of nine respectively six teacher categories combined with three student categories: Student Accept 6 respectively 5 times, Student Neutral 9 respectively 5 times and Student Deny in both groups one time. See Table 5.3 for the 15 T-S sequences which included a Student Accept or a Student Neutral response. Besides frequencies (Fr) per sequence per group (training and control) per measurement moment (pre-test and post-test) also the expected values (row total of 9 rows x column total of 2 columns / overall total) (E), Z-scores (Z) and their adjusted scores (K: Kappa values) are given.

In both groups in both the pre-test and the post-test the teachers mostly asked a Recall question, mostly followed by a Student Accept response. Students in the training group responded with Deny on average two times in the pre-test and one time in the post-test, those in the control group did this three times respectively four times. Other student response categories were used still less.

Besides Recall questions the teachers in the training group used all other question categories. All student responses in the training group on Reasoning questions were Neutral. The second most occurring sequence in the training group was Teacher Surface reasoning question - Student Neutral response in the pre-test and Teacher Deep reasoning question - Student Neutral response in the post-test. The teachers in the control group besides Recall questions only used Hybrid and Surface reasoning as question categories. In the control group the sequence Teacher Recall question - Student Neutral response occurred second most in both the pre-test and the post-test.

Table 5.3
Frequencies, Expected Frequencies, Kappa- and Z- scores of T-S sequences (question - response)^a

	Student response							
	Accept				Neutral			
	Pre-test		Post-test		Pre-test		Post-test	
	Training	Control	Training	Control	Training	Control	Training	Control
Teacher question								
Recall	Fr=117 E=77.57 Z=6.40** K=.36	Fr=75 E=57.85 Z=3.62** K=.37	Fr=174 E=108.41 Z=8.99** K=.47	Fr=110 E=87.86 Z=3.81** K=.27	Fr=71 E=110.42 Z=-6.40** K=-.51	Fr=29 E=46.15 Z=-3.62** K=-.30	Fr=75 E=140.59 Z=-8.99** K=-.60	Fr=59 E=81.14 Z=-3.81** K=-.25
Hybrid	Fr=53 E=44.57 Z=1.73* K=.13	Fr=17 E=19.47 Z=-0.86 K=-.16	Fr=70 E=76.19 Z=-0.99 K=-.06	Fr=18 E=15.08 Z=1.11 K=.21	Fr=55 E=63.43 Z=-1.73* K=-.19	Fr=18 E=15.53 Z=0.86 K=.13	Fr=105 E=98.81 Z=0.99 K=.16	Fr=11 E=13.92 Z=-1.11 K=-.19
Hybrid prompt	Fr=18 E=22.28 Z=-1.21 K=-.14	(Fr=5)	Fr=38 E=44.85 Z=-1.40 K=-.12	(Fr=5)	Fr=36 E=31.72 Z=1.21 K=.19	(Fr=5)	Fr=65 E=58.15 Z=1.40 K=.15	(Fr=3)
Surface reasoning	(Fr=10)	(Fr=2)	(Fr=4)	(Fr=5)	Fr=101 E=65.20 Z=7.27** K=.78	Fr=28 E=13.31 Z=5.50** K=.88	Fr=123 E=71.70 Z=9.50** K=.93	Fr=42 E=22.57 Z=5.84** K=.80
Deep reasoning	(Fr=5)	(Fr=2)	(Fr=9)	(Fr=0)	Fr=63 E=39.94 Z=5.86** K=.82	(Fr=6)	Fr=156 E=93.16 Z=10.32** K=.87	(Fr=4)
Reasoning prompt	(Fr=2)	(Fr=1)	(Fr=7)	(Fr=0)	(Fr=8)	(Fr=3)	Fr=58 E=36.70 Z=5.42** K=.75	(Fr=0)
Teacher response ^b								
Accept	Fr=58 E=68.09 Z=-1.73* K=-.10	Fr=80 E=134.61 Z=-8.42** K=-.51	Fr=112 E=174.16 Z=-7.05** K=-.28	Fr=75 E=110.05 Z=-6.81** K=-.40	Fr=107 E=96.91 Z=1.73* K=.15	Fr=162 E=107.39 Z=8.42** K=.41	Fr=288 E=225.84 Z=7.05** K=.36	Fr=154 E=109.95 Z=6.81** K=.37
Neutral	Fr=168 E=138.65 Z=3.89** K=.15	Fr=250 E=195.79 Z=7.70** K=.35	Fr=302 E=196.37 Z=11.48** K=.41	Fr=224 E=186.11 Z=5.26** K=.22	Fr=168 E=197.35 Z=-3.89** K=-.21	Fr=102 E=156.21 Z=-7.70** K=-.28	Fr=149 E=254.63 Z=-11.84** K=-.54	Fr=134 E=171.89 Z=-5.26** K=-.20
Prompt	Fr=32 E=33.84 Z=-0.43 K=-.04	Fr=19 E=13.35 Z=2.36* K=.53	Fr=111 E=67.49 Z=7.35** K=.50	Fr=8 E=6.24 Z=1.03 K=.31	Fr=50 E=48.16 Z=0.43 K=.05	(Fr=5)	Fr=44 E=87.51 Z=-7.35** K=-.64	(Fr=4)
Total ^c	Fr=463	Fr=455	Fr=829	Fr=445	Fr=659	Fr=363	Fr=1075	Fr=411

^a Frequencies which did not reach the threshold (for training group at least 16 and for the control group at least 7) are between parentheses.

^b Teacher utterance other than a question.

^c These totals are without the numbers given above between parentheses.

* $p < .05$, ** $p < .01$.

Of the T-S sequences starting with a teacher response in the pre-test the training group showed most Teacher Neutral - Student Accept and Teacher Accept - Student Neutral sequences. In the post-test they showed far most Teacher Neutral - Student Accept sequences. In the post-test they showed the Teacher Prompt - Student Accept sequence more than three times as much as in the pre-test. In both the pre-test and the post-test the control group showed Teacher Neutral - Student Accept sequence far the most, followed by Teacher Accept - Student Neutral. These teachers hardly used Prompt responses.

Teacher-Student-Teacher (T-S-T) Sequence

Each T-S-T sequence consists of three utterances (questions or responses) starting with a teacher utterance followed by a student utterance and ending with a teacher utterance. In the pre-test of the training group 20 different T-S-T sequences, together 514 sequences, occurred on average at least one time per dialogue (the threshold). For the control group this concerned 25 different T-S-T sequences, together 672 sequences. In the post-test of the training group 36 different T-S-T sequences, together 1075 sequences reached the threshold. For the control group this concerned 29 different T-S-T sequences, together 750 sequences. See Table 5.4 for the frequencies of combinations of teacher questions and teacher responses in the 46 different T-S-T sequences (the S consisting of an Accept or Neutral Student response) which showed up on average at least one time per dialogue in the training group and/or the control group.

Five other different T-S-T sequences which also reached the threshold but did occur least frequently are not shown in Table 5.4 to avoid making the table overcrowded. Sequence T Recall - S Deny - T Recall occurred in the pre-test of the training group (16 times) respectively the control group (7 times) and in the post-test of control group (10 times). Another four sequences occurred in the control group of which two sequences were combinations of T Accept followed by S Recall and respectively T Accept and T Neutral (used in the pre-test and post-test 7-9 times). T Recall - S Deny - T Accept was used 9 times in the post-test. The fifth T Recall - S Recall - T Neutral was used 7 times in the post-test.

In the post-test compared to the pre-test in the training group 20 different new T-S-T sequences reached the threshold. The teacher categories concerned: (1) four different sequences containing two questions: Recall and Hybrid, Surface reasoning used twice, Deep reasoning and Recall, Deep reasoning used twice, (2) twelve different sequences containing one teacher question and one teacher response of which the questions were: Recall (two times), Hybrid and Hybrid prompt (both three times), Deep reasoning (four times), (3) four different sequences of two responses: Prompt used twice and three sequences in which

Prompt was the second teacher utterance. In the control group in the post-test compared to the pre-test six different new sequences reached the threshold which all contained one or two T Recall and T Surface reasoning questions.

Table 5.4
Frequencies of teacher questions and responses in T-S-T sequences

			Student response							
			Accept				Neutral			
			Pre-test		Post-test		Pre-test		Post-test	
			Train- ing	Con- trol	Train- ing	Con- trol	Train- ing	Con- trol	Train- ing	Con- trol
Teacher question	S ^b	Teacher question								
Recall	S ^b	Recall	23	22	30	27				17
Recall	S ^b	Hybrid			16					
Recall	S ^b	Surface reasoning	17		18	8				
Surface reasoning	S ^b	Recall					22		18	11
Surface reasoning	S ^b	Hybrid					18		19	
Surface reasoning	S ^b	Surface reasoning							17	8
Deep reasoning	S ^b	Recall							22	
Deep reasoning	S ^b	Deep reasoning							20	
Teacher question	S ^b	Teacher response ^a								
Recall	S ^b	Accept		17	27	26		11	17	18
Recall	S ^b	Neutral	35	24	41	40	23	8		15
Hybrid	S ^b	Accept							26	
Hybrid	S ^b	Neutral		7			19		20	
Hybrid prompt	S ^b	Accept							17	
Surface reasoning	S ^b	Accept					16	11	24	11
Surface reasoning	S ^b	Neutral						7		9
Deep reasoning	S ^b	Accept							35	
Deep reasoning	S ^b	Neutral							29	
Teacher response ^a	S ^b	Teacher question								
Accept	S ^b	Recall				15	22	33	24	26
Accept	S ^b	Hybrid						7	25	
Accept	S ^b	Hybrid prompt							19	
Accept	S ^b	Surface reasoning						8		
Accept	S ^b	Deep reasoning							25	
Neutral	S ^b	Recall	29	29	29	37	26	9		29
Neutral	S ^b	Hybrid	18	7	28	10			16	
Neutral	S ^b	Hybrid prompt			23					
Neutral	S ^b	Surface reasoning				10				
Neutral	S ^b	Deep reasoning			20					
Teacher response ^a	S ^b	Teacher response ^a								
Accept	S ^b	Accept		20		11	23	76	109	68
Accept	S ^b	Neutral	21	48	34	42	24	29	51	49
Accept	S ^b	Prompt							19	
Neutral	S ^b	Accept	16	41	38	32	28	28	33	41
Neutral	S ^b	Neutral	60	159	109	128	58	46	41	54
Neutral	S ^b	Prompt			28					
Prompt	S ^b	Neutral			28					
Prompt	S ^b	Prompt			21					
		Total	219	374	490	386	279	273	626	356

^a Teacher utterance other than a question.

^b Student utterance is an Accept or Neutral response, see the column headings.

5.4.4 Effects of the training on the occurrence of utterances and sequences

Ancova analyses were conducted on the data of question and response categories percentages of single teacher utterances and sequences which were used on average at least one time per dialogue in both groups on at least one of the two moments of measurement (pre-test and post-test). This threshold was not reached for two question categories: Request (both groups) and Reasoning prompt (Control group) and for one response category: Deny (Training group). See Table 5.5 for the results of teacher utterances which reached the threshold. Six of the eight Ancovas revealed a significant difference between the training and control group on the post-test, controlled for the pre-test.

Table 5.5

Ancovas with bootstrap of teacher utterances in the dialogues, post-test Mean % and SD, corrected for pre-test differences

Teacher utterance	Training	Control	F	Bootstrap	
	M% (SD)	M% (SD)		p	Part η^2
<i>Teacher question</i>					
Recall	13.38 (5.69)	23.16 (4.15)	15.61	.00*	.44
Hybrid	9.62 (5.17)	3.50 (2.92)	5.02	.03**	.20
Hybrid prompt	6.44 (3.57)	.94 (.61)	13.06	.00*	.41
Surface reasoning	7.96 (8.42)	6.07 (5.52)	0.46	.49	.02
Deep reasoning	9.43 (4.42)	.88 (1.76)	9.12	.00*	.35
<i>Teacher response^a</i>					
Accept	18.22 (9.58)	24.89 (7.49)	0.62	.40	.03
Neutral	22.0 (8.17)	38.0 (6.24)	18.66	.00*	.48
Prompt	7.86 (5.30)	1.70 (2.18)	3.96	.01**	.18

* $p < .01$, ** $p < .05$.

^aTeacher utterance other than a question.

Of the T-S sequences the threshold was not reached by both groups for Surface reasoning - S Accept. Further the threshold was not reached by the control group for three sequences starting with a question (Hybrid, Deep reasoning and Reasoning prompt) all in combination with Accept respectively Neutral, and for one sequence of two response categories (Prompt - Neutral). See Table 5.6 for the outcomes of the 10 Ancovas which reached the threshold. Five of these Ancovas revealed a significant difference between the two groups on the post-test, controlled for the pre-test.

Table 5.6

Ancovas with bootstrap of T-S sequences in the dialogues, post-test Mean % and SD, corrected for pre-test differences

Teacher-Student sequence		Training	Control	Bootstrap		
		M% (SD)	M% (SD)	F	p	Part η^2
<i>Teacher question</i>	<i>Student response</i>					
Recall	Accept	8.99 (1.19)	12.48 (1.48)	2.50	.10	.11
Recall	Deny	0.91 (.28)	2.72 (.27)	13.14	.01**	.40
Recall	Neutral	3.51 (.56)	6.90 (1.31)	8.49	.02**	.30
Hybrid	Accept	3.64 (.63)	1.88 (.58)	1.62	.17	.08
Hybrid	Neutral	5.22 (1.00)	1.19 (.71)	6.23	.02**	.24
Surface reasoning	Neutral	7.45 (2.09)	5.30 (2.06)	0.33	.55	.02
<i>Teacher response^a</i>	<i>Student response</i>					
Accept	Accept	5.40 (.84)	7.47 (1.00)	0.54	.54	.03
Accept	Neutral	12.45 (1.80)	15.74 (2.20)	1.21	.20	.06
Neutral	Accept	14.49 (1.84)	22.36 (1.44)	3.21	.07	.14
Neutral	Neutral	6.92 (.93)	13.60 (2.27)	9.83	.02**	.33
Prompt	Accept	5.48 (.89)	.76 (.58)	10.35	.01*	.34

* $p < .01$, ** $p < .05$.

^aTeacher utterance other than a question.

Of the T-S-T sequences the threshold was not reached by both groups for sequences containing a reasoning prompt question and by the control group for sequences containing a hybrid prompt a deep reasoning question or a prompt response. See Table 5.7 for the outcomes of the 22 Ancovas which reached the threshold. Seven of these Ancovas revealed a significant difference between the two groups on the post-test, controlled for the pre-test.

Table 5.7

Ancovas with bootstrap of T-S-T sequences in the dialogues, post-test Mean % and SD, corrected for pre-test differences

Teacher-Student-Teacher sequence			Training M% (SD)	Control M% (SD)	F	p	Bootstrap Part η^2
<i>Teacher question</i>	Student ^a	<i>Teacher question</i>					
Recall	Accept	Recall	1.40 (.42)	3.19 (.94)	3.99	.12	.17
Recall	Deny	Recall	.24 (.11)	1.05 (.14)	15.87	.01*	.44
Recall	Accept	Surface reasoning	1.18 (.33)	1.04 (.47)	0.57	.46	.03
Surface reasoning	Neutral	Recall	1.14 (.29)	1.38 (.68)	0.10	.79	.01
Surface reasoning	Neutral	Surface reasoning	.96 (.74)	1.03 (.68)	0.34	.44	.02
<i>Teacher question</i>		<i>Teacher response^b</i>					
Recall	Accept	Accept	1.48 (.32)	2.76 (.49)	4.07	.08	.17
Recall	Neutral	Accept	.81 (.23)	2.05 (.44)	7.66	.05**	.28
Recall	Accept	Neutral	2.02 (.45)	4.46 (.50)	7.10	.01**	.26
Recall	Neutral	Neutral	.51 (.20)	1.79 (.60)	7.24	.01**	.27
Surface reasoning	Neutral	Accept	1.35 (.33)	1.45 (.70)	0.01	.90	.00
<i>Teacher response^b</i>		<i>Teacher question</i>					
Accept	Neutral	Recall	.86 (.28)	3.00 (.65)	11.81	.02**	.37
Accept	Neutral	Hybrid	1.13 (.27)	.46 (.17)	2.10	.09	.10
Neutral	Accept	Recall	1.48 (.33)	4.20 (.63)	17.08	.00*	.46
Neutral	Neutral	Recall	.70 (.18)	3.29(.81)	20.12	.01*	.50
Neutral	Accept	Hybrid	1.50 (.41)	1.11 (.55)	0.24	.61	.01
<i>Teacher response^b</i>		<i>Teacher response^b</i>					
Accept	Neutral	Accept	4.61 (1.18)	6.48 (1.88)	0.16	.69	.01
Accept	Accept	Neutral	1.82 (.53)	4.24 (.57)	2.34	.25	.11
Accept	Neutral	Neutral	2.28 (.56)	5.18 (.90)	3.81	.10	.16
Neutral	Accept	Accept	1.59 (.54)	3.43 (.40)	0.13	.75	.01
Neutral	Neutral	Accept	1.25 (.35)	4.00 (.62)	15.97	.00*	.44
Neutral	Accept	Neutral	4.86 (1.53)	11.62 (2.04)	1.77	.15	.08
Neutral	Neutral	Neutral	1.95 (.39)	5.21 (1.47)	8.22	.01**	.29

* $p < .01$, ** $p < .05$.

^a Student utterance.

^b Teacher utterance other than a question.

5.5 Conclusions and discussion

Two research questions were central in this study. The first was a descriptive one concerning questions posed and responses given by teachers in feedback dialogues between a teacher and an individual student. The second research question concerned the possible effect of a training programme.

In total 46 feedback dialogues between teachers and individual students in senior secondary nursing education were analysed. The feedback dialogues lasted about 20 minutes in the pre-test and 25 minutes in the post-test. Per minute both participants (teacher and student) had four or five turns. Both teachers and students uttered on average between 14-19 words per turn.

In the training group the average duration of the feedback dialogues and the average number of turns increased, while the average number of words teachers used per turn decreased. These aspects remained about the same for the control group. The results indicate that the teachers in the training group became more efficient and their students used more words in the feedback dialogues.

A detailed coding scheme was developed consisting of seven teacher and four student question categories and four teacher and three student response categories. The decision of coding quite a number of different teacher question categories is in line with Thomas Farrar (1986) who stated the need to match the complexity of questions, as well as their meanings in the social context and their relations with the kind of dialogue.

Per question and response category a frequency of occurrence of on average at least once per dialogue was used as a threshold for including a category in further analysis. Based on this threshold, teachers showed to use six of seven question categories, while students used only Recall questions.

The results show that Recall questions and Accept and Neutral responses were used far the most by the teachers. Hybrid prompt, Deep reasoning, and Reasoning prompt questions and prompt responses were used less often. Of all possible Teacher-Student sequences in both groups the combination Recall question - Accept response occurred far the most, with on the second place Surface reasoning - Neutral response. In the training group teachers and students used few deny responses and less than teachers and students in the control group did.

In the control group the number of different Teacher-Student-Teacher sequences which reached the threshold in the pre-test remained approximately the same in the post-test, while teachers of the training group extended their repertoire from 18 to 38 different T-S-T sequences. The new sequences included 16 times one or two teacher questions and four times a prompt response.

We conclude that the teachers in the training group became accustomed to a broader repertoire of single questions, especially deep reasoning questions and reasoning prompts. Their dialogues contained much more different sequences and these sequences included more questions and prompts which potentially stimulate students to elaborate and reflect (Nückles et al., 2009; Wirth, 2009). Therefore we assume that the training programme did contribute to the teachers' competence development in supporting students' reflection skills.

In this study the teachers in the training group used many questions (about 50% of the utterances), much more in comparison to the control group (about 25-30% of the utterances). An explanation might be that the teachers in the training group already began to show other behaviour during the pre-test due to their involvement by that time (namely in the standard setting study, see chapter 3) (Vollmeyer & Rheinberg, 2005).

The results show that the teacher question categories vary in extensiveness of the student answer which followed. For example a Recall question followed by a single student word compared with a Surface reasoning question followed by 231 student words in a turn in the training group and by 254 student words in a turn in the control group. This is in accordance with the distinction made by Greasser and Person (1994) between short and long answer questions. Questions that aim to elicit inference processes may make students think about their learning process and may ask for reflection (reflection prompts; Gé & Land, 2004; Kintsch, 2005). In the training programme attention was paid by the trainer to the fact that hybrid questions are potentially confusing because student can perceive them as verifications, for example to be answered with 'Yes' or 'No', while teachers want to support students' reflection and thus to stimulate elaborations of thinking activities. Therefore teachers were stimulated to pose open questions in feedback dialogues and listen more and talk less (Roter et al., 2004).

In the feedback dialogues teachers in both groups hardly used continuers as '(Uh)mhm' and 'Okay' (i.e., Chi et al., 2001). Teachers of the training group used '(Uh)mhm' only in the post-test. Possibly they used the continuer as response to have time to think about a next question or response and then the student took the turn. Teachers and students in the training group used 'Yes' much more in the post-test than in the pre-test. This might mean that they confirmed that they had comprehended the other or that it was correct what the other said (Erkens & Janssen, 2008).

In the control group both students and teachers used longer elaborations in the post-test than in the pre-test. Students of the training group used longer elaborations in the post-test than in the pre-test, while for their teachers this was the opposite. We conclude that verbal dominance in the training group shifted

somewhat from teacher to student (Chin, 2006; Blatt et al., 2008). This result might be due to the training programme in which teachers learned that long elaborations are not useful for supporting students' reflection but in stead active listening and caring for wait time are important (Good & Brophy, 2000). Moreover during the time available for the dialogues students in the training group got more possibilities to reflect by elaborating about thoughts, feelings, and behaviour and to formulate intentions.

We assume that the feedback dialogues of both teacher groups were interactive and constructive, because in the dialogues many turns occurred and continuers were hardly used (Chi et al., 2001). However, the dialogues in the training group, especially in the post-test, appeared to be the most challenging ones because in these dialogues teachers dominated the least (Chin, 2006).

A limitation of the effect study was that random selection of schools or teachers and matching of groups for the quasi experimental design were not feasible due to the turbulent circumstances during the study in Dutch senior secondary nursing education. Another limitation concerns the coding of Hybrid and Hybrid prompt questions. Questions that are hybrid can both be coded as a recall question and as one of the reasoning question categories. We only coded the hybrid question itself but not the question category to which the hybrid question referred. Therefore we do not know exactly how many questions were recall questions and how many questions concerning reasoning question categories were posed by the teachers in both groups in the pre-test and the post-test.

In this study reflection was chosen as a dependent variable. We conceptualised reflection as a part of self-regulation but at the same time a certain extent of self-regulation is a condition for being able to comprehend and fruitfully use feedback. Further research might be focused on the conceptualisation of the relation between self-regulation and reflection on the one hand and feedback and reflection on the other.

In future research the effect of hybrid questions could be studied in depth by focusing on teacher-student sequences starting with a hybrid question or hybrid question prompt and computing both their frequencies and the frequencies of the teacher categories to which they refer and determining which student utterances follow. These insights could be used in a revised training programme in which teachers also can be instructed more explicitly about reasoning question and prompt categories and their possible effect on a student's reflection. Further research might focus on how individual students appreciate teachers' questions and how teachers' questions influence students' thinking processes and reflection skills. Moreover research can be done to establish what students learn from teachers' feedback by giving them a reflection task before and after the feedback.

The results of this study might be used to develop guidelines for teaching reflection skills in nursing education and in pre-service nursing teacher education. In the study the effect of teachers asking questions on individual students' reflection was established. The teachers were not used to pose open questions in feedback dialogues. In a revised training programme teachers might be instructed explicitly about reasoning question categories and their possible effects on reflection. Moreover teachers can be trained in teaching students to reflect by giving them explicit instruction about how to pose questions to themselves and how they can deal with questions in written reflection tasks. In a training programme explicit attention is desirable for the introduction of theory about reflecting and the use of reflection tasks for reflection skills development. These instruction and tasks should be consistent across courses in school and internship, which is possible when theory is linked to practice (Asselin, 2011). An important condition for this to occur is having a clear and accepted curriculum for the development of reflection (Aronson, 2011).

Appendix A

Coding scheme question categories

Question category	Short description	Example	Reference
1a. Recall T ^a 1b. Recall S ^b	Question to which a reply is expected that requests few e.g., a verification, property or specification, complementation of a concept, quantitative data	T ^a : Where is your internship? S ^b : Is it correct now?	Greasser & Person (1994); Chin (2006)
2. Hybrid T ^a	Question that the student may perceive as a short answer question, but which is intended to obtain a long answer	T ^a : Do you know what the cause is from that?	Greasser & Person (1994)
3. Hybrid prompt T ^a	Question that the student may encourage productive learning strategies, which the student in principle is capable of, but not spontaneously shows or shows on a (too) low level	T ^a : Has your changed attitude ensured that the contact with your colleagues is better now?	Berthold et al. (2007); Nückles et al. (2009); Ifenthaler (2012)
4a. Request T ^a 4b. Request S ^b	Question to do a particular action.	T ^a : Tell me. S ^b : Explain just once more, please.	Greasser & Person (1994)
5a. Surface reasoning T ^a 5b. Surface reasoning S ^b	Question to which an answer longer than one sentence is expected but hardly an argument, e.g. demand for a definition, example, comparison, interpretation or judgment	T ^a : What else did go well? S ^b : What do you mean by explain?	Greasser & Person (1994)
6a. Deep reasoning T ^a 6b. Deep reasoning S ^b	Question to which an answer longer than one sentence is expected and that exposes a clear way of reasoning e.g., about what led to an event, cause and effect, purposefulness, expectations, a means which enables to do something or an instrumental-procedural question	T ^a : Why do you think that would be of influence? S ^b : How could I do it the next time?	Greasser & Person (1994)
7. Reasoning prompt T ^a	Question that a student may encourage to productive learning strategies, which the student in principle is capable of, but not spontaneously shows or shows on a (too) low level.	T ^a : You mention 'panting in your neck' of people who want to hurry. How does that influence you?	Greasser & Person (1994); Berthold et al. (2007); Nückles et al. (2009); Ifenthaler (2012)

^aTeacher, ^bStudent.

Appendix B

Coding scheme response categories

Response category	Short description	Example	References
1a. Neutral T ^a 1b. Neutral S ^b	A reaction which is no acceptance or negation of what the other said before	T ^a : Yes, but I ask this myself. S ^b : I am doing an internship in a hospital.	Chin (2006); Blatt et al. (2008), Erkens & Janssen (2008)
2a. Deny T ^a 2b. Deny S ^b	A negation of what the other said	T ^a : No, that is not right. S ^b : No, that is not what I mean.	Chin (2006); Blatt et al. (2008); Erkens & Janssen (2008)
3a. Accept T ^a 3b. Accept S ^b	Acceptation and/or confirmation of what the other said; a positive utterance	T ^a : Yes, indeed. S ^b : Yes, I can do it that way.	Chin (2006); Blatt et al. (2008); Erkens & Janssen (2008)
4. Prompt T ^a	A tip or hint which can stimulate the student to use learning strategies which the student is capable to in principal but is not showing spontaneously or on a too low level	T ^a : A next time you could use the alarm button.	Berthold et al. (2007); Nückles et al. (2009); Ifenthaler (2012)

^aTeacher, ^bStudent.

Appendix C

Example of coding and sequences (questions and responses)

The example is a part of a feedback dialogue between a teacher and student of the training group in the post-test. Codes of teacher questions and responses and of student responses are included. Teacher questions concern Recall and Hybrid prompt and teacher responses concern Accept, Neutral and Prompt. Student responses concern Accept and Neutral. The digits show how different sequences were distinguished for analyses; digit 1 shows which utterances form together sequence 1, digit 2 shows which utterances form together sequence 2, et cetera.

T ^a or S ^b	Utterances ^c	Category	T - S seq	T - S - T seq
T	And could you think of something uhm, an example or name, of which you say that you have it sometimes experienced? That you talk about a critical situation?	Hybrid prompt	1	1
S	Now uhm, critical, acute uhm, yes acute a gentleman who has fallen and then everywhere lies blood of which you think yes, you don't see immediately where it does come from. And then it is oohjee, what should I ...	Neutral	1	1
T	Yes.	Neutral	2	1 - 2
S	And then nevertheless behave in a right way by taking the man inside, if he can still walk the distance, cleaning up everything and then quietly looking what it is.	Neutral	2	2
T	Mhmm, yes.	Accept	3	2 - 3
S	And then, based on this, phone a resident or possibly a specialist.	Neutral	3	3
T	That you take action?	Recall	4	3 - 4
S	Yes, then take action yes. Uhm, yes I then always, I always keep quiet in view of such things. I always first look carefully and talk with this man and then uhm, reassure him and that. Yes such situations or a situation of uhm a man who got a bad result that he had a tumour.	Accept	4	4
T	Yes.	Accept	5	4 - 5
S	Yes, in that case it is first something of ohjee, because you are the first to hear it, because I was the first person who came there when they just had received the result. Then it is for a moment sort of yes, how do I react on the spot. So first I will quietly listen for a while to the story. And uhm you know, what the people have to say themselves, and even then I do not react much, I simply let them talk and then I ask if they have something to ask or want something from me or if I can do something.	Accept	5	5
T	Yes, and it is to your satisfaction when you are going home in the evening?	Hybrid prompt	6	5 - 6
S	Yes, I actually feel quite good about this.	Accept	6	6
T	Yes.	Accept	7	6 - 7
S	Because I think I can fairly well imagine myself in other people.	Neutral	7	7
T	You do also write it here.	Neutral	8	7 - 8
S	Yes, I added it there.	Accept	8	8
T	That is one of your qualities.	Prompt	9	8 - 9

^aTeacher, ^bStudent, ^cUtterances, as part of a teacher - student feedback dialogue in the post-test of the training group

6 Teachers' feedback to support students' reflection⁷

Abstract

Feedback dialogues between a teacher and an individual student have the potential to support students' reflection. Little is known about the feedback teachers give in such one-to-one dialogues to support reflection. In this study 46 videotaped feedback dialogues between teachers and individual nursing students were analysed from teachers who participated in a training programme to further their competences in giving feedback (n = 16) and teachers of a control group (n = 7).

The study aimed to get more insight into: (1) What are the relations between teachers' feedback and students' reflection during feedback dialogues? (2) What is the impact of a teacher training programme about giving feedback on teachers' repertoire in giving feedback and students' reflection skills? All teachers' utterances were coded with a coding scheme including nine categories of feedback and two categories of additional information. All students' utterances were coded with a coding scheme consisting of eight reflective thinking activities, four categories concerning content information and appreciation, and a single category for other information. The data were analysed using Multi Episode Protocol Analysis and Ancova, focusing on single teacher (T) and student (S) utterances and sequences of two (T-S) and three (T-S-T) utterances.

Teachers on average did use seven out of nine feedback categories and both additional information categories at least once per feedback dialogue. Students did so for six out of eight thinking activities, three out of four content-related categories and the single other information category. Corrected for pre-test differences the two teacher groups (training and control), on the post-test significantly differed on several of the single teacher and student categories and sequences of two and three teacher and student categories. The results indicate that in the post-test of the training group teacher feedback in the dialogues was primarily Suggestive feedback. Didactic feedback and Corrective feedback were used less. Suggestive and Corrective (positive) feedback were often followed by students showing thinking activities. Our careful conclusion is that the training seemed to contribute to changes in teachers' feedback repertoire during feedback dialogues which have potential to support students' reflection.

⁷ A shorter version of this chapter is submitted as manuscript to an international peer-reviewed journal.

6.1 Introduction

In senior secondary vocational education, such as nursing education, coaching students on their reflection is a new task for teachers. It is expected that feedback by teachers on students' reflections under certain circumstances can further students' reflection skills development (Higgins, Hartley & Skelton, 2002; Kluger & DeNisi, 1996; Quinton & Smallbone, 2010; Sadler, 1998). Feedback can be seen as information communicated to students to change their thinking and behaviour and to improve their learning (Shute, 2008). In order to learn from feedback it is important that both the feedback sender (often a teacher) and the feedback receiver (mostly a student) are active and interact (Nicol & Macfarlane-Dick, 2006; Shuell, 1988). Therefore much is to be expected from verbal teacher-student feedback dialogues (Blatt, Confessore, Kallenberg & Greenberg, 2008; Brinko, 1993). During feedback dialogues teachers can check students' understanding of the feedback given and explain any vagueness while the students can verify their interpretation of the information. Both participants can make connections between the information and relevant actions for further improvement (Hattie & Timperley, 2007).

In feedback dialogues interaction becomes visible in sequences of teacher and student utterances (Ruiz-Primo & Furtak, 2007). Teachers and students exchange information about students' progress in learning and performance towards learning goals and standards (Nicol & Macfarlane-Dick, 2006). Students can use this information to close the gap between current and desired performance (Sadler, 2010). Research about verbal feedback in education mainly concerns the learning content to be discussed in groups (Chin, 2006) and in tutor conversations (Chi, 1996), but does not regard verbal feedback that is useful to support students' reflection. The guidance of students who reflect in feedback dialogues asks for new teacher competences and these might be developed in a teacher training programme.

The aims of this study are twofold. First to investigate the relations between teachers' feedback and students' reflection during feedback dialogues about how to become a nurse. Second to research whether teachers' feedback and students' reflection can be improved by a teacher training programme about giving feedback. The research questions are: What are the relations between teachers' feedback and students' reflection during feedback dialogues? What is the impact of a teacher training programme about giving feedback on teachers' repertoire in giving feedback and students' reflection skills?

6.2 Theoretical framework

Learning functions and autonomy support

In senior secondary vocational education it is common that teachers and students have feedback dialogues in which they exchange information about students' current and desired performance and how to close the gap between these two (Butler & Winne, 1995; Sadler, 2010). In such dialogues feedback from teachers which furthers students' reflection and deep learning is essential (Higgins et al., 2002; Nicol & Macfarlane-Dick, 2006; Quinton & Smallbone, 2010; Sadler 1998). For this to occur it is important that the feedback fits the students' learning process and needs (Black & Wiliam, 1998). Actually this is about finding an appropriate balance between autonomous student learning and teacher directed learning (Shuell, 1988). To learn from feedback students should actively process the information and relate it to other available information in awareness of the desired outcome.

Shuell (1988) distinguished learning functions which can be engaged by teachers or students, for example monitoring learning as checking for understanding (teachers) and monitoring performance as self testing (students). Vermunt and Verloop (1999) distinguished three categories of learning functions: cognitive, affective, and regulative activities teachers can use as a substitute for students' activities leading to a strong teacher regulation or for activating students in a regulation shared with students. Cognitive learning activities, for example structuring and analysing, can lead to learning outcomes in terms of knowledge. Students employ affective learning activities, for example activities concerning attribution, appreciation, and coping with emotions. Regulative learning activities are meta-cognitive activities that students use to decide about the learning content, to control the cognitive and affective learning activities, and to steer the outcomes of their learning. The relative importance of teacher-initiated and student-initiated activities depends on the situation and the task.

Learning from instruction involves interaction and therefore teachers better avoid long explanations because these are not interactive, nor constructive (Chi, Siler, Jeong, Yamauchi & Hausmann, 2001). Ideally teachers do not give more guidance than necessary and they support students' autonomy to enhance their engagement (Reeve, Jang, Carrell, Jeon & Barch, 2004). Comparably, we assume that students' learning from reflection also benefits from teachers' autonomy support which is in balance with students' need to monitor their own reflection process (Shuell, 1988).

Teacher feedback

Chi (1996) and Chi and colleagues (2001) distinguished three groups of feedback types: 'suggestive feedback', 'didactic explanation', and 'corrective feedback'. Suggestive feedback is feedback that gives students a cue to think about or that evokes ideas. Such feedback can be a means to support students' reflection without being very directive (Van den Boom, Paas & Van Merriënboer, 2007). Examples of suggestive feedback are summarizing, checking whether teacher and student understood each other, and diagnosing. Didactic explanation is feedback in which a teacher explains certain content as a reaction to a student's mistakes. This explanation provides knowledge that can be helpful for example to understand what is relevant in reflection but does not provide opportunities for exercising reflection skills. Examples of didactic explanations are information about reflection methods and about the meaning of reflective thinking. Corrective feedback is information which prevents that the student goes into a wrong direction. We suppose that these three groups of feedback types used by teachers in learning situations will also be used by teachers in feedback dialogues to support students' reflection. Besides corrective feedback, which may represent negative information, it is important to give students also positive information. A confirmation, for example: "Well done, that you called the doctor immediately when the patient got unconscious!" is specific and emphasises concrete successful behaviour. Such confirmation can be seen as informative appreciation which might strengthen a student's self-confidence (Seligman & Csikszentmihalyi, 2000). Another categorisation of feedback types comes from Chin (2006). She identified four feedback types in science lessons based on research about feedback and questioning: 'confirmation-direct instruction', 'correction-direct instruction', 'focusing and zooming', and 'constructive challenging'. The first two types have similarities with a combination of didactic explanation and corrective feedback (Chi, 1996; Chi et al., 2001). In confirmation-direct instruction students' action or answer is confirmed, followed by direct instruction while in correction-direct instruction corrective information is followed by explanation of standards. A teacher who uses the third feedback type 'focusing and zooming' accepts an answer and continues asking questions that elicit, probe and elaborate student's thinking. The fourth feedback type 'constructive challenging' implies that a teacher gives evaluative or neutral comments followed by a reformulated question. In both the third and fourth feedback type a teacher continues to ask questions. This is an important activity that has the potential to guide student's thinking for example about similarities and differences in situations, alternative ways of acting, declaring and searching why things worked out or not.

Shute (2008) distinguished two types of feedback, aiming at verification respectively elaboration. Verification feedback consists of information about the correctness of an answer or reaction from a student. Such feedback is typically rather short. Elaboration feedback explains why an answer or reaction is or is not correct and is based on an analysis of mistakes and misunderstandings. This feedback type is a combination of corrective feedback and an explanation based on an analysis of the cause why something went wrong. The student gets additionally the chance to revise the reaction. Shute also distinguished a feedback type in which the teacher gives hints and instructions to guide a student in the right direction. This feedback type resembles the category 'focusing and zooming' of Chin (2006) if the feedback concerns the student's reflective thinking. Another type of elaboration feedback is a teacher's information about how to go further (Shute, 2008). In a dialogue 'looking forward' can be seen as a next step after diagnosing what was essential in the situation or experience about which the student reflected.

Reflection skills and thinking activities

Dewey (1933) described reflection as a specific investigative way of thinking in reaction to doubt in a certain situation. The reflection leads to a conclusion. Since then much has been written about reflection but no unambiguous and agreed-upon description is available (see, e.g., Atkins & Murphy, 1993; Hatton & Smith, 1995; Mansvelder-Longayroux, Beijaard & Verloop, 2007; Procee, 2006). A common core is that reflection is a systematic way of thinking which is linked to the own behaviour. In this study by reflection we mean a conscious and deliberate process of thinking about and interpreting situations, events, experiences, available knowledge, insights and emotions, with the aim to get to know the own behaviour and to learn from it (Ertmer & Newby, 1996; Hatton & Smith, 1995; Kelchtermans, 2007; Korthagen, Kessels, Koster, Lagerwerf & Wubbels, 2001; Lee 2005). Reflection can be seen as a specific set of thinking activities to support learning and therefore in this study from now on the term 'thinking activities' is used.

Reflection is a thinking process that is not visible and can only be studied indirectly. Several studies have shown that it is possible to investigate reflection by operationalising it in certain cognitive, affective and regulative activities in different contexts (Dekker-Groen et al., 2011; Mansvelder-Longayroux et al., 2007; Oosterbaan, Van der Schaaf, Baartman & Stokking, 2010; Shuell, 1988; Van der Schaaf, Baartman, Prins, Oosterbaan & Schaap 2011; Vermunt & Verloop, 1999). In this study reflection is operationalised in terms of eight thinking activities: Describe, Analyse, Structure, Evaluate, Attribute, Explain, Conclude,

and Intend. These thinking activities were developed in an earlier study and were successfully used in nursing education to measure teachers' beliefs about students' reflection (Dekker-Groen et al., 2011). See Appendix B at the end of this chapter for a short description and an example of the thinking activities.

Sequences in feedback dialogues

In a dialogue the exchange between the teacher and the student is a natural unit of analysis. Besides analysis of two successive utterances analysis of longer sequences can give additional insight in the teacher-student dialogue, for example the sequences IRF: teacher Initiation, student Response, teacher Feedback/Follow-up (Mercer, 2004) and IRE: teacher Initiation, student Response, teacher Evaluation (Van Zee & Minstrell, 1997). Van Zee and Minstrell stated that a shift from traditional classroom discourse towards more reflective practices can be made by changing the nature of the third turn in these sequences from an evaluative to a reflective utterance. This happens only when teachers do not pose questions on which they already know the answer but questions instead that invite students to think about issues which could be relevant, such as "Does that make sense to you?" or "Which alternative solutions may be successful in such circumstances?" In line with this proposed shift Greasser, Person and Hu (2002) pleaded for a five-step dialogue frame to get students more actively involved. Their dialogue frame used by tutors builds on an IRE toss and a QAF pattern: teacher Question, student Answer, teacher Feedback. Characteristics of the five-step dialogue frame are tutors' high frequent use of deep-reasoning questions, short feedback and questioning which becomes more specific. In such an approach students get opportunities to practice reflection and to improve their reflection skills. Probably analysing a sequence of five turns will give additional insights in teacher-student interaction during a feedback dialogue to support a student's reflection.

6.3 Methods

Design

We used a quasi experimental pre-post test design including an experimental and control group to explore teachers' and students' utterances during feedback dialogues and to examine the effects of a teacher training in giving feedback (see Figure 6.1). Both the pre-test and the post-test consisted of the registration of a feedback dialogue of a teacher with an individual student. The time lag between the pre-test (t1) and post-test (t3) was between 12 and 15 months.

	t1	t2	t3
Experimental group	O1	X	O2
Control group	O3		O4

Figure 6.1 Quasi experimental pre-post test control group design

The experimental group participated in a training programme (t2). We will call this group the training group.

Intervention

The intervention was a teacher training programme for supporting nursing students' reflection skills development (Dekker-Groen, Van der Schaaf & Stokking, 2013). In three group meetings of 90 minutes information and instruction was given about the function of questioning, general and specific questions, question types and their relation with feedback and reflection. Moreover video fragments of representative feedback dialogues between teachers and students were used as a means for teachers to discuss with colleagues and to exchange experiences (Borko, Jacobs, Eiteljorg & Pittman, 2008; Van Es & Sherin, 2008). Each group meeting was followed by a period of independent practice of new behaviour in interaction with students. In individual meetings of 45 minutes video interaction analysis was used. This is a method to give teachers insight in their actions (Caris-Verhallen, Kerkstra, Bensing & Grypdonck, 2000; Roter et al., 2004). The researcher, who is also video interaction coach, discussed with the teacher video fragments of the teacher's own feedback dialogues with a student and guided the teacher to reflect on experiences and to formulate new goals. After each individual meeting there was a period of independent practice.

Participants

Six schools of senior secondary vocational education for nursing participated in the study. For the training group a purposeful sample was used based on the following selection criteria: being interested in developing the own competences and having enough time to participate in the training. Additional selection criteria for both the training group and the control group were: being responsible for teaching and guiding students for a minimum number of 20 hours per week and not being close to retirement. As a result 19 teachers joined the training programme and eight teachers formed a control group. Three teachers of the

training group and one teacher of the control group could not participate in the post-test due to departure to another job or personal reasons. Therefore two feedback dialogues of 16 teachers of the training group and seven teachers of the control group could be analysed.

The training group consisted of 12 female and four male teachers, on average 51 years old and with 13 years of teaching experience. The teachers in the control group were all female, on average 47 years old and with 16 years of teaching experience. The differences in age and experience between the two teacher groups were not statistically significant.

Teachers of the control group worked on other locations of the same institutes as teachers of the training group. The control group did not receive any information about the research questions nor the training programme and did not have contact with the teachers from the training group. The education programme did not differ between locations and all teachers used the same methods and same materials. All teachers posed a student to participate in a feedback dialogue. Students who participated were in their second year of nursing education, mostly female (per group 79% or more), and on average 20 years old. The students in the post-test were other students than in the pre-test.

Instrumentation

Based on the literature about feedback and reflection referred to before, two coding schemes were developed for respectively coding teachers' and students' verbal behaviour categories in feedback dialogues. Also a coding map was developed in which we included a description and an example for each code and a number of coding rules. For example each utterance gets only one code (teacher category or student category) also when in an utterance more than one code might be given. In that case that code is chosen which is seen as most important feedback category to support the student's reflection (see at the end of the next section) or the most important thinking activity (see at the end of the section after the next section). Also the context of the feedback dialogues was taken into account. Due to this coding map it was clear what was understood by each code and when a certain code had to be used (Chi, 1997).

Coding scheme for teachers' feedback categories

For coding teachers' feedback we developed a scheme based on the three groups of categories used in tutor conversations: suggestive, didactic and corrective feedback (Chi, 1996; Chi et al., 2001). To these groups we added categories from literature concerning feedback (Hattie & Timperley, 2007; Shute, 2008),

educational research (Chin, 2006; Collins, Brown & Newman, 1989), and positive psychology (Seligman & Csikszentmihalyi, 2000).

Suggestive feedback encloses five categories: Summarise, Check, Focus, Diagnose and Forward. At the start of feedback dialogues, to support students' reflection, teachers may give feedback in the form of a summary which is coded as Summarise (category 1). This can help students because it becomes clear what teachers perceive as core of the conversation. Sometimes a summary is a kind of repetition by which teachers emphasise something or pose the same question again, for example because the student avoided the question. Teachers can use Check (category 2) to control whether teacher and student are on the same trail (Chi, 1996). A feedback strategy intended as a check is often ended by an additional question such as "Is it right?" Teacher and student can complement or correct each other if necessary.

In a feedback dialogue meant to support a student's reflection it is very important that teachers Focus (category 3) their feedback on students' explorations of current processes or experiences. By focusing and zooming in at the core a relation can be made with possible or desirable behaviour (Chin, 2006; Shute, 2008). Diagnose (category 4) can be seen as combining the teacher activities evaluating and concluding because a diagnosis is based on an assessment and an (implicit) conclusion. Such feedback might cognitively challenge students (Chin, 2006; Shute, 2008) for example to devise alternatives. Further it is important that students decide what could remain the same in their behaviour and what should be improved. By using Forward (category 5) teachers let students pay attention to future actions (Hattie & Timperley, 2007; Shute, 2008).

The second feedback group Didactic feedback refers to two feedback categories: Expound and Modeling (categories 6 and 7) as ways to support students' reflection. In a dialogue teachers not only give feedback on reflections, but sometimes also explanations about theory, for example the use of thinking categories that may be necessary (i.e., didactic explanation, Chi, 1996; Chin, 2006; Shute, 2008). Moreover it may be helpful when teachers show how to reflect (modeling, Collins et al., 1989).

The third feedback group Corrective feedback (e.g. Chi, 1996, Chin, 2006; Shute, 2008) contains the categories Positive appreciation and Negative appreciation. Positive appreciation (category 8), for example concerns teachers' compliments about students' execution of a task, provides students with information that it went well: (Seligman & Csikszentmihalyi, 2000). Negative appreciation (category 9) is information to prevent misconceptions or to help students understand why things went wrong.

Finally, we distinguished an utterance group ‘Other’ containing two categories. First, teachers can give procedural information in a dialogue: Inform procedural (category 10). Examples are: agree about the goals, monitor progress and make appointments. Second, utterances which can not be coded by the categories mentioned before, are not understandable, or contain very few words because the speaker was interrupted are coded with Inform otherwise (category 11). See Appendix A for the scheme of 11 teacher categories. We assume that Suggestive feedback will be the most important group to evoke students’ reflections, followed by Didactic feedback and then Corrective feedback.

Coding scheme for students’ utterances

The coding scheme of students’ utterances firstly includes the group ‘Thinking activity’ which consists of eight categories of reflective thinking: Describe, Analyse, Structure, Evaluate, Attribute, Explain, Conclude and Intend (Dekker-Groen et al., 2011).

In reflection conversations not every utterance of a student will be reflective. Therefore, we added Content information, a group of two cognitive activities (Vermunt & Verloop, 1999) concerning the categories Summarise and Amplify. Summarise will be used if students repeat what they said earlier or say something comparable. It is also a form of feedback for the teacher about what students mean to say. Amplify will be used when students say something which is not a thinking activity. Amplify and Describe have in common that both categories are about facts. Amplify is coded, on the one hand, if students give information as an answer on a question of the teacher about practical issues, for example the name of the internship institute or how long the internship still lasts. On the other hand we use Amplify to code student’s utterance when the teacher’s utterance is a ‘Yes’ or ‘(Uh)mhm’ between two student utterances which both could be coded with the same thinking activity. The reason for this is that it is very difficult to determine for a coder which feedback category is meant by a teacher’s ‘Yes’ or ‘(Uh)mhm’. It may be a verbal sign for students to go on (acknowledgement response; Chi et al., 2001) or a confirmation or an appreciation. By not choosing a thinking activity we code the student’s utterance conservatively.

In addition a category Content appreciation was included to be able to code content with Appreciate positive or Appreciate negative. Finally the category Inform otherwise was added to code utterances which were not heard well or which concerned only a few words for example because the student was interrupted by the teacher and could not finish her utterance.

See Appendix B for this scheme of 13 student categories. We assume that with regard to reflection the group of categories ‘Thinking activity’ is the most important.

Segmentation and interrater agreement

All videotaped feedback dialogues were transcribed verbatim (Mercer, 2004). Next, transcripts were imported in the program Multi Episode Protocol Analysis (MEPA) (Erkens, 2005). Speech turn-taking was used as the segmentation criterion in each transcript, because it fits both the natural course of conversation and the research questions (Chi, 1997). Moreover turn-taking is a criterion for segmentation that can be objectively applied. All words spoken in one turn (by the teacher or the student) will be called an utterance. Each utterance of a teacher and of a student was coded with a teacher category, respectively a student category.

Two research assistants independently coded three randomly chosen protocol segments of 100-102 utterances (resulting in 302 allotted codes; 150 to teachers’ utterances and 152 to students’ utterances), each from a different feedback dialogue between a teacher and a student who did not participate in the study but were from the same schools as in the sample for the study. Interrater agreement and percentage agreement were computed. Cohen’s Kappa for teachers’ utterances was ‘fair’ (.58) and for students’ utterances ‘good’ (.70) (Banerjee, Capozzoli, McSweeney & Sinha, 1999; De Wever, Schellens, Valcke & Van Keer, 2006; Strijbos & Stahl, 2007). Percentage agreement of the allotted codes for the teachers was 68% and for the students 75%. Next, the two research assistants independently coded all 46 feedback dialogues (both pre-test and post-test of each of 16 (training) and 7 (control) teachers).

Data collection

Data was gathered in the form of videos of feedback dialogues between teachers and individual students to support the student’s reflection skills. Two feedback dialogues per teacher were videotaped, one before and one after the training programme, resulting in 46 video recordings in total. In the training group this took place in the summer of 2009 and the summer and autumn of 2010. In the control group for each teacher the length of the period between the two videotaped dialogues was similar. All teachers got the same written request to videotape a feedback dialogue with a student who participated in their mentor group. They asked each student’s consent about the use of the videotape for research purposes. In both groups one feedback dialogue in the pre-test concerned feedback on how successful a practical task in the curriculum was done, all the other dialogues were

about progress in the study and in the internship. In the post-test the feedback dialogues of all teachers were about the progress in the study and in the internship. All teachers used certain information or material from the student as input for the feedback dialogue, such as a student's portfolio, a written reflective report, or the written answers on a reflection task. The dialogues in the training group lasted on average 18 minutes in the pre-test and 24 minutes in the post-test. For the control group the duration was on average 23 minutes in the pre-test and 27 minutes in the post-test.

Data analysis

Means and standard deviations of the number and duration of utterances of the videotaped feedback dialogues were computed. Differences in number and duration of utterances and words between the training group and the control group were analysed using independent two-tailed *t*-tests, and within each condition the differences between pre-test and post-test were analysed using two-tailed paired *t*-tests. For the significant results the effect sizes were computed.

In MEPA all utterances per feedback dialogue were coded. Utterances could be questions or responses. Next, frequencies of the teacher utterances and the student utterances were computed separately for the training group and the control group. To control for differences in number of utterances in the dialogues within both groups and between pre-test and post-test also percentages were calculated.

All analyses of sequences were done with a teacher utterance as starting point because in this study the teachers' feedback was central. Through analyses of sequences we can determine which feedback categories are used to initiate reflection and which teacher feedback follows on a student utterance.

Sequences of two utterances were analysed to determine how often a certain teacher utterance was followed by a certain student utterance (T-S sequence). In MEPA four different scores of these T-S sequences were computed: Fr-score (frequency), E-score (expected succession frequency), Z-score, and K-score (Kappa score). The Z-score was computed to determine whether the difference between the frequency and the expected frequency of a certain student utterance which followed on a certain teacher utterance was statistically significant (Jeong, 2005). K-scores are corrected Z-scores. In the computation of the K-score a correction is made for differences in the number of segments within and between groups (Wampold & Margolin, 1982). K-scores are between -1 and 1; negative scores indicate that the sequence occurs less often than expected, positive scores indicate that the sequence occurs more often than expected.

Also percentages were computed of the combinations of suggestive feedback categories with teacher question categories respectively response categories in both the pre-test and post-test of the training and the control group. In addition, percentages were computed of the combinations of teacher question categories respectively response categories with student thinking activities in both the pre-test and the post-test of the training and the control group.

Additionally sequences of three utterances were analysed, i.e. Teacher utterance-Student utterance-Teacher utterance (T-S-T sequence). See Appendix C for an example of how 17 successive utterances were divided in sequences and coded. In each feedback dialogue we determined which teacher feedback categories were used and in combination with which student utterance. Moreover we analysed differences between pre-test and post-test in T-S-T sequences.

A large number of sequences occurred only (very) infrequently, so we decided to use a threshold for analysing and reporting about utterances and sequences. We used a minimum total frequency of 16 in the training group and of 7 in the control group to ensure that in both groups on average the utterance or sequence occurred at least one time per dialogue. This threshold should be reached on at least one of the two moments of measurement (pre-test and post-test).

To estimate the effects of the training Ancova was used to control for differences in pretest percentages of teacher and student categories between the training group and the control group (Huck & McLean, 1975). Ancova was used for analysing teacher-student (T-S) sequences and teacher-student-teacher (T-S-T) sequences. These analyses on the differences between the two groups were preceded by a bootstrap procedure because the sample was small and the data was not normally distributed (Field, 2009). Bootstrapping was done by creating 2000 bootstrap samples with replacement. Besides the F-value of Ancova both the p-value and the partial η^2 of the bootstrap were computed.

6.4 Results

6.4.1 Duration and number of segments

The duration of videotaped feedback dialogues was for the training group in the pre-test on average $M = 17.4$ minutes, $SD = 7.5$ and in the post-test $M = 24.1$ minutes, $SD = 8.5$, and for the control group in the pre-test $M = 22.6$, $SD = 10.6$ and in the post-test $M = 26.6$, $SD = 8.9$. Two-tailed paired t -tests on the duration of feedback dialogues revealed a significant difference between pre-test and post-test of the training group ($t(15) = -2.87$, $p = .012$, $\eta^2 = .36$).

The number of segments in the feedback dialogues was for the training group in the pre-test $n = 2487$ ($M = 155.4$, $SD = 72.2$) and in the post-test $n = 3994$ ($M = 249.6$, $SD = 131.1$) and for the control group in the pre-test $n = 1781$ ($M = 254.6$, $SD = 171.9$) and in the post-test $n = 1923$ ($M = 274.7$, $SD = 120.3$). Two-tailed paired t -tests on the number of turns between the pre-test and post-test of the training group revealed a significant difference ($t(15) = -2.77$, $p = .014$, $\eta^2 = .34$).

6.4.2 Teacher and student utterances in feedback dialogues

Teacher utterances

For the coding of teacher utterances 11 categories were distinguished (see Appendix A). In both teacher groups (training and control) two categories did not reach the threshold of occurring on average at least once per dialogue, namely: Modeling (Didactic feedback) and Appreciate negative (Corrective feedback). See Table 6.1 for the frequencies and percentages of the teacher categories which did reach the threshold.

Table 6.1

Frequencies and percentages of single teacher utterances

	Pre-test				Post-test			
	Training		Control		Training		Control	
	Freq	Perc	Freq	Perc	Freq	Perc	Freq	Perc
Feedback								
<i>Suggestive</i>								
Summarise	67	5.4	35	3.9	124	6.2	31	3.2
Check	170	13.6	131	14.7	218	10.9	215	22.3
Focus	413	33.1	103	11.6	813	40.6	105	10.9
Diagnose	180	14.4	72	8.1	167	8.4	78	8.1
Forward	20	1.6	16	1.8	28	1.4	8	0.8
<i>Didactic</i>								
Expound	197	15.8	309	34.7	192	9.6	291	30.2
<i>Corrective</i>								
Appreciate positive	112	9.0	173	19.4	343	17.1	174	18.0
<i>Other</i>								
Inform procedural	30	2.4	29	3.3	30	1.6	31	3.2
Inform otherwise	43	3.4	17	1.9	76	3.8	23	2.4
<i>Total^a</i>	1249	100	891	100	2001	100	964	100

^a Frequencies and percentages in this row are based on all teacher categories, including Modeling and Appreciation negative.

In the feedback dialogues of the pre-test teachers in the training group used Suggestive feedback the most (68.1%), especially Focus (33.1%), for the control group this was 40.1% respectively 11.6%. Teachers in the control group used Didactic feedback as Expound the most (34.7%). For the training group this was 15.8%. Both groups differed in the use of Diagnose (Suggestive feedback) and Appreciate positive (Corrective feedback); training group 14.4% respectively 9% and control group 8.1% respectively 17.1%.

In the post-test the difference between groups in use of Focus increased because the use increased in the training group and did not change in the control group. Frequencies in the use of Expound remained about the same in both groups and percentages decreased almost comparably. Moreover the percentages of the use of Diagnose became almost the same in both groups due to a decrease in the training group. Frequencies in the use of Appreciate positive increased three times for the training group and remained the same for the control group, resulting in almost doubled percentages for the training group and an essentially unaltered percentage for the control group. The category Other occurred less than 6% in both the pre-test and post-test of both teacher groups.

Combinations of a teacher suggestive feedback category and a question or a response category

Teacher utterances which were coded as a feedback category were additionally coded as a question category or a response category. We distinguished seven question categories (Recall, Hybrid, Hybrid prompt, Surface reasoning, Deep reasoning, Reasoning prompt, Request) and four response categories (Accept, Neutral, Deny, Prompt) (see chapter 5 Appendix A). Table 6.2 shows the frequencies and percentages of ten combinations of three suggestive feedback categories with six question categories and nine sequences of five suggestive feedback categories with three response categories which all occurred on average at least one time per dialogue in the training group and the control group in the pre-test or post-test or both.

Table 6.2
Frequencies and percentages of combinations of teacher feedback and question or response categories^a

		Pre-test				Post-test			
		Training		Control		Training		Control	
Two codes per teacher utterance		Freq	Perc	Freq	Perc	Freq	Perc	Freq	Perc
<i>Suggestive feedback and question</i>									
Check	Recall	140	16.47	108	30.25	192	14.22	179	40.96
Check	Hybrid	(12)	(1.41)	7	1.96	(10)	(0.74)	12	2.75
Focus	Recall	42	4.94	10	2.80	41	3.04	16	3.66
Focus	Hybrid	89	10.47	23	6.44	170	12.59	17	3.89
Focus	Hybrid prompt	43	5.06	9	2.52	107	7.93	(6)	(1.37)
Focus	Surface reasoning	106	12.47	30	8.40	127	9.41	43	9.84
Focus	Deep reasoning	60	7.06	8	2.24	160	11.85	(4)	(0.92)
Focus	Reasoning prompt	(6)	(1.29)	(5)	(1.40)	61	4.52	(0)	(0.00)
Diagnose	Recall	20	2.35	(2)	(0.56)	(14)	(1.04)	(1)	(0.23)
Diagnose	Hybrid	18	2.12	(5)	(1.40)	(4)	(0.30)	(2)	(0.46)
<i>Suggestive feedback and response</i>									
Summarise	Accept	13	1.53	9	2.52	18	1.33	(6)	(1.37)
Summarise	Neutral	42	4.94	26	7.28	93	6.89	24	5.49
Check	Neutral	(11)	(1.29)	12	3.36	(10)	(0.74)	19	4.35
Focus	Neutral	22	2.59	11	3.08	59	4.37	14	3.20
Focus	Prompt	30	3.53	(5)	(1.40)	71	5.26	(5)	(1.14)
Diagnose	Accept	(14)	(1.65)	18	5.04	24	1.78	19	4.35
Diagnose	Neutral	69	8.12	27	7.56	73	5.41	51	11.76
Diagnose	Prompt	40	4.71	17	4.76	41	3.04	4	0.92
Forward	Neutral	(7)	(0.82)	7	1.96	(7)	(0.52)	(5)	(1.14)
<i>Total^b</i>		734	86.36	322	90.17	1237	91.64	398	91.17

^a Frequencies which did not reach the threshold (for training group at least 16 and for the control group at least 7) and their percentages are between parentheses.

^b These totals are without the numbers given above between parentheses.

Table 6.2 shows that in both the training group and the control group the combination of the teacher categories Check and Recall occurred far the most. Focus was used in combination with five different question categories, more in the training group than in the control group. All five Suggestive feedback categories were used in combination with Neutral response, of which Check only in the control group. Summarise and Diagnose were used in combination with Accept response. Focus and Diagnose were used in combination with Prompt response, of which Focus only in the training group.

Student utterances

For the coding of student utterances 13 categories were distinguished (see Appendix B). In both groups the students used two thinking activities respectively one Content information category on average less than one time per dialogue, namely Structure and Attribute respectively Summarise, so these categories did not reach the threshold. See Table 6.3 for the frequencies and percentages of the remaining nine student categories which reached the threshold.

Table 6.3
Frequencies and percentages of student utterances^a

	Pre-test				Post-test			
	Training		Control		Training		Control	
	Freq	Perc	Freq	Perc	Freq	Perc	Freq	Perc
<i>Thinking activity</i>								
Describe	93	7.5	40	4.5	205	10.3	70	7.3
Analyse	43	3.5	8	0.9	100	5.0	7	0.7
Evaluate	154	12.4	71	8.0	205	10.3	73	7.6
Explain	148	12.0	97	10.9	239	12.0	111	11.6
Conclude	70	5.7	45	5.1	195	9.8	46	4.8
Intend	(8)	(0.7)	(5)	(0.6)	26	1.3	(1)	(0.10)
<i>Content information</i>								
Amplify	314	25.4	243	27.3	224	11.2	293	30.6
<i>Content appreciation</i>								
Appreciate positive	293	23.7	313	35.2	636	31.9	290	30.2
Appreciate negative	50	4.0	14	1.6	35	1.8	22	2.3
<i>Other</i>								
Inform otherwise	52	4.2	48	5.4	104	5.2	41	4.3
<i>Total^b</i>	1238	100	890	100	1993	100	959	100

^a Frequencies which did not reach the threshold (for training group at least 16 and for the control group at least 7) and their percentages are between parentheses.

^b Frequencies and percentages in this row are based on all student categories, including Structure, Attribute, and Summarise.

In the feedback dialogues of the pre-test students in the training group used utterances concerning Thinking activities the most (42.2%), especially Evaluate and Explain (12.4% respectively 12%). In the post-test again Thinking activity utterances were used the most (50.8%) concerning an about equally division between four Thinking activities namely Describe and Evaluate (both 10.3%), Explain (12.0%) and Conclude (9.8%). Students of the control group used in the pre-test Content appreciation the most (36.8%), followed by Thinking Activities (31.5%) and as third Content information (28.0%). In the post-test the three

percentages were about equal. Other activities occurred 4%-5% in both the pre-test and post-test of both student groups.

6.4.3 Sequences of teachers' feedback categories and students' thinking activities

Teacher-Student Sequences (T-S)

Each sequence consists of a pair of two utterances starting with a teacher utterance followed by a student utterance. Because eleven different teacher categories and thirteen different student categories were distinguished, in total 143 different sequences were possible. In the training group and in the control group together in total 24 different T-S sequences occurred on average at least one time per dialogue in the pre-test or the post-test (the threshold for analysing and reporting). See Table 6.4 for the results of 20 of the sequences, the two teacher groups taken together. Besides frequencies (see Fr) also expected values (see E) are included per column per moment of measurement (pre-test and post-test). (The expected values per cell were calculated as the total frequency of the concerning row times the total frequency of the concerning column divided by the overall total frequency of all rows and columns.) Moreover significant Z-scores (see Z) and their adjusted scores (see K: Kappa values) are given.

Four of the 24 frequencies which did occur least frequently are not shown in the table to avoid making the table overcrowded. Summarise-S Appreciate positive occurred in the training group 22 times (pre-test) and 80 times (post-test) and in the control group 16 and 12 times. Expound-Amplify occurred in the training group 60 and 22 times, in the control group 66 and 80 times. Expound-S Appreciate positive occurred in the training group 61 and 111 times, in the control group 162 and 131 times. Inform procedural-S Appreciate positive occurred in the training group 17 and 14 times, in the control group 16 and 20 times.

Table 6.4 shows that in the training group T-S sequences starting with Check respectively Diagnose occurred fewer and those starting with Focus respectively Appreciate positive occurred much more in the post-test in comparison to the pre-test. In the control group T-S sequences starting with Check occurred more and those starting with Appreciate positive occurred fewer in the post-test in comparison to the pre-test, while the occurrence of T-S sequences starting with Focus or Diagnose was comparable in the pre-test and the post-test.

The Z-scores and K-scores in Table 6.4 show that 60 T-S sequences occurred significantly more and two T-S sequences significantly less than could be expected. The occurrence of eleven T-S sequences was as expected.

Table 6.4 *Frequencies, Expected Frequencies, Kappa- and Z- scores of T-S sequences^b*

Teacher utterance	Student utterance															
	Thinking activity								Content information		Content appreciation					
	Describe		Evaluate		Explain		Conclude		Amplify		Appreciate positive					
	^a	pre	post	pre	post	pre	post	pre	post	pre	post	pre	post			
<i>Suggestive feedback</i>																
Check	T	(Fr=10)	(Fr=12)	(Fr=10)	Fr= 23	(Fr=8)	Fr= 23	(Fr=0)	(Fr=0)	Fr= 52	Fr= 25	Fr= 69	Fr= 105			
					E=11.15	E=13	E=21.54	E=12.18	E=19.85	E=34.59						
					Z=3.75	Z=2.94	Z=7.35	Z=3.89	Z=12.16	Z=13.42						
					K=.06	K=.05	K=.21	K=.06	K=.33	K=.38						
Focus	C	(Fr=6)	Fr= 22	Fr= 12	Fr= 8	Fr= 7	Fr= 25	(Fr=6)	(Fr=0)	Fr= 44	Fr= 74	Fr= 44	Fr= 63			
					E=7.81	E=5.21	E=8.14	E=7.11	E=12.38	E=17.82	E=32.67	E=22.96	E=32.34			
					Z=5.49	Z=3.16	Z=3.92	Z=3.92	Z=6.93	Z=8.33	Z=5.03	Z=6.21				
					K=.23	K=.10	K=.13	K=.23	K=.23	K=.19	K=.17					
Focus	T	Fr= 45	Fr= 108	Fr= 70	Fr= 81	Fr 73	Fr= 134	Fr= 34	Fr= 107	Fr= 84	Fr= 77	Fr= 49	Fr= 151			
					E=15.32	E=41.58	E=25.37	E=41.58	E=24.38	E=48.48	E=11.53	E=39.55	E=51.73	E=45.44	E=48.1	E=129.01
					Z=8.46	Z=11.84	Z=10.01	Z=7.03	Z=11.11	Z=14.19	Z=7.34	Z=12.32	Z=5.25	Z=5.4	Z=2.37	
					K=.38	K=.41	K=.35	K=.24	K=.39	K=.45	K=.38	K=.43	K=.12	K=.18	K=.04	
	C	Fr= 10	Fr= 17	Fr= 14	Fr= 15	Fr= 33	Fr= 34	Fr= 13	Fr= 11	Fr= 10		Fr= 19	Fr= 12			
					E=2.31	E=3.81	E=4.09	E=3.98	E=5.59	E=6.05	E=2.6	E=2.51	E=14.01	E=18.05	E=15.79	
					Z=5.28	Z=7.08	Z=5.15	Z=5.8	Z=12.28	Z=12.05	Z=6.74	Z=5.59				
					K=.20	K=.20	K=.15	K=.16	K=.30	K=.28	K=.25	K=.20				

^a T = Training group (for each T in the table), C = Control group (for each C in the table).

^b Frequencies which did not reach the threshold are noted between parentheses.

(Table 6.4 continues)

Table 6.4 (Continued) *Frequencies, Expected Frequencies, Kappa- and Z- scores of T-S sequences*

Teacher utterance	Student utterance												
	Thinking activity						Content information		Content appreciation				
	Describe		Evaluate		Explain		Conclude		Amplify		Appreciate positive		
	^a	pre	post	pre	post	pre	post	pre	post	pre	post	pre	post
<i>Suggestive feedback</i>													
Diagnose	T	(Fr=8)	(Fr=0)	Fr= 31	Fr= 13	Fr= 24	Fr= 7	(Fr=0)	(Fr=10)	Fr= 36	(Fr=9)	Fr= 51	Fr= 113
				E=11.08	E=8.54	E=10.65	E=9.96			E=22.6		E=21.02	E=26.5
				Z=6.41		Z=4.38				Z=3.13		Z=7.23	Z=18.71
				K=.14		K=.10				K=.09		K=.19	K=.62
C	(Fr=5)	(Fr=0)	Fr= 7	Fr= 15	Fr= 13	Fr= 9	Fr=8	(Fr=0)	(Fr=6)	Fr= 14	Fr= 28	Fr= 28	
			E=2.86	E=2.95	E=3.91	E=4.49			E=11.85	E=12.62	E=11.73		
			Z=2.55	Z=7.3	Z=4.83	Z=2.34				Z=4.87	Z=5.26		
			K=.06	K=.17	K=.13	K=.06				K=.26	K=.25		
<i>Corrective feedback</i>													
Appreciate positive	T	(Fr=9)	Fr= 53	Fr= 16	Fr= 55	Fr= 16	Fr= 50	(Fr=5)	Fr= 50	Fr= 42	Fr= 65	(Fr=8)	Fr= 31
			E=17.54	E=6.9	E=17.54	E=6.33	E=20.45		E=16.69	E=14.06	E=19.17		E=54.43
			Z=9.09	Z=3.66	Z=9.6	Z=3.84	Z=7.05		Z=8.74	Z=8.15	Z=11.27		Z=-3.62
			K=.19	K=.09	K=.20	K=.09	K=.14		K=.19	K=.29	K=.22		K=-.43
C	Fr= 10	Fr= 14	Fr= 16	Fr= 17	Fr= 23	Fr= 13	Fr= 7	Fr= 10	Fr= 91	Fr= 91	Fr= 14	Fr= 11	
	E=3.87	E=6.28	E=6.88	E=6.55	E=9.4	E=9.96	E=4.36	E=4.13	E=23.54	E=26.29	E=30.32	E=26.02	
	Z=3.31	Z=3.29	Z=3.74	Z=4.36	Z=4.8			Z=3.07	Z=15.74	Z=14.37	Z=-3.43	Z=-3.35	
	K=.17	K=.12	K=.14	K=.16	K=.16			K=.14	K=.45	K=.44	K=-.54	K=-.58	

^aT = Training group (for each T in the table), C = Control group (for each C in the table).

Teacher question categories followed by a student category

Students can respond in different ways on teacher questions. More insight into this interaction can be reached by determining which teacher utterances coded as one of the question categories (Recall, Hybrid, Hybrid prompt, Surface reasoning, Deep reasoning, Reasoning prompt, Request, see Appendix A chapter 5) were followed by which of the possible student categories. Table 6.5 shows the frequencies and percentages of the combinations of teacher question categories followed by a student utterance which occurred on average at least once per dialogue in both groups (training and control) in the pre-test or post-test or both.

Table 6.5

Frequencies and percentages of teacher questions followed by student utterances

Teacher utterance	Student utterance	Pre-test				Post-test			
		Training		Control		Training		Control	
		Freq	Perc	Freq	Perc	Freq	Perc	Freq	Perc
<i>Teacher question</i>	<i>Thinking activity</i>								
Recall	Describe	14	2.23	8	3.56	16	1.66	18	5.96
Recall	Evaluate	14	2.23	10	4.44	24	2.49	8	2.65
Recall	Explain					22	2.28	25	8.28
Recall	Conclude							7	2.32
Hybrid	Describe	14	2.23			35	3.63		
Hybrid	Evaluate	30	4.78			19	1.97	7	2.32
Hybrid	Explain	16	2.55	15	6.67	31	3.22	8	2.65
Hybrid	Conclude					27	2.80		
Hybrid prompt	Explain					24	2.49		
Surface reasoning	Describe					19	1.97	8	2.65
Surface reasoning	Analyse					18	1.87		
Surface reasoning	Evaluate	24	3.82			19	1.97	7	2.32
Surface reasoning	Explain	27	4.30	13	5.78	26	2.70	20	6.62
Deep reasoning	Describe					22	2.28		
Deep reasoning	Analyse					22	2.28		
Deep reasoning	Evaluate					17	1.77		
Deep reasoning	Explain	17	2.71			36	3.74		
Deep reasoning	Conclude					32	3.32		
<i>Teacher question</i>	<i>Content information</i>								
Recall	Amplify	70	11.15	43	19.11	26	2.70	70	23.18
Hybrid	Amplify	19	3.03			18	1.87		
Surface reasoning	Amplify	37	5.89			25	2.60		
<i>Teacher question</i>	<i>Content appreciation</i>								
Recall	Appreciate positive	82	13.06	46	20.44	144	14.95	64	21.19
Recall	Appreciate negative	25	3.98	9	4.00	15	1.56	14	4.64
Hybrid	Appreciate positive	21	3.34			37	3.84		
Hybrid prompt	Appreciate positive					25	2.60		
<i>Teacher question</i>	<i>Other</i>								
Deep reasoning	Inform otherwise					16	1.66		
<i>Total</i>		410	65.30	144	64.00	715	74.22	256	84.78

Table 6.5 shows that in both groups (training and control) Recall-S Appreciate positive occurred often in the pre-test and the post-test. In the training group Recall-Amplify occurred often in the pre-test, while in the control group this combination occurred the most in the post-test. In the training group in the post-test eleven new combinations occurred, with the teacher question categories Deep reasoning (5 times), Hybrid prompt and Surface reasoning (both 2 times), Recall and Hybrid (both 1 time). In the control group new combinations occurred with the teacher question categories Recall and Surface reasoning (both 2 times) and Hybrid (1 time). Four teacher question categories were each followed by at least four different thinking activities. Two of these teacher question categories were also followed by Amplify respectively Appreciate positive.

Teacher-Student-Teacher Sequences (T-S-T)

T-S-T sequences which reached the threshold for analysing and reporting in this study consisted of 53 different sequences of two teacher feedback categories with a student utterance in between. In the pre-test and post-test of the training group this concerned 13 respectively 31 different T-S-T sequences, in the pre-test and post-test of the control group 25 respectively 27 different T-S-T sequences. See Table 6.6 for these T-S-T sequences. Two T-S-T sequences (not included in table 6.6) contained the student category Other and were left out.

The sequences in Table 6.6 show that in the post-test compared to the pre-test the training group reached the threshold for 20 new T-S-T sequences. In these sequences teacher utterances were mostly combinations of Focus and Appreciate positive (nine times). Four other sequences contained Appreciate positive and three other sequences contained Focus. The student utterances accompanying these teacher category sequences were five times Appreciate positive, three times Describe, Evaluate, Explain and Amplify, two times Explain, and one time Analyse. In the training group three sequences did not reach the threshold in the post-test anymore: Focus-Explain-Diagnose, Diagnose-Evaluate-Focus, and Expound-Amplify-Expound.

In the control group in the post-test compared to the pre-test nine new T-S-T sequences reached the threshold including four times the teacher categories Check, Diagnose, Expound, and Appreciate positive and three times the category Focus. The accompanying student utterances were four times Appreciate positive, two times Explain, and one time Describe, Amplify, and Appreciate positive. Seven sequences in the control group did not reach the threshold in the post-test anymore: four times Focus and Diagnose, three times Expound and Diagnose, and one time Forward and Appreciate positive.

Table 6.6

Frequencies of Teacher-Student-Teacher Sequences (T-S-T) occurring on average at least once per dialogue in training group (T) and control group (C)

Teacher Feedback category			Student Utterance category															
			Thinking activity						Content information				Content appreciation					
			Describe		Analyse		Evaluate		Explain		Conclude		Amplify		Appr positive		Appr negative	
T	S	T	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post		
<i>Two times Suggestive feedback</i>																		
Summarise	S ^a	Summarise															T17	
Summarise	S ^a	Focus															T32	
Check	S ^a	Check										C11	C28	C11			T16 C16	C8
Check	S ^a	Focus												T22		T41 C11		
Focus	S ^a	Check		C8					T17 C7									
Focus	S ^a	Focus	T17	T48		T35	T30	T38	T26	T49	T16	T47	T30	T23	T17		T94	
Focus	S ^a	Diagnose							T18 C10									
Diagnose	S ^a	Focus					T16								T23 C7		T51	
Diagnose	S ^a	Diagnose													C8		T21	
<i>Suggestive and Didactic feedback</i>																		
Check	S ^a	Expound											C8	C16	C14		C19	
Focus	S ^a	Expound													C9		T16	
Diagnose	S ^a	Expound															C13	
Forward	S ^a	Expound															C7	
Expound	S ^a	Check											C10	C14		C16	C20	
Expound	S ^a	Focus													T20 C10		T38 C11	
Expound	S ^a	Diagnose															C9	
<i>Suggestive and Corrective feedback</i>																		
Check	S ^a	Appr pos								C10				C13	C17			
Focus	S ^a	Appr pos		T25				T20	C11	T41		T17		C17	C21			
Appr pos	S ^a	Check												C17	C21			
Appr pos	S ^a	Focus		T18				T17				T18	C7	T29 C8				
Appr pos	S ^a	Diagnose												C8				
<i>Two times Didactic feedback</i>																		
Expound	S ^a	Expound								C7				T24 C23	C34	T16 C114	T39 C83	
<i>Didactic and Corrective feedback</i>																		
Expound	S ^a	Appr pos												C23	C18			
Appr pos	S ^a	Expound												C21	C22	C9		
<i>Two times Corrective feedback</i>																		
Appr pos	S ^a	Appr pos		T22				T16	C9	T20 C11				C43	T16 C23			
Total			T17	T113 C8		T35	T46	T91	T44 C30	T127 C35	T16	T82	T54 C176	T91 C211	T98 C205	T365 C182	C8	

^a See for the student utterance categorie(s) the column heading(s).

6.4.4 Effects of the training on the occurrence of utterances and sequences

Ancova tests were done on teacher and student categories which were used at least on average once in each feedback dialogue of both the training and control group in the pre-test or the post-test or both. Concerning the frequencies of the teacher categories this threshold was reached for nine of 11 categories. See Table 6.7 for the Mean %, SD, and Ancova results. Three Ancovas revealed a significant difference between the two groups (training and control): Check, Focus, and Expound. Concerning the frequencies of the student categories the threshold was reached for 10 of 13 categories, See Table 6.7 for the Mean %, SD, and Ancova results. Three Ancovas revealed a significant difference between the students of the two groups: Analyse, Conclude, and Amplify.

Table 6.7 *Ancovas with bootstrap of single teacher and student utterances*

	Training	Control	Ancova	Bootstrap	
	M% (SD)	M% (SD)	F	p	Part η^2
Teacher utterance					
<i>Suggestive feedback</i>					
Summarise	6.26 (1.13)	3.24 (1.77)	1.84	.20	.08
Check	11.14 (1.26)	24.03 (2.05)	25.05	.00*	.56
Focus	43.07 (3.45)	11.73 (3.00)	9.32	.01*	.32
Diagnose	8.95 (1.40)	9.10 (2.51)	.00	.96	.00
Forward	1.88 (.52)	1.05 (.32)	1.38	.20	.09
<i>Didactic feedback</i>					
Expound	8.53 (1.79)	27.31 (4.01)	11.74	.02**	.37
<i>Corrective feedback</i>					
Appreciate positive	14.94 (2.05)	17.43 (2.61)	.59	.40	.03
<i>Other</i>					
Inform procedural	2.31 (.62)	3.10 (.74)	.40	.62	.02
Inform otherwise	3.37 (.64)	2.29 (.35)	2.09	.12	.10
Student utterance					
<i>Thinking activity</i>					
Describe	9.07 (1.35)	7.88 (2.80)	.00	.96	.00
Analyse	5.37 (1.04)	1.07 (.48)	3.33	.03**	.16
Evaluate	10.70 (1.20)	8.36 (1.34)	.66	.34	.03
Explain	13.50 (1.62)	12.54 (1.91)	.40	.40	.02
Conclude	11.87 (1.55)	5.25 (.83)	6.15	.01**	.24
<i>Content information</i>					
Amplify	10.53 (1.39)	29.80 (3.91)	47.19	.00*	.70
<i>Content appreciation</i>					
Appreciate positive	29.92 (2.42)	28.66 (3.86)	.31	.58	.02
Appreciate negative	1.89 (.38)	1.6 (.63)	.03	.84	.00
<i>Other</i>					
Inform otherwise	5.12 (.58)	3.88 (1.07)	1.07	.40	.05

* $p < .01$, ** $p < .05$.

Concerning the specific combinations of a teacher suggestive feedback and a question or a response category (see Table 6.2 for the frequencies and percentages), the Mean %, SD, and Ancova results are shown in Table 6.8, pertaining to twelve teacher category combinations which reached the threshold. Five Ancovas revealed a significant difference between the two groups (training and control): Check and Recall, Focus and Hybrid, Focus and Hybrid prompt, Focus and Deep reasoning, and Diagnose and Prompt.

Table 6.8

Ancovas with bootstrap of combinations of a teacher feedback and a question or response category

Two codes per teacher utterance		Training	Control	Bootstrap		
		M% (SD)	M% (SD)	F	p	Part η^2
<i>Suggestive feedback and question</i>						
Check	Recall	8.55 (1.11)	18.26 (1.52)	18.88	.00*	.49
Focus	Recall	1.75 (.30)	1.48 (.56)	.08	.83	.00
Focus	Hybrid	7.49 (1.28)	1.60 (.56)	4.94	.03**	.20
Focus	Hybrid prompt	6.24 (.75)	1.28 (.22)	15.66	.00*	.44
Focus	Surface reasoning	6.59 (1.60)	4.94 (1.93)	.55	.55	.03
Focus	Deep reasoning	7.71 (.94)	.46 (.48)	13.30	.00*	.40
<i>Suggestive feedback and response</i>						
Summarise	Accept	.89 (.32)	.59 (.18)	.66	.34	.03
Summarise	Neutral	4.16 (.68)	2.29 (1.51)	1.68	.03**	.08
Focus	Neutral	2.54 (.67)	1.30 (.53)	1.51	.17	.07
Diagnose	Accept	1.04 (.26)	2.04 (.60)	2.91	.20	.13
Diagnose	Neutral	3.34 (.44)	5.49 (1.37)	4.41	.13	.18
Diagnose	Prompt	2.09(.45)	.39 (.28)	4.67	.03**	.19

* $p < .01$, ** $p < .05$.

The T-S sequences which reached the threshold consisted of 24 different sequences. See Table 6.9 for the Mean %, SD, and Ancova results concerning these sequences. The percentages in this Table have been calculated taking *per* teacher as 100% the total of *all* sequences (including those not reaching the threshold). Six Ancovas revealed a significant difference between the two teacher groups (training and control): Focus-Explain, Focus-Conclude, Check-Amplify, Appreciate positive-Amplify, Expound-Amplify, and Inform otherwise-Appreciate positive.

Table 6.9

Ancovas with bootstrap of Teacher feedback - Student category sequences (T-S)

Teacher-Student sequence		Training	Control	Bootstrap		
		M% (SD)	M% (SD)	F	p	Part η^2
<i>Suggestive feedback</i>	<i>Thinking activity</i>					
Check	Evaluate	1.29 (.43)	.99 (.29)	.02	.85	.00
Check	Explain	1.22 (.39)	2.89 (.67)	5.09	.07	.20
Focus	Describe	5.46 (.93)	1.85 (.85)	2.16	.09	.10
Focus	Evaluate	4.32 (.70)	1.67 (.70)	3.06	.05	.13
Focus	Explain	8.17 (1.27)	3.95 (1.16)	3.25	.04**	.14
Focus	Conclude	7.23 (1.35)	1.25 (.35)	8.00	.01*	.29
Diagnose	Evaluate	.98 (.36)	1.73 (.59)	2.91	.33	.13
Diagnose	Explain	.38 (.14)	1.14 (3.38)	3.19	.14	.14
<i>Corrective feedback</i>	<i>Thinking activity</i>					
Appreciate positive	Describe	1.89 (.57)	1.65 (.52)	.06	.78	.00
Appreciate positive	Evaluate	2.49 (.40)	1.91 (.57)	.42	.48	.02
Appreciate positive	Explain	2.31 (.55)	1.31 (.44)	1.25	.18	.06
Appreciate positive	Conclude	2.32(.49)	1.10 (.40)	2.32	.12	.10
<i>Suggestive feedback</i>	<i>Content information</i>					
Check	Amplify	1.38 (.36)	8.30 (1.29)	46.43	.00*	.70
Focus	Amplify	3.82 (.85)	.70 (.36)	2.75	.14	.12
Diagnose	Amplify	.58 (.21)	1.64 (.75)	2.51	.26	.11
<i>Suggestive feedback</i>	<i>Content appreciation</i>					
Summarise	Appreciate positive	4.02 (.91)	1.25 (4.64)	3.61	.06	.15
Check	Appreciate positive	5.27 (.95)	7.04 (1.13)	.87	.29	.04
Focus	Appreciate positive	6.71 (1.32)	1.20 (.43)	4.13	.06	.17
Diagnose	Appreciate positive	5.76(.96)	3.27 (.80)	3.12	.08	.14
<i>Corrective feedback</i>	<i>Content information</i>					
Appreciate positive	Amplify	2.48 (.66)	8.72 (1.95)	9.91	.03**	.62
<i>Corrective feedback</i>	<i>Content appreciation</i>					
Appreciate positive	Appreciate positive	1.57 (.37)	1.20 (.29)	.21	.59	.01
<i>Content information</i>	<i>Content information</i>					
Expound	Amplify	1.13 (.31)	7.92 (1.60)	43.80	.00*	.69
<i>Content information</i>	<i>Content appreciation</i>					
Expound	Appreciate positive	4.68 (1.19)	11.28 (3.36)	1.38	.21	.07
<i>Other</i>	<i>Content appreciation</i>					
Inform procedural	Appreciate positive	.73 (.21)	1.93 (.55)	6.11	.04**	.23
Inform otherwise	Appreciate positive	.49 (.15)	.18(.13)	1.40	.17	.07

* $p < .01$, ** $p < .05$.

Concerning the specific sequences of teacher question categories followed by a student category (see Table 6.5 for the frequencies and percentages), the Mean %, SD, and Ancova results are shown in Table 6.10, pertaining to 11 different sequences of a teacher question category and a student category which reached the threshold. Two Ancovas revealed a significant difference between the two groups (training and control): Recall-Explain and Recall-Amplify.

Table 6.10
Ancovas with bootstrap of sequences of a Teacher question and a Student category

Teacher utterance	Student utterance	Training	Control	F	Bootstrap	
		M% (SD)	M% (SD)		p	Part η^2
<i>Teacher question</i>	<i>Thinking activity</i>					
Recall	Describe	.67 (.23)	1.90 (.76)	4.24	.15	.18
Recall	Evaluate	1.31 (.40)	.99 (.21)	.18	.53	.01
Recall	Explain	1.07 (.41)	2.81 (.48)	6.40	.04**	.24
Hybrid	Evaluate	1.25 (.39)	.78 (.43)	.28	.57	.01
Hybrid	Explain	1.70 (.39)	.94 (.41)	.80	.32	.04
Surface reasoning	Describe	1.27 (.41)	1.00 (.64)	.10	.75	.01
Surface reasoning	Evaluate	.95 (.28)	.83 (.48)	.00	.98	.00
Surface reasoning	Explain	1.49 (.76)	2.60 (.97)	.78	.38	.04
<i>Teacher question</i>	<i>Content information</i>					
Recall	Amplify	1.40 (.36)	7.98 (1.44)	40.30	.00*	.67
<i>Teacher question</i>	<i>Content appreciation</i>					
Recall	Appreciate positive	7.34 (1.10)	7.21 (.99)	.00	.93	.00
Recall	Appreciate negative	.64 (.19)	1.35 (.31)	4.14	.05	.17

* $p < .01$, ** $p < .05$.

The T-S-T sequences which reached the threshold consisted of 14 sequences including five teacher categories (Check, Focus, Diagnose, Expound, and T Appreciate positive) combined with three student categories (Explain, Amplify, and S Appreciate positive). See table 6.11 for the Mean %, SD and Ancova results concerning these sequences. Two Ancovas revealed a significant difference between the two teacher groups: Check-S Appreciate positive-Check, and Focus-S Appreciate positive-Explain.

Table 6.11
Ancovas with bootstrap of Teacher-Student-Teacher (T-S-T) sequences

Teacher-Student-Teacher sequence			Training	Control	Bootstrap			
			M% (SD)	M% (SD)	F	p	Part η^2	
<i>Suggestive feedback</i>	- <i>Student utterance</i>	- <i>Suggestive feedback</i>						
Check	Appreciate positive	Check	.64 (.27)	1.91 (.49)	5.23	.05*	.21	
Check	Appreciate positive	Focus	2.32 (.57)	1.40 (.64)	.94	.34	.05	
Focus	Explain	Check	1.18 (.28)	.85 (.45)	1.01	.29	.05	
Focus	Explain	Diagnose	.49 (.21)	.69 (.35)	.27	.62	.01	
Diagnose	Appreciate positive	Focus	2.65 (.70)	.59 (.22)	3.54	.11	.15	
Diagnose	Appreciate positive	Diagnose	1.21 (.33)	.40 (.19)	2.83	.05	.12	
<i>Suggestive feedback</i>	- <i>Student utterance</i>	- <i>Didactic feedback</i>						
Focus	Appreciate positive	Expound	.76 (.34)	.46 (.17)	.25	.49	.01	
Expound	Amplify	Expound	.35 (.22)	3.42 (1.19)	15.88	.05*	.43	
Expound	Appreciate positive	Focus	1.97 (.24)	1.08 (.37)	4.13	.08	.17	
Expound	Appreciate positive	Expound	1.36 (.61)	6.68 (2.37)	1.87	.21	.09	
<i>Suggestive feedback</i>	- <i>Student utterance</i>	- <i>Corrective feedback</i>						
Focus	Explain	Appreciate positive	2.20 (.55)	.37 (.18)	3.41	.04*	.15	
Appreciate positive	Amplify	Focus	1.10 (.45)	.78 (.26)	.20	.62	.01	
<i>Corrective feedback</i>	- <i>Student utterance</i>	- <i>Corrective feedback</i>						
Appreciate positive	Explain	Appreciate positive	1.03 (.28)	1.13 (.48)	.08	.81	.00	
Appreciate positive	Amplify	Appreciate positive	.57 (.23)	1.84 (.72)	.61	.46	.03	

*p < .05.

6.5 Conclusions and discussion

In this study feedback dialogues between teachers and individual students in nursing education were explored and effects of a teacher training programme were analysed. The first research question concerned the relations between teachers' feedback and students' reflection. The second research question was about the impact of a teacher training programme on teachers' repertoire about giving feedback to support students' reflection skills.

To answer the first research question analyses were done concerning frequencies of single teacher and student utterances and sequences of utterances. Results showed that teachers in the control group used especially Check, mostly followed by students' Appreciate positive. Teachers in the training group used especially Focus. In T-S-T sequences Focus was often used twice with in between a Student's Appreciate positive or a Thinking activity. Also combinations of Suggestive and Corrective feedback and of twice Corrective feedback included a

Thinking activity. The results indicate that in the post-test of the training group teacher feedback in the dialogues was primarily Suggestive feedback. Didactic feedback and Corrective feedback were used less. Suggestive and Corrective (positive) feedback were often followed by students showing thinking activities. This feedback possibly supported students' reflection (Van den Boom et al., 2007). Students' Appreciate positive response might indicate their positive engagement (Reeve et al, 2004).

Results about effects of the training programme indicate that the participating teachers learned to focus on students' thinking activities. Utterances in the post-test which were coded with Focus contained more often a Hybrid, Hybrid prompt or a Deep reasoning question. These question types potentially stimulate elaborations (Greasser & Person, 1994). Students more often showed Thinking activities, especially Analyse, Explain, and Conclude, which also was found by Oosterbaan and colleagues (2010). Teachers used more Appreciate positive and less Check en Expound and their students used less Amplify. These results might indicate that students felt more self-confident and teachers and students comprehended each other better (Seligman & Csikszentmihalyi, 2000). These factors might also indicate a positive learning environment and more autonomy for students (Reeve et al., 2004; Shuell, 1988). In the post-test on average each student in the training group formulated intentions (Shute, 2008).

Although the differences in percentages of the use of utterance categories concerning both groups of teachers on the two moments were small, the findings seem consistent and might indicate a shift in the feedback dialogues from dialogues in which information is being exchanged toward dialogues in which students more often reflect on their experiences. Our careful conclusion is that the training seemed to contribute to changes in teachers' feedback repertoire during feedback dialogues which have potential to support students' reflection.

In the coding scheme concerning teacher feedback we included categories which were supposed to be important in feedback dialogues in education (e.g. Chi, 1996; Chin 2006). Suggestive feedback was the first category. Check was used very often, mostly in combination with a Recall question. It seems that this feedback category was often used in dialogues in which especially information was exchanged and the teacher searched for confirmation or correction of information (Chin, 2006). In our study Check does not seem to have the potential to further reflection, neither does Diagnose. Focus however seemed to support reflection because in the post-test Focus was used significantly more by the training group followed by students' thinking activities. This is in line with Chin (2006) who assumes that this feedback category initiates longer answers. Also Greasser and Person (1994) describe question categories which might initiate longer answers, such as the

reasoning questions in this study. According to them, Hybrid questions might be asked by the student by a verification or an elaboration. Also prompts might be followed by elaborations (e.g. Wirth, 2009), as we saw in this study.

Forward was used on average one or two times per dialogue. That teachers use this category might show that they see looking forward by formulating concrete intentions as part of a feedback dialogue. This seems to fit with the importance attached to making plans concrete (Hattie & Timperley, 2007; Shute, 2008). Expound was hardly used by teachers in the training group and modeling was even used less. Expound can be seen as instruction or information by the teacher which is needed for the student to proceed (Chi, 1996; Chi et al., 2001; Chin, 2006) and seems to be more efficient in group lessons than in dialogues with an individual student. Procedural information and information not concerning reflection were used in less than 5% of the utterances. That more than 95% of the utterances concerned content relevant in relation to teacher feedback and students' reflection skills development fits the interpretation that the feedback dialogues were purposeful (Sinatra & Pintrich, 2010).

Concerning the students' utterances in the dialogues, we used Thinking activities as operationalisations of reflection comparable with operationalisations in other studies (Mansvelder-Longayroux et al., 2007; Oosterbaan et al., 2010; Van der Schaaf et al., 2011). In the post-test of the training group all of the categories Describe, Evaluate, Explain, and Conclude were used in about 10% of students' utterances. For Analyse this was 5% and for Intend 1%. Structure and Attribute were used on average less than one time by per dialogue. In the post-test Analyse and Conclude were used significantly more by students in the training group than by students in the control group. The same pertains to the T-S sequences including Explain and Conclude and the T-S-T sequences including Explain. The frequencies of the Thinking activities varied. Possibly part of these activities are used by students more implicitly. Nevertheless it seems fruitful to operationalise reflection in a series of possible thinking activities (cf. Van der Schaaf et al., 2011; Quinton & Smallbone, 2010).

The second category Content information was operationalised as Summarise and Amplify. Students hardly used Summarise but they used Amplify rather often. This especially pertains to students in the control group. Amplify was coded when students gave information about practical issues and when between two Thinking activities the teacher used an acknowledgement response (Chi et al., 2001).

The third category Content Appreciation was mainly used in a positive way and hardly in a negative way. In the training programme we used video interaction guidance in which the core principle was attuned interaction (Kennedy, 2011). Teachers learnt about the importance of receiving students' utterances and about

how to let students receive teachers' utterances. This may be an explanation of the increase in the use of Appreciate positive by both teachers and students.

A limitation of the study is that the results concern those Feedback categories and Thinking activities of which the frequency exceeded a certain threshold. Moreover in this study per utterance only one code could be allocated, while for some utterances several codes might be applicable. Consequently the number of coded teacher feedback utterances and student thinking activities was lower than their actual use. A second limitation is that the study was based on a restricted, non-probabilistic sample of teachers and students in a specific period, which restricts the possibilities to generalise the findings and conclusions.

This study has practical relevance as it concerned the effects of bringing in expertise about giving feedback on reflection in a training programme for teachers. Students can be taught to self-reflect by giving them explicit instruction how to reflect. This instruction should be a part of a consistent reflective structure across courses in which theory that is taught is linked to practice (cf. Asselin, 2011). In the same vein in a teacher training programme explicit attention is desirable for knowledge concerning reflection skills development (theory and reflection tasks) and curriculum development (cf. Aronson, 2011). The training content developed (about feedback, reflection, and coaching using video interaction) and the insights into the effects of the programme on students' reflection in nursing education might also be useful in other contexts and for other target groups (for example in pre-service teacher education or education for social work professions).

A suggestion for future research can be to compare different ways of coding: coding only one feedback category per utterance (this study) and coding all possible feedback categories per teacher utterance respectively to make a comparison between coding only one thinking activity per utterance and coding all possible thinking activities per student utterance. Such comparisons can give insight in the most useful way of coding. Also different approaches of coding can be compared: an analytical approach and more holistic approach for example by using levels of reflection (cf. Mezirow, 1991; Wallman, Kettis Lindblad, Hall, Lundmark & Ring, 2008). Levels and content of reflection were not investigated. Future research can be done into the relation between thinking activities, levels of reflection and content of reflection (cf. Oosterbaan et al., 2010).

In the study T-S and T-S-T sequences were analysed to get more insight in the interaction of teachers and students in feedback dialogues to support students' reflection. Further research is important because concerning these sequences still not much is known. Further study might also give more insight into how to guide feedback dialogues intended to support students' reflection.

Appendix A

Coding scheme teacher utterances

Utterance category	Reference	Short description	Example
Feedback			
<i>Suggestive</i>			
1. Summarise	Chi (1996)	Naming the core, also repeating or saying something in other words	So, to summarise, in fact the teachers have confirmed what you are doing well
2. Check	Chi (1996)	Checking whether teacher and student comprehend each other well	Do you mean ...? So actually a fairly good internship?
3. Focus	Chin (2006)	Asking more questions to stimulate student's thinking	What is a good guidance for you?
4. Diagnose	Chi (1996); Chin (2006)	Concluding implicitly or explicitly or assessing	You know, that is very important what you say
5. Forward	Hattie & Timperley (2007); Shute (2008)	Giving information or posing a question, let the student review the feedback and/or reflect aiming at the future	How would you, looking back to that specific situation, handle it possibly the next time in another way?
<i>Didactic</i>			
6. Expound	Chi (1996); Chi et al. (2001); Chin (2006)	Explaining, amplifying or clarifying	In fact it consists of the whole list what is in your workbook
7. Modeling	Collins et al. (1989)	Thinking aloud, modeling	I considered then which role I played myself and what I could attribute to colleagues
<i>Corrective</i>			
8. Appreciate positive	Seligman & Csikszent- mihalyi, (2000)	Making a positive remark about the content or the learning process of the student, or complimenting	Okay. Now a fine intention.
9 Appreciate negative	Corrective feedback: Chi (1996); Chi et al. (2001)	Making a negative remark about the content, confronting the student or giving corrective feedback	No, I still need it. Lets look. Shall I just get a piece of paper for you?
<i>Other</i>			
10. Inform procedural		Making a procedural remark, e.g. about the continuation of the feedback dialogue	Just a moment, I'll switch off the camera
11. Inform otherwise		An utterance not being feedback on the content or the learning process of the student.	Okay, thank you

Appendix B

Coding scheme student utterances

Utterance category	Short description	Example
<i>Thinking activity</i> ^a		
1. Describe	Telling precisely and systematically the main points about a happening, situation, process or experience as observed, which is occasion for reflection	Yes acute a gentleman who has fallen and then everywhere lies blood of which you think yes, you don't see immediately where it does come from. And then it is oohjee, what should I do
2. Analyse	Work out which aspects play a role in the person, practice and/or theory by which one becomes aware	And that I did notice yes, it is not only about the care and little things and so. These people are sitting there and they only are very alone there.
3. Structure	Taking together aspects, placing them in chronological sequence to recognize patterns and to search for connections	And then there became a line visible for me of hee, that actually does come back
4. Evaluate	Appraising or assessing a happening, situation, process or experience in relation to a goal or norm	It's going better, easier, more pleasant now because I know what I have to do and what I can expect
5. Attribute	Attributing a role or cause to oneself, an other person, or the situation	It did happen because I didn't pay attention well
6. Explain	Naming reasons or causes of behaviour or feelings	...because then I become more calm... I do it this way because
7. Conclude	Drawing conclusions about a happening, situation, process or experience in relation to thinking and behaving	So that doesn't word. Actually I have to do it otherwise
8. Intend	Deciding, having an intention and saying what you are going to do	I want to try telling to ten first
<i>Content information</i> ^b		
9. Summarise	Saying the same or in somewhat different words; feeding back	Yes how do I approach it
10. Amplify	Giving information concerning the content, giving explanation	Pfff, then I have to take the booklet before me
<i>Content appreciation</i>		
11. Appreciate positive ^c	Making a positive remark about what the teacher said, confirming or complimenting	Yes. When you are simply very busy, right
12. Appreciate negative ^d	Making a negative remark about what the teacher said or giving corrective information	No, that's all in
<i>Other</i>		
13. Inform otherwise	Making an utterance which is not feedback on the content or learning process	My pleasure

^aBased on Dekker-Groen et al. (2011), ^bVermunt and Verloop (1999), ^cSeligman and Csikszentmihalyi (2000), ^dChi (1996) and Chi et al. (2001).

Appendix C

Example of coding and analyses of sequences (teacher's feedback and student's reactions)

The example is a part of a feedback dialogue between a teacher and student of the training group in the post-test. Codes of teacher and student categories are included. Teacher categories concern Suggestive feedback (Check, Focus), Didactic feedback (Expound) and Corrective feedback (Appreciate positive). Student categories concern Thinking activity (Describe, Evaluate, Explain), Content information (Amplify) and Content appreciation (Appreciate positive). The digits show how different sequences were distinguished for analyses; digit 1 shows which utterances form together sequence 1, digit 2 shows which utterances form together sequence 2, et cetera.

T ^a or S ^b	Utterances ^c	Category	T - S seq	T - S - T seq
T	And could you think of something uhm, an example or name, of which you say that you have it sometimes experienced? That you talk about a critical situation?	Focus	1	1
S	Now uhm, critical, acute uhm, yes acute a gentleman who has fallen and then everywhere lies blood of which you think yes, you don't see immediately where it does come from. And then it is oohjee, what should I	Describe	1	1
T	Yes.	Appreciate positive	2	1 - 2
S	And then nevertheless behave in a right way by taking the man inside, if he can still walk the distance, cleaning up everything and then quietly looking what it is.	Describe	2	2
T	Mhmm, yes.	Appreciate positive	3	2 - 3
S	And then, based on this, phone a resident or possibly a specialist.	Amplify	3	3
T	That you take action?	Check	4	3 - 4
S	Yes, then take action yes. Uhm, yes I then always, I always keep quiet in view of such things. I always first look carefully and talk with this man and then uhm, reassure him and that. Yes such situations or a situation of uhm a man who got a bad result that he had a tumour.	Describe	4	4
T	Yes.	Appreciate positive	5	4 - 5
S	Yes, in that case it is first something of ohjee, cause you are the first to hear it, because I was the first person who came there when they just had received the result. Then it is for a moment sort of yes, how do I react on the spot. So first I quietly listen for a while to the story. And uhm you know, what the people have to say themselves, and even then I do not react much, I simply let them talk and then I ask if they have something to ask or want something from me or if I can do something.	Describe	5	5
T	Yes, and it is to your satisfaction when you are going home in the evening?	Focus	6	5 - 6
S	Yes, I actually feel quite good about this.	Evaluate	6	6
T	Yes.	Appreciate positive	7	6 - 7
S	Because I think I can fairly well imagine myself in other people.	Explain	7	7
T	You do also write it here,.	Expound	8	7 - 8
S	Yes, I added it there.	Appreciate positive	8	8
T	That is one of your qualities.	Focus	9	8 - 9

^a Teacher, ^b Student, ^c 17 successive utterances within a continuous teacher - student feedback dialogue.

7 Conclusions and discussion

Introduction

The research reported in this thesis started with the premise that students in nursing education have to be prepared for continuous learning during the career. The development of reflection skills, as part of self-regulated learning, stimulates lifelong learning. Being able to support the development of these skills of their students requires that teachers expand their pedagogical repertoire with interventions and forms of interaction that stimulate reflection.

The research question of this thesis was: How can teachers' competences in supporting nursing students' reflection skills be developed by a teacher training programme?

The sub-questions were:

(a) What are the characteristics the programme should have in view of current insights in the literature on teacher professional development?, (b) How can the programme be designed in view of the knowledge about teachers' interventions and interactions which might contribute to students' reflection skills development?, (c) What are the feasibility and the appreciation of the programme?, (d) What are the effects of the programme?, and (e) Which conclusions can be drawn about the value of a knowledge-based approach for extending teachers' pedagogical repertoire?

This thesis includes five empirical studies which build upon each other. The studies have been conducted in senior secondary nurse education level 4, because in the nursing profession reflection skills are key, as nurses have to take well-considered decisions and actions in every-day practice.

In the first study content standards were developed concerning what teachers should know and be able to do to support students' reflection skills development. In the second study for eight teacher competences performance standards were developed, distinguishing four levels of performance. The third study concerned the development, conducting and evaluation of a programme in which teachers of five schools in secondary nursing education were trained on four of these competences. Videotapes of feedback dialogues of teachers with individual students were analysed in the studies four and five.

Conclusions per sub-question

(a) What are the characteristics a programme should have in view of current insights in the literature on teacher professional development?

In the literature on teacher training and professional development several approaches can be distinguished which differ in the activities they emphasise: 1) instructing teachers, letting them exercise skills and giving them feedback (training-application approach; Joyce & Showers, 2002; Tillema & Veenman, 1987); 2) letting teachers think about their current behaviour and alternatives (developmental-reflective approach; Kuipers et al., 2010); 3) letting teachers discuss and exchange experiences with colleagues (social learning approach; Bandura, 1977); 4) informing teachers about available knowledge and best practices (expertise-based approach; Timperley et al., 2007).

We tried to combine these approaches, of which the first, third and fourth were used in group meetings and periods of individual practice and the second and fourth were pertinent in meetings between the trainer and the participating teachers individually.

On a more concrete level in the literature consensus has been growing about the following core features of effective teacher professional development: a clear focus on relevant content, ample possibilities for active learning, coherence between all parts of the programme, sufficient duration and intensity, and collective participation (Desimone, 2009; Garet et al., 2001; Penuel et al., 2007).

We succeeded in implementing most of these characteristics into the training programme. In all meetings with the participating teachers relevant knowledge was brought in about the new curriculum goals and materials and relevant competences such as posing questions, giving feedback and stimulating reflection. Teachers worked focused and in an active way to prepare, conduct and evaluate feedback dialogues they had with an individual student (Sinitra & Pintrich, 2010). Active learning was also stimulated in the group meetings by urging teachers to exchange experiences and to discuss approaches. Coherence was realised by making connections between content and activities of the programme and by linking these with the teachers' knowledge, beliefs, skills, and needs (Tillema & Imants, 1995). The programme was spread in time over 12-15 months to ensure the teachers would have enough time for independent practice. Only collective participation was restricted because per school not the whole team but only a rather small number of teachers could participate.

(b) How can a programme be designed in view of the knowledge about teachers' interventions and interactions which might contribute to students' reflection skills development?

The answer to this sub-question has to cover three aspects. First, the content of the programme in view of the available knowledge about the competences the teachers needed to develop, including the pedagogical content knowledge concerning the reflection skills the students should develop and the interventions and interaction forms that can be expected to support the students' development of these skills. Second, the educational design of and the activities in the programme that would stimulate and support the participating teachers' competence development. Third, the method to be used in developing the programme.

The first study resulted in a framework of content standards including relevant teacher knowledge about reflection as well as about teacher behaviour expected to be effective for supporting students' reflection. The content standards were formulated as tasks, because tasks are concrete and can be judged on criteria such as relevance and feasibility. Using a Delphi method in two rounds in a panel of experts, including teachers, consensus was reached about the framework. In this framework reflection was specified in terms of reflection objects (what to reflect upon) and thinking activities (how to reflect). Reflection was seen as a conscious and well-considered process of thinking about and making interpretations of events and situations aiming at improved comprehension and learning in order to behave more effectively and consciously about the impact on others (e.g. Hatton & Smith, 1995; Kelchtermans, 2007, Lee, 2005).

We considered this framework as too extensive for use in a teacher training programme. Moreover standards had to be developed. Therefore in the second study the framework was integrated with the general teacher competence descriptions of the Ministry of Education, Culture and Science (2007) used for human resource development purposes in senior secondary vocational education. The result was a new framework of eight teacher competences, for each of which a rubric was developed with attributes and levels of performance. Using a comprehensive standard setting procedure in six teams consensus was reached about the formulation of the competences, the minimum level teachers should reach (the performance standards), and the judgemental model to be used.

In deciding about the content of the programme it was considered that to be optimally effective the content should not only be relevant but also feasible and fitting the teachers' needs (Garet et al., 2001; Tillema & Imants, 1995). The eight teacher competences resulting from the first two studies can be divided in four competences concerning conditions such as interpersonal and organisational skills

and four competences concerning subject content and pedagogical aspects. The latter group was the most specific and urgent. Consequently these competences were chosen as the content of the programme: giving students instruction and assignments in reflection; coaching students in developing reflection skills; supporting development through feedback; holding reflection conversations and giving assessments.

The educational design and activities that the programme preferably should characterise were already described by the answer on sub-question (a).

The method used to develop the programme was a step by step one, synchronous to the three successive cycles of the training programme. Each cycle included a group meeting, a first period of independent exercise, individual meetings, and a second period of independent exercise, and was adapted to the teachers' remarks and needs. In this way experiences and evaluations concerning each cycle could be used in developing the next cycle. An example of this is that we learned during the first cycle how much importance the teachers attached to the setting of goals for themselves and for their students to be reached in feedback dialogues and how difficult this was for them. Based on this knowledge we could pay more attention to this in the second cycle of the programme.

(c) What are the feasibility and the appreciation of the programme?

The training programme was conducted at five schools. Due to turbulent circumstances in senior secondary vocational education it was not easy to plan meetings. Nevertheless at each school three group meetings of 90 minutes per meeting took place. In addition, each teacher had three individual meetings (45-minutes per meeting). Teachers on average estimated that they spent a total of 26 hours on the programme (including contact time, time for working on tasks, and independent practice), during one school year. In the literature no exact indication is given about the intensity and duration needed for change in teaching behaviour to occur. There is some support for activities spread over a semester that include 20 hours or more of contact time (Desimone, 2009). The span of time in our programme was longer and the contact time was less. The teachers practiced their skills in regular feedback dialogues and in working on training tasks including the preparation of these dialogues and learning by reflection afterwards.

The available time was intensively used. In the group meetings the teachers actively participated, observed video fragments, discussed interpretations, got instruction, exchanged experiences, discussed subject matter as well as teacher behaviours and ways to improve them (Borko et al., 2008; Van Es & Sherin, 2008). Video clips of specific aspects of teacher competences were highlighted and

discussed to help teachers to improve their behaviour (Good & Brophy, 2000). In the individual meetings the trainer showed video clips of a teacher's feedback dialogue. Next, the teacher observed the clips, the trainer gave feedback and both discussed the teacher's behaviour (Caris-Verhallen et al., 2000; Kennedy, 2011). Each video clip lasted about a minute and zoomed in on a specific behaviour, such as posing specific questions. The teacher reflected on her behaviour and its effect on the student (Good & Brophy, 2000; Kennedy, 2011).

Apart from the meetings the teachers hardly consulted each other. This might be attributed to lack of time but also to the fact that every teacher was working on own learning goals concerning the competences central in the training programme. Actually it would not have been very useful to pay attention to common goals such as developing a joint view on the place of reflection in the nursing curriculum, because only three to five teachers per team participated in the training programme. So, the core feature 'collective participation' was hardly realised. Exchange between the participants during group meetings was restricted, as only limited time for these meetings was available. Consulting a colleague has been the least done learning activity, so it is not obvious that this activity contributed to the results of the training.

According to the participants' evaluations of each meeting and their appraisals of the overall programme, it can be concluded that the participating teachers on average were fairly satisfied with the programme. All teachers were particularly enthusiast about assignments that appeal on skills they could use immediately in their own feedback dialogues with students.

(d) What are the effects of the programme?

The effects of the programme were investigated in several ways.

In study three (chapter 4) on several moments and with different instruments the teachers of the training group reported about the possible effects of the training. According to the participating teachers they on average increased in: focus on the content of students' reflections and the results of their guidance; trying things out and looking for more information; self confidence; developing another view on guiding students' reflection. After the programme the teachers indicated they had developed their knowledge, skills and attitudes during the programme and also developed in conducting a number of specific activities. Additionally the teachers were asked about five activities: attending to relations between theory, practice, actions and experiences; giving feedback on possibilities for improvement; giving feedback that stimulates students' thinking in general; supporting students' thinking activities that are specific for reflection; and guiding students' personal

and professional development. Teachers in the training group scored higher on these activities than the control group. According to the teachers in the training group their students increased in the extent to which they reflected on different objects and this increase was statistically significant, also compared with the control group. These indications of effects obviously do not represent overwhelming evidence, but all effects found were in the intended direction.

The results fit with the conclusion of Desimone and colleagues (2002) that competence development focussed on specific instructional practices increases teachers' use of intended practices and that active learning adds to this. The overall conclusion from study three is that when teachers want to actively develop their competences in supporting students' reflection skills are offered a coherent, purposeful programme including instruction, exercise, and feedback, can improve their relevant practice with a fairly modest investment of their time.

In studies four and five (chapters 5 and 6) per teacher two videotaped feedback dialogues with an individual student were analysed. In study four the focus was on the teachers' questions and responses. The results showed that in the post-test the teachers in the training group used more questions in more different categories and more prompt responses which potentially stimulated students to elaborate and reflect (Nückels et al., 2009; Wirth, 2009). Their students made longer elaborations in the post-test than in the pre-test, while for the teachers this was opposite and even significantly so. In the control group both students and teachers made longer elaborations in the post-test than in the pre-test. We conclude that in the training group the verbal dominance shifted somewhat from teacher to student, indicating that students' reflection might have been stimulated (Chin, 2006; Blatt et al., 2008). This result might be due to the training programme in which teachers learned that long elaborations are not useful for supporting students' reflection but in stead active listening and caring for wait time are important (Good & Brophy, 2000, Kennedy, 2011).

In study five the relations between teachers' feedback and students' reactions were analysed. Teachers' feedback in the control group was especially Expound, while their students used mainly Amplify and Positive appreciation. Teachers' feedback in the training group was primarily of suggestive nature, and much less didactic and corrective feedback. The students' thinking activities mostly enclosed: describe, analyse, evaluate, explain, conclude, and intend. The suggestive feedback category Focus was the most used teacher feedback category in the training group, still more in the post-test than in the pre-test. It was mostly followed by a student thinking activity, while the teacher category Expound seemed to be counter-productive in interrupting students' reflection process. The results indicate that the teachers participating in the training learned to focus on students' thinking

processes. Further, according to the post-test, they mostly used specific question categories such as Hybrid, Hybrid prompt and Deep reasoning questions, which potentially stimulate students' elaborations (Greasser & Person, 1994). Their students in the post-test more often showed the thinking activities Analyse, Explain and Conclude. These thinking activities are shown to be prerequisite for reflection by Oosterbaan and colleagues (2010). The teachers of the training group in the post-test used more Appreciate positive and less Check en Expound and their students used less Amplify. These results might indicate that teachers and students did comprehend each other better and the learning environment was positive with more autonomy for students (Reeve et al., 2004; Seligman & Csikszentmihalyi, 2000; Shuell, 1988).

(e) Which conclusions can be drawn about the value of a knowledge-based approach for extending teachers' pedagogical repertoire?

In the first two studies, in which criteria and standards were developed, we explicitly used available knowledge from literature and experts about both the learning goals the teachers had to reach with their students (concerning their reflection skill development) and the teachers' interventions and interactions with students which can be expected to further these goals (such as giving certain feedback, posing certain questions and ways of furthering the dialogue). These criteria, standards and relevant knowledge from an additional literature study were explicitly used in developing the content of the training programme and in conducting this programme. To study the effects of this programme measures were developed based on the knowledge from literature and experts. Given the framework for this thesis and the conditions available this was the maximum extent to which we could make it plausible that the effects found could be attributed ultimately at least partially to the knowledge-base used. A more decisive answer on this sub-question would need a more ambitious experimental design comparing several different approaches of supporting teachers' development of the same competences under the same conditions.

General conclusion

The results fit the conclusion that teacher competence development focussed on specific instructional activities and active learning can be expected to be effective (Desimone et al., 2002). In addition also the involvement of the teachers and other stakeholders in the development of the standards and the training programme will have contributed to the content and activities being perceived as relevant and

feasible and to the effects found (McKenney, Nieveen & Van den Akker, 2006). The programme encountered support and the teachers took ownership of the process (Smith, 2011). The use of video clips in the group meetings had the benefit that the teachers watched examples of colleagues' feedback dialogues (Van Es & Sherin, 2010).

The training programme was based on a combination of three of the four approaches discussed in the chapters one and four of this thesis. First, components of the programme such as instruction, exercise and feedback were typical for a training-application approach (Joyce & Showers, 2002; Tillema & Veenman, 1987). Second, discussion about current practice and reflection on alternatives fitted into a developmental-reflective approach (Kuijpers et al., 2010). Third, giving information about and using relevant knowledge indicated an expertise-based knowledge-development approach (Putman & Borko, 1997; Tillema & Imants, 1995; Timperley et al., 2007). The fourth approach mentioned before, that of social learning (Bandura, 1977), was less central, as the time available for discussion and exchange between the teachers was restricted.

Two additional specific features are: 1) The meetings with the teachers during the programme and also the feedback dialogues between the teachers and their students were authentic situations to them and this contributes to the ecological validity (De Bruijn & Leeman, 2011). 2) The coaching of the teachers by the trainer, using video interaction guidance, might have stimulated their reflection and empowerment (Garmston, Linder & Whithaker, 1993; Kennedy, 2011; Tillema & Smith, 2009; Veenman & Denessen, 2001).

In view of the available evidence we are inclined to conclude that the training programme might well have stimulated the teacher learning processes we modeled in chapter 1 (see Figure 1.1). The teachers were given information, instruction and tasks from which they might learn and they were stimulated to develop and use their knowledge base, to set goals and decide upon strategies for the feedback dialogues with their students, to try new behaviour, to have dialogues with the trainer, their students and their colleagues and receive their feedback, and to reflect upon this feedback. On the level of teachers and students no barriers were encountered. Students seemed to be sensitive for teachers' prompts. Such cues are facilitators for the development of self-regulation (Nicol & Macfarlane-Dick, 2006). The single barrier that will have restricted their learning and development most was probably institutional: the permanent turbulence in their work situation (Yang & Carless, 2012).

Limitations

The studies had to contend with some limitations. This may have influenced or restricted the results. The first pertains to the number of participants. It was planned to involve 28 teachers working in four schools in the training programme: four teams of seven teachers each. Due to the turbulent circumstances in senior secondary vocational education fewer teachers were involved and no complete team was able to participate. The same circumstances restricted the time available for the meetings and for independent exercises. Unfortunately not all 19 teachers who participated in the programme could deliver complete data and in the observational studies only data from 16 participating teachers were available. This result is possibly also due to the length of the period that part of the teachers already participated in our studies, not only in study three (chapter 4) but also in the second or even the first study (chapters 3 and 2).

Other limitations concerned the videotaped teacher-student dialogues and their coding. The students participating in the post-test were mostly other students than in the pre-test, because in the meantime students had been assigned to other teachers as their mentor. Although the teachers invited students for the post-test matching those in the pre-test, it can not be excluded that this has influenced the results. The interrater reliability of the coding of the protocols of these dialogues was mediocre. This was possibly due to the complexity of the coding scheme (Hawkins & Power, 1999). Moreover, we chose to code only one teacher category and one student category per turn, also when more categories were applicable. The reported frequencies of the categories will therefore represent an underestimation of their use.

Finally, the researcher was also the developer of the training programme based on information and feedback of teachers and managers. Moreover the researcher was the trainer in the group meetings and the coach in the individual meetings with the teachers. Although she has tried to remain as objectively as possible and to keep all decisions and activities transparent and standardised, such a combination of roles is intrinsically vulnerable.

Implications for educational practice

Several results of the studies might be relevant for educational practice in senior secondary vocational education for health professions. First, content standards and performance standards were developed and validated which can be used for purposes of in-service teacher training, teacher assessment and quality control policies. Second, the training programme developed can be conducted with other

teachers and in other schools, and also parts of it can be used, for example the exercises or the coaching of individual teachers based on videotaped feedback dialogues. Third, the results of the studies three, four and five (chapters 4, 5, 6) into the effects of the training give insights in aspects of teachers' behaviour that can obviously be changed and might be effective in instructing, interacting with students, posing questions and giving feedback. Finally, the elaborated view on reflection including reflection objects and thinking activities and the developed training programme might also be relevant in other contexts, for example teacher education and social work education.

Suggestions for further research

It would be interesting and might be relevant to constructively replicate the studies three, four and five. Suggestions are to vary the time available for the meetings and the number of meetings; to include students with teachers who are their mentors for a longer period of time; to let them have more than one dialogue; not only focus on the student's internship but on her personal and professional development in a longer period of time; to coach teachers in action by means of ear bud headphones.

In future research also different ways of coding can be compared, for example coding only one feedback category per teacher utterance and coding all possible feedback categories per utterance, respectively coding only one student thinking activity per utterance and coding all possible thinking activities per utterance.

The feedback and reflection utterances might also be studied on different aspects of content, modality and quality, such as their relevance for the specific learning goal, their cognitive complexity, or their affective load, including the affective effects of students' reflection utterances on the teacher.

Also different teacher competence development activities can be studied, for example by having two teachers on the same competence level discuss their feedback dialogues (same-ability reciprocal coaching) or a junior teacher and a senior teacher (cross-ability peer coaching). In the same vein, feedback dialogues could be studied of teachers with a small group of students, and within a group of students coaching each other. In addition, the students can be asked about their perceptions and appreciations of the feedback they received.

Finally, the teachers' learning processes during similar training programmes could be studied in more depth, by investigating how teachers interpret and elaborate upon the new knowledge they receive (information and feedback from the trainer and the students) and how they restructure their existing knowledge, beliefs and intentions.

References

- Aguirre, J., & Speer, N.M. (2000). Examining the relationship between beliefs and goals interteacher practice. *Journal of Mathematical Behavior*, 18 (3), 327-356.
- Airasian, P.W., & Gullickson, A. (1994). Examination of teacher self-assessment. *Journal of Personnel Evaluation in Education*, 8, 195-203.
- Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. In J. Kuhl & J. Beckman (Eds.), *Action-control: from cognition to behavior* (pp. 11-39). Heidelberg, Germany: Springer.
- Ajzen, I., & Fishbein, M. (2005). The influence of attitudes on behavior. In D. Albarracín, B.T. Johnson, & M.P. Zanna (Eds.), *The handbook of attitudes* (pp. 173-221). Mahwah, New Jersey: Erlbaum.
- Aronson, L. (2011). Twelve tips for teaching reflection at all levels of medical education. *Medical Teacher*, 33, 200-205.
- Arter, J., & Chappuis, J. (2007). *Creating & recognizing quality rubrics*. New Jersey: Pearson Education, Inc.
- Asselin, M.E. (2011). Using reflection strategies to link course knowledge to clinical practice: The RN-to-BSN student experience, *Journal of Nursing Education*, 50 (3), 125-133.
- Atkins, S., & Murphy, K. (1993). Reflection: A review of the literature. *Journal of Advanced Nursing*, 18, 1188-1192.
- Ball, D.L., Thames, M.H., & Phelps, G. (2008). Content knowledge for teaching. What makes it special? *Journal of Teacher Education* 59 (5), 389-407.
- Bandura, A. (1977). *Social learning theory*. New York: Prentice-Hall, Inc.
- Banerjee, M., Capozzoli, M., McSweeney, L., & Sinha, D. (1999). Beyond kappa: A review of interrater agreement measures. *The Canadian Journal of Statistics*, 27, 3-23.
- Berk, R.A. (1996). Standard Setting: The next generation (where few psychometricians have gone before!). *Applied Measurement in Education*, 9 (3), 215-235.
- Berthold, K., Nückles, M., & Renkl, A. (2007). Do learning protocols support learning strategies and outcomes? The role of cognitive and metacognitive prompts. *Learning and Instruction*, 17, 564-577.

- Biesma, R.G., Pavlova, M., Vaatstra, R., Van Merode, G.G., Czabanowska, K., Smith, T., & Groot, W. (2008). Generic versus specific competencies of entry-level Public Health graduates: Employers' perceptions in Poland, the UK, and the Netherlands. *Advances in Health Sciences Education* 13, 325-343.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-74.
- Blatt, B., Confessore, S., Kallenberg, G., & Greenberg, L. (2008). Verbal interaction analysis: Viewing feedback through a different lens. *Teaching and Learning in Medicine*, 20 (4), 329-333.
- Bolhuis, S., & Voeten, M.M. (2004). Teachers' conceptions of student learning and own learning. *Teachers and Teaching: Theory and Practice*, 10 (3), 77-98.
- Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher*, 33 (8), 3-15.
- Borko, H., Jacobs, J., Eiteljorg, E., & Pittman, M.E. (2008). Video as a tool for fostering productive discussions in mathematics professional development. *Teaching and Teacher Education*, 24, 417-436.
- Boud, D., Keogh, R., & Walker, D. (1985). Promoting reflection in learning: A model. In D. Boud, R. Keogh & D. Walker (Eds.). *Reflection: Turning experience into learning* (pp.18-40). London: Kogan Page.
- Boursicot, K.A.M., Roberts, T.E., & Pell, G. (2006). Standard setting for clinical competence at graduation from medical school: A comparison of passing scores across five medical schools. *Advances in Health Sciences Education*, 11, 173-183.
- Bradbury-Jones, C., Hughes, S.M., Murphy, W., Parry, L., & Sutton, J. (2009). A new way of reflecting in nursing: The Peshkin approach. *Journal of Advanced Nursing*, 65 (11), 2485-2493.
- Branch, W.T., & Paranjape, A. (2002). Feedback and reflection: Teaching methods for clinical settings. *Academic Medicine*, 77 (12), 1185-1188.
- Bransford, J., Derry, S., Berliner, D., & Hammerness, K. (2005). Theories of learning and their roles in teaching. In: L. Darling-Hammond, J. Bransford, P. LePage, K. Hammerness, & H. Duffy (Eds.). *Preparing teachers for a changing world: What teachers should learn and be able to do* (pp. 40-87). San Francisco: Jossey-Bass.
- Brinko, K.T. (1993). The practice of giving feedback to improve teaching: What is effective? *The Journal of Higher Education* 64 (5), 574-593.
- Burton, A.J. (2000). Reflection: Nursing's practice and education panacea? *Journal of Advanced Nursing*, 31 (5), 1009-1017.
- Butler, D.L. (1998). The strategic content learning approach to promoting self-regulated learning: A report of three studies. *Journal of Educational Psychology*, 90, 682-697.

- Butler, D.L., & Winne, P.H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, 65 (3), 245-281.
- Cader, R., Campbell, S., & Watson, D. (2005). Cognitive continuum theory in nursing decision-making. *Journal of Advanced Nursing* 49 (4), 397-405.
- Cadman Slater, S., & Boulet, J.R. (2001). Predicting holistic ratings of written performance assessments from analytic scoring. *Advances in Health Sciences Education*, 6, 103-119.
- Caris-Verhallen, W.M.C.M., Kerkstra, A., Bensing, J.M., & Grypdonck, M.H.F. (2000). Effects of video interaction analysis training on nurse-patient communication in the care of the elderly. *Patient Education and Counselling*, 39 (1), 91-103.
- Chi, M.T.H. (1996). Constructing self-explanations and scaffolded explanations in tutoring. *Applied Cognitive Psychology*, 10, 33-49
- Chi, M.T.H. (1997). Quantifying qualitative analysis of verbal data: a practical guide. *The Journal of the Learning Sciences*, 6, 271-315.
- Chi, M.T.H., De Leeuw, N., Chiu, M.H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science*, 18, 439-477.
- Chi, M.T.H., Siler, S.A., Jeong, H., Yamauchi, T., & Hausmann, R.G. (2001). Learning from human tutoring. *Cognitive Science*, 25, 471-533.
- Chin, C. (2006). Classroom interaction in Science: Teacher questioning and feedback to students' responses. *International Journal of Science Education*, 11, 1315-1346.
- Clarke, D., & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teaching and Teacher Education*, 18, 947-967.
- Clayton. M.J. (1997). Delphi: a technique to harness expert opinion for critical decision making tasks in education. *Educational Psychology*, 17 (4), 373-386.
- Collins, A., Brown, J.S., & Newman, S.E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L.B. Resnick (Ed.). *Knowing, learning, and instruction: Essays in honour of Robert Glaser* (pp. 453-494). Hillsdale, New Jersey: Erlbaum.
- Colo & Calibris (2007). *Landelijke Kwalificaties MBO, Beroepsgroep: Mbo-verpleegkundige, versie 2007-2008*. [National Qualifications Senior Secondary Vocational Education, Profession group: Senior Secondary Vocational Education-nurse, version 2007-2008]. Zoetermeer/Bunnik: Colo/Calibris. Retrieved December 12, 2007, from <http://www.calibris.nl>.
- Davis, E.A. (2003). Prompting middle school science students for productive reflection: Generic and directed prompts. *The Journal of the Learning Sciences*, 12 (1), 91-142.

- De Bruijn, E., & Leeman, Y. (2011). Authentic and self-directed learning in vocational education: Challenges to vocational educators. *Teaching and Teacher Education, 27*, 694-702.
- De Corte, E. (2000). Marrying theory building and the improvement of school practice: a permanent challenge for instructional psychology. *Learning and Instruction, 10*, 249-266.
- Dekker-Groen, A.M., Van der Schaaf, M.F., & Stokking, K.M. (2011). Teacher competences required for developing reflection skills of nursing students. *Journal of Advanced Nursing, 67* (7), 1568-1579.
- Dekker-Groen, A.M., Van der Schaaf, M.F., & Stokking, K.M. (2012). Performance standards for teachers supporting nursing students' reflection skills development. *Journal of Nursing Education and Practice, 2* (1), 9-19.
- Dekker-Groen, A.M., Van der Schaaf, M.F., & Stokking, K.M. (2013). A teacher competence development programme for supporting students' reflection skills. To be published in *Teachers and Teaching: Theory and Practice*, April, 2013. The iFirst version has been made available on the website of the publisher in November, 2012.
- Desimone, L.M. (2009). Improving impact studies of teachers' professional development: Towards better conceptualizations and measures. *Educational Researcher, 38* (3), 181-199.
- Desimone, L.M., Porter, A.C., Garet, M.S., Suk Yoon, K., & Birman, B.F. (2002). Effects of professional development on teachers' instruction: Results from a three-year longitudinal study. *Educational Evaluation and Policy Analysis, 24* (2), 81-112.
- De Wever, B., Schellens, T., Valcke, M., & Van Keer, H. (2006). Content analysis schemes to analyze transcripts of online asynchronous discussion groups: A review. *Computers & Education, 46*, 6-28.
- Dewey, J. (1933). *How we think: A restatement of the relation of reflective thinking to the educative process*. New York: D.C. Heath and Company.
- Dexter, P., Applegate, M., Backer, J., Claytor, K., Keffer, J., Norton, B., & Ross, B. (1997). A proposed framework for teaching and evaluating critical thinking in nursing. *Journal of Professional Nursing, 13* (3), 160-167.
- Elshout-Mohr, M., Oostdam, R., & Overmaat, M. (2002). Student assessment within the context of constructivist educational settings. *Studies in Educational Evaluation, (28)* 4, 369-390.
- Eraut, M. (1994). *Developing professional knowledge and competence*. London: Falmer Press.
- Ericsson, K.A. (2004). Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Academic Medicine, 79* (10), 70-81.

- Erkens, G. (2005). *Multi Episode Protocol Analysis (MEPA)*. Version 4.10. Retrieved March 2, 2009, from <http://edugate.fss.uu.nl/mepa>.
- Erkens, G., & Janssen, J. (2008). Automatic coding of online collaboration protocols. *International Journal of Computer-Supported Collaborative Learning*, 3, 447-470.
- Ertmer, P., & Newby, T. (1996). The expert learner: Strategic, self-regulated, and reflective. *Instructional Science*, 24, 1-24.
- European Commission Education & Training (2011). *Teacher Education*. Retrieved June 10, 2011, from http://ec.europa.eu/education/school-education/doc832_en.htm
- Evans, L. (2008). Professionalism, professionalism and the development of education professionals. *British Journal of Educational Studies*, 56 (1), 20-38.
- Fenstermacher, G.D. (1987). On understanding the connections between classroom research and teacher change. *Theory into Practice*, 26 (1), 3-7.
- Field, A. (2009). *Discovering statistics using SPSS* (3th ed.). London: Sage Publications Ltd.
- Fishman, B.J., Marx, R.W., Best, S., & Tal, R.T. (2003). Linking teacher and student learning to improve professional development in systemic reform. *Teaching and Teacher Education*, 19, 643-658.
- Flammer, A. (1981). Towards a theory of question asking. *Psychological Research*, 43 (4), 407-420.
- Garet, M.S., Porter, A., Desimone, L., Birman, B.F., & Suk Yoon, K. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, 38 (4), 915-945.
- Garmston, R., Linder, C., & Whitaker, J. (1993). Reflections on cognitive coaching. *Educational Leadership*, 51 (2), 57-61.
- Gé, X., & Land, S.M. (2004). A conceptual framework for scaffolding ill-structured problem-solving processes using question prompts and peer interactions. *Educational Technology Research and Development*, 52 (2), 5-22.
- Gibbs, G., & Simpson, C. (2005). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education*, 3-31.
- Gillespie, A. (2007). The social basis of self-reflection. In J. Valsiner & A. Rosa (Eds.). *The Cambridge handbook of socio-cultural psychology* (pp. 678-691). Cambridge: Cambridge University Press.
- Good, T.L., & Brophy, J.E. (2000). *Looking in classrooms*. New York: Longman.
- Goodman, C.M. (1987). The Delphi technique: a critique. *Journal of Advanced Nursing*, 12 (6), 729-734.

- Goodwin, A.L. (2008). Defining teacher quality. Is consensus possible? In M. Cochran-Smith, S. Feiman-Nemser, D.J. McIntyre, & K.E. Demers (Eds.). *Handbook of research on teacher education. Enduring questions in changing contexts* (pp. 399-403). New York: Routledge, Taylor & Francis Group.
- Gorney, B., & Ness, R.G. (2000). Evaluation dimensions for full-time head coaches at NCAA Division 11 Institutions. *Journal of Personnel Evaluation in Education, 14* (1) 47-65.
- Govaerts, M.J.B. (2008). Educational competencies or education for professional competence? *Medical Education, 42*, 234-236.
- Graesser, A.C., & Person, N.K. (1994). Question asking during tutoring. *American Educational Research Journal, 31*, 104-137.
- Graesser, A.C., Person, N.K., & Hu, X. (2002). Improving comprehension through discourse processing. *New Directions for Teaching and Learning, 89*, 33-44.
- Graesser, A.C., Person, N.K., & Magliano, J.P. (1995). Collaborative dialogue patterns in naturalistic one-to-one tutoring. *Applied Cognitive Psychology, 9*, 495-522.
- Greatorex, J., & Dexter, T. (2000). An accessible analytic approach for investigating what happens between the rounds of a Delphi study. *Journal of Advanced Nursing, 32* (4), 1016-1024.
- Guskey, T.R. (1986). Staff development and the process of teacher change. *Educational Researcher, 15*(5), 5-12.
- Guskey, T.R. (2002). Professional development and teacher change. *Teachers and Teaching: Theory and Practice, 8* (3), 381-391.
- Hafner, J.C., & Hafner, P.M. (2003). Quantitative analysis of the rubric as an assessment tool: An empirical study of student peer-group rating. *International Journal of Science Education, 25*, 1509-1528.
- Hambleton, R.K., & Pitoniak, M.J. (2006). Setting performance standards. In R.L. Brennan (Ed.), *Educational measurement* (4th ed.) (pp. 433-469). Westport: Praeger.)
- Harden, R.M., & Crosby, J. (2000). AMEE Guide 20: The good teacher is more than a lecturer- the twelve roles of the teacher. *Medical Teacher, 22*, 334-347.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77* (1), 81-112.
- Hatton, N., & Smith, D. (1995). Reflection in teacher education: towards definition and implementation. *Teacher and Teacher Education, 11*, 33-49.
- Hawkins, K. & Power, C.B. (1999). Gender differences in questions asked during small decision-making group discussions. *Small Group Research, 30* (2), 235-256.
- Higgins, R., Hartley, P., & Skelton, A. (2002). The conscientious consumer: Reconsidering the role of assessment feedback in student learning. *Studies in Higher Education, 27* (1), 53-64.

- Huck, S.W., & McLean, R.A. (1975). Using a repeated measure ANOVA to analyze the data from a pretest-posttest design: A potentially confusing task. *Psychological Bulletin*, 82 (4), 511-518.
- Ifenthaler, D. (2012). Determining the effectiveness of prompts for self-regulated learning in problem-solving scenarios. *Educational Technology & Society*, 15 (1), 38-52.
- Ilgén, D.R., & Davis, C.A. (2000). Bearing bad news: Reactions to negative performance feedback. *Applied Psychology: An International Review*, 49 (3), 550-565.
- Jeong, A. (2005). A guide to analyzing message-response sequences and group interaction patterns in computer-mediated communication. *Distance Education*, 26 (3), 367-383.
- Jeong, A., Clark, D.B., Sampson, V.D., & Menekse, M. (2011). Sequential analysis of scientific argumentation in asynchronous online discussion environments. In S. Puntambekar, G. Erkens & C. Hmelo-Silver (Eds.), *Analyzing interactions in CSCL: Methodology, approaches, and issues*. New York: Springer.
- Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, 2, 130-144.
- Joyce, B., & Showers, B. (2002). *Student achievement through staff development*. New York/London: Longman.
- Kane, M. (1994). Validating the performance standards associated with passing scores. *Review of Educational Research*, 64 (3), 425-461.
- Kearsley, G.P. (1976). Questions and question asking in verbal discourse: A cross-disciplinary review. *Journal of Psycholinguistic Research*, 5, 355-375.
- Keeney, S., Hasson, F., & McKenna, H. (2006). Consulting the oracle: ten lessons from using the Delphi technique in nursing research. *Journal of Advanced Nursing*, 53 (2), 205-212.
- Kelchtermans, G. (2007). Capturing the multidimensionality of teacher professionalism: broad and deep reflection. In J. van Swet, P. Ponte & B. Smit (Eds.), *Postgraduate programmes as platform: A research-led approach* (pp. 97-109). Rotterdam: Sense Publishers.
- Kennedy, H. (2011). What is video interaction guidance (VIG)? In H. Kennedy, M. Landor & L. Todd (Eds.), *Video interaction guidance. A relationship-based intervention to promote attunement, empathy and wellbeing* (pp. 20-42). London: Jessica Kingsley Publishers.
- Kintsch, E. (2005). Comprehension theory as a guide for the design of thoughtful questions. *Top Long Disorders*, 25 (1), 51-64.
- Kluger, A.N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119 (2), 254-284.

- Koedinger, K.R., & Alevan, V. (2007). Exploring the assistance dilemma in experiments with cognitive tutors. *Educational Psychology Review*, 19, 239-264.
- Korthagen F.A.J., Kessels, J., Koster, B., Lagerwerf, B., & Wubbels, T. (2001). *Linking practice and theory: The pedagogy of realistic teacher education*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Korthagen, F., & Vasalos, A. (2005). Levels in reflection: core reflection as a means to enhance professional growth. *Teaching and Teacher Education*, 11 (1), 47-71.
- Korthagen, F., & Wubbels T. (2001). Evaluative Research on the Realistic Approach and on the Promotion of Reflection. In Korthagen, F.A.J., Kessels, J., B. Koster, B. Lagerwerf & T. Wubbels (Eds.), *Linking practice and theory: the pedagogy of realistic teacher education* (pp. 88-107). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Krathwohl, D.R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into Practice*, 41 (4), 212-218.
- Kuijpers, J.M., Houtveen, A.A.M., & Wubbels, Th. (2010). An integrated professional development model for effective teaching. *Teaching and Teacher Education*, 26, 1687-1694.
- Kwakman, K. (2003). Factors affecting teachers' participation in professional learning activities. *Teaching and Teacher Education*, 19, 149-170.
- LaBoskey, V. (1993). A conceptual framework for reflection in preservice teacher education. In J. Calderhead & P. Gates (Eds.), *Conceptualizing reflection in teacher development* (pp. 23-38). London: Falmer Press.
- Lane, S., & Stone, C.A. (2006). Performance Assessment. In R.L. Brennan (Ed.), *Educational measurement* (4th ed.) (pp. 387-431). Westport: Praeger.
- Lee, H.-J. (2005). Understanding and assessing preservice teachers' reflective thinking. *Teaching and Teacher Education*, 21, 699-715.
- Linstone, H.A., & Turoff, M. (Eds.) (1975). *The Delphi method: Techniques and applications*. London: Addison-Wesley Publishing Company.
- Lizzio, A., & Wilson, K.L. (2004) Action learning in higher education: An investigation of its potential to develop capability. *Studies in Higher Education*, 29, 469-488.
- Mackintosh, C. (1998). Reflection: a flawed strategy for the nursing profession. *Nurse Education Today*, 18 (7), 553-557.
- Mann, K., Gordon, J., & MacLeod, A. (2009). Reflection and reflective practice in health professions education: A systematic review. *Advances in Health Science Education*, 14, 595-621.
- Mansvelder-Longayroux, D., Beijaard, D., & Verloop, N. (2007). The portfolio as a tool for stimulating reflection by student teachers. *Teaching and Teacher Education*, 23 (1), 47-62.

- McKenney, S., Nieveen, N. & Van den Akker, J. (2006). *Design research from a curriculum perspective*. In J. van den Akker, K. Gravemeijer, S. McKenney & N. Nieveen (Eds.), *Educational design research* (pp. 67-90). London: Routledge.
- Meirink, J.A., Meijer, P.C., & Verloop, N. (2007). A closer look at teachers' individual learning in collaborative settings. *Teachers and Teaching: Theory and Practice*, 13 (2), 145-164.
- Mena Marcos, J.J., & Tillema, H. (2006). Studying studies on teacher reflection and action: An appraisal of research contributions. *Educational Research Review*, 1, 112-132.
- Mercer, N. (2004). Sociocultural discourse analysis: Analysing classroom talk as a social mode of thinking. *Journal of Applied Linguistics*, 1 (2), 137-168.
- Mertler, C.A. (2001). Designing scoring rubrics for your classroom. *Practical Assessment, Research and Evaluation*, 7(25), Retrieved June 8, 2011, from <http://PAREonline.net/getvn.asp?v=7&n=25>.
- Mezirow, J. (1991). *Transformative dimensions of adult learning*. San Francisco: Jossey-Bass.
- Ministry of Education, Culture and Science. (2007). *Law on educational professions*. Retrieved November 15, 2007, from <http://www.minocw.nl/onderwijspersoneelkwaliteit/266/Wet-BIO-en-bekwaamheidseisen.html>.
- Morrison, G.R., Ross, S.M., & Kemp, J.E. (2004). *Designing effective instruction* (4th ed.). Hoboken, New York: J. Wiley & Sons.
- Mory, E.H. (2004). Feedback research revisited. In D. Jonassen (Ed.), *Handbook of research on educational communication and technology* (pp. 745-783). Mahwah: Lawrence Erlbaum.
- Movisie (2007). *Beroepscompetentieprofiel MBO-verpleegkundige ten behoeve van onderwijs-experimenten*. [Profession competence profile senior secondary vocational education-nurse for experiments in education]. Retrieved January 10, 2011, from http://www.movisie.nl/116476/def/home/publicaties/publicaties/beroepscompetentieprofiel_mboverpleegkundige
- Murry, J.W.J., & Hammons, J.O. (1995). Delphi: a versatile methodology for conducting qualitative research. *The Review of Higher Education*, 18 (4), 423-436.
- NBPTS (National Board for Professional Teaching Standards). (2011). *The standards*. Retrieved June 10, 2011, from http://www.nbpts.org/the_standards.
- Nicol, D.J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, 31 (2), 199-218.
- Nückles, M., Hübner, S., & Renkl, A. (2009). Enhancing self-regulated learning by writing learning protocols. *Learning and Instruction*, 19, 259-271.
- O'Donovan, M. (2007). Implementing reflection: Insights from pre-registration mental health students. *Nurse Education Today*, 27 (6), 610-616.

- Oosterbaan, A., Van der Schaaf, M., Baartman, L., & Stokking, K. (2010). Reflection during portfolio-based conversations. *International Journal of Educational Research*, 49, 151-160.
- Opfer, V.D., & Pedder, D. (2011). Conceptualizing teacher professional learning. *Review of Educational Research*, 81 (3), 376-407.
- Penuel, W.R., Fishman, B.J., Yamaguchi, R., & Gallagher, L.P. (2007). What makes professional development effective? Strategies that foster curriculum implementation. *American Educational Research Journal*, 44, 921-958.
- Platzer, H., Blake, D., & Ashford, D. (2000). Barriers to learning from reflection: A study of the use of group work with post-registration nurses. *Journal of Advance Nursing*, 31 (5), 1001-1008.
- Platzer, H., & Snelling, J. (1997). Promoting reflective practitioners in nursing: A review of and journals to facilitate reflection. *Teaching in Higher Education*, 2 (2), 103-121.
- Procee, H. (2006). Reflection in education: A Kantian epistemology. *Educational Theory*, 56 (3), 237-253.
- Putman, R., & Borko, H. (1997). Teacher learning: Implications of new views of cognition. In B.J. Biddle, T.L. Good & I.F. Goodson (Eds.), *The international handbook of teachers and teaching* (Vol. 2, pp. 1223-1296). Dordrecht: Kluwer.
- Quinton S., & Smallbone, T. (2010). Feeding forward: Using feedback to promote student reflection and learning – a teaching model, *Innovations in Education and Teaching Innovations*, 47 (1), 125-135.
- Reeve, J., Jang, H., Carrell, D., Jeon, S., & Barch, J. (2004). Enhancing students' engagement by increasing teachers' autonomy support. *Motivation and Emotion*, 28 (20), 147-169.
- Rodgers, C. (2002). Defining reflection: another look at John Dewey and reflective thinking. *Teachers College Record*, 4 (4), 842-866.
- Roelofs, E., & Sanders, P. (2007). Towards a framework for assessing teacher competence. *European Journal of Vocational Training*, 40 (1), 123-139.
- Rosenstein, B. (2000). Video use for program evaluation, a conceptual perspective. *Studies in Education Evaluation*, 26, 373-394.
- Rosenstein, B. (2002). Video use in social science research and program evaluation. *International Journal of Qualitative Methods*, 1 (3), 1-38.
- Ross, J.A., & Bruce, C.D. (2007). Teacher self-assessment: A mechanism for facilitating professional growth. *Teaching and Teacher Education*, 23, 146-159.
- Roter, D.L., Larson, S., Shinitzky, H., Chernoff, R., Serwint, J.R., Adamo, G., & Wissow, L. (2004). Use of an innovative video feedback technique to enhance communication skills training. *Medical Education*, 38, 145-157.

- Ruiz-Primo, M.A., & Furtak, E.M. (2007). Exploring teachers' informal formative assessment practices and students' understanding in the context of scientific inquiry. *Journal of Research in Science Teaching*, 44 (1), 57-84.
- Sadler, D.R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119-144.
- Sadler, D.R. (1998). Formative assessment: Revisiting the territory. *Assessment in Education: Principles, Policy and Practice*, 5, 77-85.
- Sadler, D.R. (2010). Beyond feedback: Developing student capability in complex appraisal. *Assessment and Evaluation in Higher Education*, 35(5), 535-550.
- SCB (Stichting Consortium Beroepsoponderwijs) [Foundation Consortium Vocational Education]. (2007). *Zorg en welzijn* [Care and welfare]. Retrieved November 15, 2007, from <http://www.consortiumbo.nl/mbo/zorg-and-welzijn-en-ag>.
- Schön, D.A. (1987). *Educating the reflective practitioner. Towards a new design for teaching and learning in the professions*. San Francisco-London, Jossey-Bass.
- Schön, D.A. (1991). *The reflective practitioner. How professionals think in action*. Aldershot Hants: The Academic Publishing Group.
- Seligman, M.E.P., & Csikszentmihalyi, M. (2000). Positive psychology: An introduction. *American Psychologist*, 55, 5-14.
- Shuell, T.J. (1988). The role of the student in learning from instruction. *Contemporary Educational Psychology*, 13, 276-295.
- Shuell, T.J. (1996). Teaching and learning in a classroom context. In D.C. Berliner, & R.C. Calfee (Eds.), *Handbook of Educational Psychology* (pp. 726-763). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Shulman, L.S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57 (1), 1-22.
- Shute, V.J. (2008). Focus on formative feedback. *Review of Educational Research*, 78 (1), 153-189.
- Sinatra, G.M., & Pintrich, P.R. (2010). The role of intentions in conceptual change learning. In G.M. Sinatra & P.R. Pintrich (Eds.), *Intentional conceptual change* (pp. 1-18). New York: Routledge.
- Sireci, S. (2005). The most frequently unasked questions about testing. In R. Phelps (Ed.), *Defending standardized testing* (pp. 111-121). Mahwah, New Jersey: Erlbaum.
- Smith, K. (2011). Professional development of teachers – A prerequisite for Afl to be successfully implemented in the classroom. *Studies in Educational Evaluation*, 37, 55-61.
- Smith, K.S., & Simpson, R.D. (1995). Validating teaching competences for faculty members in higher education: a national study using the Delphi method. *Innovative Higher Education*, 19 (3), 223-234.

- Steinert, Y., Mann, K., Centeno, A., Dolmans, D., Spencer, J., Gelula, M., & Prideaux, D. (2006). A systematic review of faculty development initiatives designed to improve teaching effectiveness in medical education. *Medical Teacher, 28* (6), 497-526.
- Strijbos, J.W., & Stahl, G. (2007). Methodological issues in developing a multi-dimensional coding procedure for small-group chat communication. *Learning and Instruction, 17*, 394-404.
- Sultana, R.G. (2009). Competence and competence frameworks in career guidance: Complex and contested concepts. *International Journal Educational Vocational Guidance, 9*, 15-30.
- Tanner, C.A. (2006). Thinking like a nurse: A research-based model of clinical judgment in nursing. *Journal of Nursing Education, 45* (6), 204-211.
- Teekman, B. (2000). Exploring reflective thinking in nursing practice. *Journal of Advanced Nursing, 31* (5), 1125-1135.
- Thillmann, H., Künsting, J., Wirth, J., & Leutner, D. (2009). Is it merely a question of 'What' to prompt or also 'When' to prompt? *Zeitschrift für Pädagogische Psychologie, 23* (2), 105-115.
- Thomas Farrar, M.T. (1986). Teacher questions: The complexity of the cognitively simple. *Instructional Science, 15* (1), 89-107.
- Thompson, I. (2009). Scaffolding in the writing centre: A microanalysis of an experienced tutor's verbal and nonverbal tutoring strategies. *Written Communication, 26* (4), 417- 453.
- Tillema, H.H., & Imants, J.G.M. (1995). Training in the professional development of teachers. In T.R. Guskey & M. Huberman (Eds.), *Professional development in education: New paradigms and practices* (pp. 135-150). New York: Teachers College Press.
- Tillema, H.H., & Smith, K. (2009). Assessment orientation in formative assessment of learning to teach. *Teachers and Teaching: Theory and Practice, 15* (3), 391-405.
- Tillema, H.H., & Veenman, S.A.M. (1987). Conceptualizing training methods in teacher education. *International Journal of Educational Research, 11* (5), 519-529.
- Timperley, H., Wilson, A., Barrar, H., & Fung, I. (2007). *Teacher professional learning and development*. New Zealand: Ministry of Education.
- Van den Boom, G., Paas, F., & Van Merriënboer, J.J.G. (2007). Effects of elicited reflections combined with tutor or peer feedback on self-regulated learning and learning outcomes. *Learning and Instruction, 17*, 532-548.
- Van der Schaaf, M., Baartman, L., Prins, F., Oosterbaan, A., & Schaap, H. (2011). Feedback dialogues that stimulate students' reflective thinking. *Scandinavian Journal of Educational Research, 55* (3), 273-289.

- Van Es, E.A., & Sherin, M.G. (2008). Mathematics teachers' 'learning to notice' in the context of a video club. *Teaching and Teacher Education, 24*, 244-276.
- Van Es, E.A., & Sherin, M.G. (2010). The influence of video clubs on teachers' thinking and practice. *Journal of Mathematics Teacher Education, 13*, 155-176.
- Van Zee, E., & Minstrell, J. (1997). Using questioning to guide student thinking. *The Journal of the Learning Sciences, 6* (2), 227-269.
- Veenman, S., & Denessen, E. (2001). The coaching of teachers: Results of five training studies. *Educational Research and Evaluation, 7* (4), 385-417.
- Vermunt, J., & Verloop, N. (1999). Congruence and friction between learning and teaching. *Learning and Instruction, 9*, 257-280.
- Vollmeyer, R., & Rheinberg, F. (2005). A surprising effect of feedback on learning. *Learning and Instruction, 15*, 589-602.
- Vygotsky, L. (1978). *Mind in society. The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Wallman, A., Kettis Lindblad, A., Hall, S., Lundmark, A., & Ring, L. (2008). A categorization scheme for assessing pharmacy students' levels of reflection during internships. *American Journal of Pharmaceutical Education, 72*, 1-10.
- Wampold, B.E., & Margolin, G. (1982). Nonparametric strategies to test independence of behavioural states in sequential data. *Psychological Bulletin, 92*, 755-765.
- Whitcomb, J., Borko, H., & Liston, D. (2009). Growing talent: Promising professional development models and practices. *Journal of Teacher Education, 60*, 207-212.
- Williams, P.L., & Webb, C. (1994). The Delphi technique: a methodological discussion. *Journal of Advanced Nursing, 19* (1), 180-186.
- Williamson, S.N. (2007). Development of a self-rating scale of self-directed learning. *Nurse Researcher, 14* (2), 66-83.
- Wirth, J. (2009). Promoting self-regulated learning through prompts. *Zeitschrift für Pädagogische Psychologie, 23* (2), 91-94.
- Wong, F.K.Y., Kember, D., Chung, L.Y.F., & Yan, L. (1995). Assessing the level of student reflection from reflective journals. *Journal of Advanced Nursing, 22*, 48-57.
- Yang, M., & Carless, D. (2012). The feedback triangle and the enhancement of dialogic feedback processes. *Teaching in Higher Education, 1-13*, iFirst article.
- Zimmerman, B.J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice, 41* (2), 64-70.

Summary

Schools in Dutch senior secondary vocational education currently are expected to offer competence-based education. This is an educational innovation that aims to improve the preparation of future professionals for continuous changes in their work. It is necessary that professional workers, for example nurses, are flexible and able to reflect on their activities and behaviour and self-regulate their work. Therefore during their education nursing students have to develop competences for lifelong learning and professional development during their career. Supporting nursing students to develop such skills asks for new competences of their teachers. Little is known about teacher competences required for supporting the development of nursing students' reflection skills.

The project reported about in this thesis aimed at competence development of teachers in giving feedback on and supporting reflection skills of nursing students. Feedback is a key element in formative assessment, both as information about how successful someone's functioning is and for suggesting how to alter the gap between current and desired functioning (Sadler, 1989; 2010). The purpose of formative assessment is furthering the learner's development, which assumes a feedback process in which a learner is actively involved. We will use the term feedback in the broad sense of formative assessment: a process of communicating information to support both students' and teachers' competence development.

Effectively contributing to teachers' development of competences in giving feedback on students' reflections demands a combination of four approaches: (1) instruction, exercise, and feedback (training-application approach, Joyce & Showers, 2002; Tillema & Veenman, 1987); (2) discussion and exchange of experiences (social learning, Bandura, 1977); (3) attention for the teachers' current practices and their reflection on alternatives (developmental-reflective approach, Kuijpers, Houtveen & Wubbels, 2010); (4) giving information and using relevant knowledge (expert-based learning, Timperley, Wilson, Barrar & Fung, 2007).

The central research question was: *How can teachers' competences in supporting nursing students' reflection skills be developed by a teacher training programme?* The research project contained five studies in senior secondary nursing education.

In Study 1 content standards were developed for education focused on supporting students' learning how to reflect. Content standards specify what teachers are expected to know and be able to do. The research question was: *What are relevant*

and acceptable descriptions of teacher competences for supporting nursing students' development of reflection skills?

A literature study and a study of national policy documents and curriculum materials offered insight into the concepts of reflection and feedback, national examination requirements for nursing education, and current innovations in Dutch vocational education. We defined reflection as a conscious and well-considered process of thinking about and interpretation of situations, events, processes, experiences and emotions, in relation to each other and to already available knowledge, aiming for better comprehension and for learning (Ertmer & Newby, 1996; Hatton & Smith, 1995; Kelchtermans, 2007). We distinguished eight thinking activities students may use during reflection: describing, analysing, structuring, evaluating, explaining, attributing, concluding and intending (Krathwohl, 2002; Oosterbaan, Van der Schaaf, Baartman & Stokking, 2010; Shuell, 1988). This resulted in pedagogical recommendations about how teachers might create a learning environment to support nursing students' reflection skills. Developing students' reflection skills often takes place during conferences in which feedback and interaction are important elements (Chi, Siler, Jeong, Yamauchi & Hausmann, 2001; Graesser & Person, 1994; Shute, 2008).

The development of a framework of teacher competences that stimulate the development of students' reflection skills was based on the general components of instructional design: learners, objectives, methods, and evaluation (Morrison, Ross & Kemp, 2004). This framework was input for a Delphi study in which 28 experts participated, including teachers and managers in nursing education. The framework contained three levels of specification: six task domains, covering 15 tasks, which represent 91 indicators. On all three levels the judgments by the experts resulted in high scale reliabilities. After two rounds a satisfying level of consensus was reached, based on minimal 75% agreement for acceptance of each indicator (Keeney, Hasson & McKenna, 2006). The experts ranked coaching as the most important competence nursing teachers should practise and as the second most important that they should develop competences to distinguish and stimulate students' thinking activities. The framework resulted in detailed descriptions about how to develop nursing students' reflection skills and how to improve education in supporting students' reflection.

In study 2 performance standards were developed and validated for nursing teachers who have to support students' reflection skills development. Performance standards specify norms for the frequency, intensity or quality of teachers' behaviour specified in the content standards. Also a judgmental model was established for making overall judgments based on the individual scores for each

task on each standard. The research questions were: *What are relevant and acceptable performance standards for teachers who have to support nursing students' development of reflection skills? Which judgmental model is, according to nursing teacher teams, convenient for assessing the teachers' competences?*

Rubrics were chosen as method to describe performance standards at four levels: Beginning, Basic, Advanced, Expert. The content standards developed in study 1 were used to develop a rubrics framework of 30 rubric attributes spread over eight competences covering five competence domains: (1) *Interpersonal*: Effective communication with individual and group; (2) *Pedagogical*: Care for a positive learning climate; (3) *Subject matter and pedagogics*: a. Give information and instruction how to reflect, b. Coach the learning process of how to reflect, c. Give feedback and teach how to receive and to ask for feedback, d. Use criteria to assess reflections; (4) *Organisational*: Design and organise the learning environment and (5) *Professional development*: Work individually and in cooperation with colleagues.

Striving for a procedure as complete and unambiguous as possible we combined three ways of standard setting (Berk, 1996; Hambleton & Pitoniak, 2006; Kane, 2004) in a standard-setting procedure including ten activities. This procedure could be executed as planned. The developed rubrics framework was discussed in six teams by in total 40 teachers and managers. The choice of the rubrics method showed to be effective because it stimulated discussions within teams and the framework appeared to be relevant and acceptable for all participants. Procedural and internal validation evidence justified the conclusion that the standard-setting procedure, including the rubrics method, had contributed to consensus within and between the participating teacher teams in nursing education. The teacher teams unanimously chose for a conjunctive judgmental model, meaning that for each standard a teacher should reach a minimum score at level two (Basic).

In study 3 a training programme was developed and conducted at five schools of nursing education. The programme focussed on the four teacher competences regarding subject matter and pedagogics: instructing, guiding, giving feedback on, and assessing students' reflection skills. In the development of this programme five research-based core features of effective teacher professional development were involved: content focus, active learning, coherence, duration, and collective participation (Desimone, 2009; Garet, Porter, Desimone, Birman & Suk Yoon, 2001; Penuel, Fishman, Yamaguchi & Gallagher, 2007). The research question was: *What are the feasibility, quality and effects of the programme?* A quasi experimental design was used in which nineteen teachers took part in the training programme and eight teachers formed a control group.

With regard to the feasibility of the programme it was concluded that three of the five core features for effective professional development were realised. The programme had a strong content focus, active learning was realised, and coherence in content and activities was provided. The programme was spread over 12-15 months to ensure the teachers would have enough time for independent practice. The time the teachers could devote to the programme however remained modest and collective participation did not happen as a team. Concerning the *quality* of the programme it was concluded that participants on average were fairly satisfied with the programme. The video interaction guidance in the individual meetings was especially appreciated.

Indications for effects of the programme did not represent overwhelming evidence, but all effects found were in the intended direction. Participants reported on average an increase in focus regarding the content of students' reflections and in self confidence. They got another view on guiding reflection, and self-estimated some competence development in conducting specific activities to support students' reflection. Self assessments showed a significant increase on the competence 'Assessing reflections'. Moreover the extent to which students reflected on a number of possible objects of reflection was higher after the training than before and, controlled for pre-measurements, also higher than in the control group. The overall conclusion is that teachers who want to actively develop their competences in supporting students' reflection skills and are offered a coherent, purposeful programme including instruction, exercise in their own setting, feedback and reflect on their activities, can improve their relevant practice with a fairly modest investment of their time.

In studies 4 and 5 we analysed verbal interactions between a teacher and an individual student to explore their feedback dialogues and to establish the effects of the programme. In both studies a quasi experimental design was used in which sixteen teachers participated in the programme and seven teachers formed a control group. In total 46 videotaped feedback dialogues were analysed. These dialogues lasted about 20 minutes in the pre-test and 25 minutes in the post-test. Per minute the teacher and the student had four or five turns. The results indicate that per turn the teachers in the training group became more efficient.

In study 4 we explored which single questions and responses were used and which sequences occurred. Effectiveness of a teacher's feedback presupposes students' self-regulation (Zimmerman, 2000). This self-regulation appears to be supported by teachers giving prompts: questions and responses which focus students on certain cognitive processes and evoke specific reflective activities (Ifenthaler, 2012; Thillmann, Künsting, Wirth & Leutner, 2009). The research questions were: *Which*

questions and responses are given in feedback dialogues between nursing teachers and individual students? What are the effects of the training programme on the questions and responses of teachers and individual students in feedback dialogues?

We combined question categories based on the Graesser-Person-Huber scheme with prompt questions (Graesser & Person, 1994; Ifenthaler, 2012; Nückels, Hübner & Renkl, 2009). We distinguished seven question categories (1) Recall; (2) Hybrid; (3) Hybrid prompt; (4) Request; (5) Surface reasoning; (6) Deep reasoning; (7) Reasoning prompt. We used response categories based on Blatt, Confessore, Kallenberg and Greenberg (2008): (1) Neutral; (2) Deny; (3) Accept and added as (4) Prompt. Teachers in the control group used about the same question categories in the same frequency in the pre-test and the post-test. They only posed more Recall questions in the post-test, a question category that is often meant to get a short answer of the student. In the post-test the training group used more Prompt responses, less Recall, Hybrid and Surface reasoning questions and more Hybrid prompt, Deep reasoning and Reasoning prompt questions. The latter question categories have the potential to let students think and express their thinking process. We concluded that the teachers in the training group became accustomed to a broader repertoire of questions, especially deep reasoning questions and reasoning prompts. Their dialogues contained more variation in sequences including more questions and prompts which potentially stimulate students to elaborate and to reflect (Nückles et al., 2009; Wirth, 2009). Therefore we assume that the programme did contribute to the teachers' competence development to support students' reflection.

In study 5 we analysed which feedback categories teachers used and how the students reacted. Essential is teacher feedback that stimulates students to reflection and deeper learning (Higgins, Hartley & Skelton, 2002; Quinton & Smallbone, 2010). The research questions were: *What is the relation between teachers' feedback and students' reflection? What are the effects of the training programme on the feedback repertoire of teachers and the utterances of individual students in feedback dialogues?*

Feedback was operationalised in five categories: one Suggestive, two Didactic and two Corrective categories (Chi, 1996; Chi et al., 2001; Chin, 2006; Shute 2008). Suggestive feedback concerned Summarise, Check, Focus, Diagnose, and Forward, Didactic feedback concerned Expound and Modeling, and Corrective feedback concerned Appreciate positive and Appreciate negative. Reflection supposes cognitive, affective and regulative activities (e.g. Mansvelder-Longayroux, Beijaard & Verloop, 2007; Oosterbaan, Van der Schaaf, Baartman, & Stokking, 2010). We operationalised these activities in eight thinking activities: Describe, Analyse, Structure, Evaluate, Attribute, Explain, Conclude and Intend

(based on study 1: Dekker-Groen, Van der Schaaf & Stokking, 2011). Moreover we added as student categories Content Information (Summarise and Amplify) and Content Appreciation (positive and negative).

In the post-test compared to the pre-test, teachers in the control group used on average more Check and less Expound. Teachers in the training group used on average especially more Focus and Appreciate positive. Focus evoked far the most thinking activities. The latter appeared significantly more with Explain and Conclude. Also sequences concerning two times Suggestive feedback, Suggestive and Corrective feedback or two times Corrective feedback evoked thinking activities. In the post-test teachers in the training group used less Expound and less Check and their students used less Amplify. These results might indicate that teachers and students comprehended each other better and that students got more autonomy. Although the differences in the use of the categories concerning the training group between the two moments were small, our careful conclusion was that the training in giving feedback seemed to contribute to changes in teachers' feedback repertoire during feedback dialogues which have potential to support students' reflection.

Looking back to the central research question: *How can teachers' competences in supporting nursing students' reflection skills be developed by a teacher training programme?* we conclude that the conducted programme was a combination of three of four modern training approaches discussed. First instruction, exercise and feedback were typical for a training-application approach. Second, discussion about current practice and reflection on alternatives fitted into a developmental-reflective approach. Third, giving information about and using relevant knowledge indicated an expertise-based approach. Social learning was less central, as the time available for discussion and exchange between the teachers was restricted.

Two additional specific features were used. First, the feedback dialogues between the teachers and their students were authentic situations. Video clips of feedback dialogues were used in the meetings with the teachers during the programme. Second, the coaching of the teachers by the trainer was done using video interaction guidance, and this might have stimulated their reflection and empowerment. Students seemed to be sensitive for teachers' reasoning questions and prompts. These are facilitators for students' development of self-regulation. One barrier that will have restricted the teachers' and students' learning and development was probably institutional: the permanent turbulence in the schools.

Samenvatting

Van scholen voor Middelbaar Beroepsonderwijs in Nederland wordt tegenwoordig verwacht dat ze competentiegericht onderwijs aanbieden. Dit is een onderwijsvernieuwing die is bedoeld om de voorbereiding te verbeteren van toekomstige professionals op continue veranderingen in hun werk. Professionele werknemers, zoals verpleegkundigen, moeten flexibel zijn, hun werk zelf regelen en kunnen reflecteren op hun activiteiten en gedrag. Studenten Verpleegkunde moeten daarom al tijdens hun opleiding competenties ontwikkelen voor levenslang leren om zich professioneel te blijven ontwikkelen gedurende hun beroepsloopbaan. Het ondersteunen van studenten Verpleegkunde om dergelijke vaardigheden te ontwikkelen vraagt van hun docenten nieuwe competenties. Er is weinig bekend over docentcompetenties die nodig zijn voor het ondersteunen van studenten Verpleegkunde bij de ontwikkeling van hun reflectievaardigheden.

Het onderzoeksproject waarover in dit proefschrift wordt gerapporteerd, richt zich op de competentieontwikkeling van docenten in het geven van feedback op en het ondersteunen van reflectievaardigheden van studenten Verpleegkunde. Feedback is een sleutelement in formatieve beoordeling: het geeft zowel informatie over hoe succesvol iemand functioneert als suggesties hoe de kloof tussen huidig en gewenst functioneren kan worden overbrugd (Sadler, 1989; 2010). Het doel van formatieve beoordeling is het bevorderen van de ontwikkeling van de lerende, hetgeen een feedbackproces veronderstelt waarin een lerende actief betrokken is. In dit proefschrift wordt de term feedback gebruikt in de brede betekenis van formatieve beoordeling: een proces van informatie communiceren om de competentieontwikkeling te ondersteunen van zowel studenten als docenten.

Effectief bijdragen aan de ontwikkeling van docentencompetenties in het geven van feedback op reflecties van studenten vraagt een combinatie van vier aanpakken: (1) instructie, oefening en feedback (training & toepassing, Joyce & Showers, 2002; Tillema & Veenman, 1987); (2) discussie over en uitwisseling van ervaringen (sociaal leren, Bandura, 1977); (3) aandacht voor de huidige praktijken van docenten en hun reflecties op alternatieven (ontwikkeling & reflectie, Kuijpers, Houtveen & Wubbels, 2010); (4) informatie geven en relevante kennis gebruiken (op kennis gebaseerd leren, Timperley, Wilson, Barrar & Fung, 2007).

De centrale onderzoeksvraag is: *Hoe kunnen docentcompetenties voor het ondersteunen van de ontwikkeling van reflectievaardigheden van studenten Verpleegkunde worden ontwikkeld door middel van een trainingsprogramma?*

Het onderzoeksproject omvat vijf studies in het Middelbaar Beroepsonderwijs.

In studie 1 zijn inhoudelijke standaarden ontwikkeld voor onderwijs gericht op het ondersteunen van het leren van studenten hoe te reflecteren. Inhoudelijke standaarden specificeren wat van docenten verwacht wordt aan kennis en kunde. De onderhavige onderzoeksvraag: *Wat zijn relevante en acceptabele omschrijvingen van docentcompetenties voor het ondersteunen van studenten Verpleegkunde in de ontwikkeling van hun reflectievaardigheden?*

Een literatuurstudie en een studie van nationale beleidsdocumenten en curriculum materialen hebben inzicht geboden in de concepten reflectie en feedback, nationale examenvereisten voor Verpleegkundeonderwijs en huidige innovaties in het beroepsonderwijs in Nederland. We definiëren reflectie als een bewust en weloverwogen proces van denken over en interpreteren van situaties, gebeurtenissen, processen, ervaringen en emoties, in relatie tot elkaar en tot al beschikbare kennis, bedoeld om deze beter te begrijpen en ervan te leren (Ertmer & Newby, 1996; Hatton & Smith, 1995; Kelchtermans, 2007). We onderscheiden acht denkactiviteiten die studenten kunnen gebruiken als ze reflecteren: beschrijven, analyseren, structureren, evaluaeren, verklaren, attribueren, concluderen en intenties formuleren (Krathwohl, 2002; Oosterbaan, Van der Schaaf, Baartman & Stokking, 2010; Shuell, 1988). Dit heeft geleid tot aanbevelingen voor hoe docenten een leeromgeving kunnen creëren om reflectievaardigheden van studenten Verpleegkunde te ondersteunen. De ontwikkeling van reflectievaardigheden van studenten vindt dikwijls plaats tijdens onderwijsleergesprekken waarin feedback en interactie belangrijke elementen zijn (Chi, Siler, Jeong, Yamauchi & Hausmann, 2001; Graesser & Person, 1994; Shute, 2008).

De ontwikkeling van een raamwerk van docentcompetenties die de ontwikkeling van reflectievaardigheden van studenten stimuleren, is gebaseerd op de algemene componenten van instructie-ontwerp: lerenden, doelstellingen, methoden en evaluatie (Morrison, Ross & Kemp, 2004). Dit raamwerk is het uitgangspunt geweest voor een Delphi studie waaraan 28 experts hebben deelgenomen, inclusief docenten en managers in Verpleegkundeonderwijs. Het raamwerk bevat drie niveaus van specificatie: zes taakdomeinen met 15 taken en 91 indicatoren. Op alle drie niveaus hebben de oordelen van de experts tot hoge schaalbetrouwbaarheden geleid. Na twee ronden was een voldoende mate van consensus bereikt, gebaseerd op minimaal 75% overeenstemming over de acceptatie van elke indicator (Keeney, Hasson & McKenna, 2006). De experts rangschikten coaching als de meest belangrijke competentie die docenten Verpleegkunde in de praktijk kunnen

toepassen en als tweede dat ze competenties dienen te ontwikkelen om denkactiviteiten van studenten te onderscheiden en te stimuleren. Het raamwerk heeft geleid tot gedetailleerde omschrijvingen over hoe reflectievaardigheden van studenten Verpleegkunde te ontwikkelen en hoe het aangeboden onderwijs te verbeteren ter ondersteuning van het leren reflecteren door studenten.

In studie 2 zijn ‘performance’ standaarden ontwikkeld en gevalideerd voor docenten Verpleegkunde die de ontwikkeling van reflectievaardigheden van studenten ondersteunen. Performance standaarden specificeren normen voor de frequentie, intensiteit of kwaliteit van het docentgedrag dat is uitgewerkt in de inhoudelijke standaarden. Ook is een beoordelingsmodel vastgesteld voor het geven van een samenvattend oordeel gebaseerd op de individuele scores per taak per standaard. De onderzoeksvragen zijn: *Wat zijn relevante en acceptabele performance standaarden voor docenten die de ontwikkeling van reflectievaardigheden moeten ondersteunen van studenten Verpleegkunde? Welk beoordelingsmodel is, volgens docententeams Verpleegkunde, geschikt om de docentcompetenties te beoordelen?*

Er is gekozen om performance standaarden te beschrijven in rubrieken met vier niveaus: Beginnend, Basis, Gevorderd, Expert. De inhoudelijke standaarden uit studie 1 zijn gebruikt om een raamwerk van rubrieken te ontwikkelen met 30 rubriekattributen, verspreid over acht competenties binnen vijf competentiedomeinen: (1) *Interpersoonlijk*: Effectieve communicatie met individu en groep; (2) *Pedagogisch*: Zorgen voor een positief leerklimaat; (3) *Vakinhoudelijk-didactisch*: a. Informatie en instructie geven hoe te reflecteren, b. Coachen van het leerproces van hoe te reflecteren, c. Feedback geven en onderwijzen hoe feedback te ontvangen en te vragen, d. Criteria gebruiken om reflecties te beoordelen; (4) *Organisatorisch*: Ontwerp en organisatie van de leeromgeving en (5) *Professionele ontwikkeling*: Individueel werken en in samenspraak met collega’s.

Strevend naar een procedure die zo compleet en ondubbelzinnig als mogelijk is, hebben we drie manieren van standaard setting gecombineerd (Berk, 1996; Hambleton & Pitoniak, 2006; Kane, 2004) in een standaard-setting procedure die tien activiteiten omvat. Deze procedure is uitgevoerd zoals gepland. Het ontwikkelde raamwerk van rubrieken is ter discussie gesteld in zes teams van in totaal 40 docenten en managers. De keuze van de rubrieken als methode is zichtbaar effectief geweest, omdat het stimuleerde tot discussies binnen teams en het raamwerk relevant en acceptabel is gebleken voor alle deelnemers. De verkregen bewijzen voor procedurele en interne validiteit onderschrijven de conclusie dat de standaard-setting procedure, inclusief het gebruik van rubrieken, heeft bijgedragen aan consensus binnen en tussen de deelnemende docententeams in het Verpleegkunde onderwijs. De docententeams hebben namelijk unaniem

voor een conjunctief beoordelingsmodel gekozen, hetgeen betekent dat een docent voor elke standaard een bepaalde minimum score dient te bereiken, en wel op niveau twee ('Basic').

In studie 3 is een trainingsprogramma ontwikkeld en uitgevoerd op vijf scholen voor Verpleegkundeonderwijs. Het programma was gericht op de vier vakinhoudelijk-didactische docentcompetenties: instructie geven, begeleiden, feedback geven en beoordelen van reflectievaardigheden van studenten. In de ontwikkeling van het programma zijn vijf op onderzoek gebaseerde kernelementen betrokken van effectieve, professionele ontwikkeling van docenten: inhoudelijk focus, actief leren, samenhang, duur en collectieve deelname (Desimone, 2009; Garet, Porter, Desimone, Birman & Suk Yoon, 2001; Penuel, Fishman, Yamaguchi & Gallagher, 2007). De onderhavige onderzoeksvraag: *Wat zijn de uitvoerbaarheid, kwaliteit en effecten van het programma?* Een quasi-experimenteel ontwerp is gebruikt waarin negentien docenten hebben deelgenomen aan het trainingsprogramma en acht docenten een controlegroep vormden.

Wat de *uitvoerbaarheid* van het programma betreft is geconcludeerd dat drie van de vijf kernelementen van effectieve professionele ontwikkeling zijn gerealiseerd. Het programma had een sterk inhoudelijke focus, actief leren werd gerealiseerd, en in samenhang tussen inhoud en activiteiten was voorzien. Het programma strekte zich uit over 12-15 maanden om er zeker van te zijn dat de docenten genoeg tijd zouden hebben om zelfstandig te oefenen. De tijd die de docenten daadwerkelijk aan het programma hebben kunnen besteden, bleef echter bescheiden en collectieve deelname gebeurde niet als team. Wat betreft de *kwaliteit* van het programma is geconcludeerd dat de deelnemers gemiddeld genomen tamelijk tevreden zijn met het programma. De video-interactie begeleiding in de individuele bijeenkomsten werd het meeste gewaardeerd.

De indicaties voor *effecten* van het programma vormden geen overweldigend bewijs, maar alle gevonden effecten wezen in de beoogde richting. Deelnemers rapporteerden gemiddeld genomen een toename van hun focus op de inhoud van reflecties van studenten en in eigen zelfvertrouwen. Zij hebben een andere kijk op het begeleiden van reflecteren ontwikkeld en ervaren zelf enige competentieontwikkeling in het uitvoeren van specifieke activiteiten ter ondersteuning van het reflecteren van studenten. Zelfbeoordelingen lieten een significante toename zien op de competentie 'Beoordelen van reflecties'. Bovendien was de mate waarin studenten reflecteren op een aantal mogelijke reflectie-objecten hoger na de training dan ervoor en, rekeninghoudend met de verschillen op de voormeting, ook hoger dan in de controlegroep. De algemene conclusie is dat docenten die hun competenties in het ondersteunen van reflectievaardigheden van studenten actief willen ontwikkelen, deze kunnen verbeteren met een vrij bescheiden

investering van hun tijd, indien zij een samenhangend, doelgericht programma krijgen aangeboden met daarin instructie, oefening in de eigen werkomgeving, feedback en reflectie op hun activiteiten in de voor hen relevante praktijksituatie.

In de studies 4 en 5 analyseerden we mondelinge interacties tussen een docent en een individuele student om hun feedbackdialogen te verkennen en de effecten van het programma vast te stellen. In beide studies is een quasi-experimenteel ontwerp gebruikt waarbij zestien docenten hebben deelgenomen aan het programma en zeven docenten een controlegroep vormden. In totaal zijn 46 feedbackdialogen geanalyseerd die op video zijn opgenomen. Deze dialogen duurden ongeveer 20 minuten in de voormeting en 25 minuten in de nameting. Per minuut voerden de docent en de student elk vier of vijf keer het woord. De resultaten geven aan dat de docenten in de trainingsgroep in de nameting minder spreektijd gebruikten.

In studie 4 bestudeerden we welke vragen en antwoorden worden gesteld en gegeven en welke volgordepatronen zich voordoen. Effectiviteit van feedback van een docent veronderstelt zelfregulatie van de studenten (Zimmerman, 2000). Deze zelfregulatie lijkt te worden ondersteund door docenten die aanmoedigingen (prompts) geven door vragen te stellen en antwoorden te geven die studenten richten op bepaalde cognitieve processen en specifieke reflectieve activiteiten oproepen (Ifenthaler, 2012; Thillmann, Künsting, Wirth & Leutner, 2009). De betreffende onderzoeksvragen zijn: *Welke vragen en antwoorden worden gegeven in feedbackdialogen tussen docenten Verpleegkunde en individuele studenten? Wat zijn de effecten van het trainingsprogramma op de vragen en antwoorden van docenten en individuele studenten in feedbackdialogen?*

Om dit te onderzoeken hebben we vraagcategorieën gebaseerd op het Graesser-Person-Huber schema gecombineerd met aanmoedigingsvragen (prompts) (Graesser & Person, 1994; Ifenthaler, 2012; Nückels, Hübner & Renkl, 2009). We onderscheiden zeven vraagcategorieën (1) vraag gericht op herinneren; (2) hybride vraag (kan leiden tot herinneren of redeneren); (3) hybride prompt (aanmoediging); (4) verzoek; (5) oppervlakkig redeneren; (6) diepgaand redeneren; (7) redeneer prompt. We gebruiken daarbij antwoordcategorieën gebaseerd op Blatt, Confessore, Kallenberg en Greenberg (2008): (1) neutraal; (2) ontkenning; (3) acceptatie, en voegen (4) prompt toe. Docenten in de controlegroep hebben ongeveer dezelfde vraagcategorieën gebruikt in dezelfde frequentie in de voormeting en de nameting. Zij stelden alleen meer herinneringsvragen in de nameting, een vraagcategorie die dikwijls is bedoeld om een kort antwoord te krijgen van de student. In de nameting gebruikte de trainingsgroep meer prompt antwoorden, minder herinnerings-, hybride en oppervlakkig redeneren vragen en meer hybrid prompt, diepgaand redeneren en redeneer prompt vragen. De laatstgenoemde categorieën hebben de potentie om studenten te laten nadenken

en hun denkproces tot uiting brengen. We concluderen dat de docenten in de trainingsgroep vertrouwd raakten met een breder repertoire aan vragen, in het bijzonder diepgaand redeneren en redeneer vragen. Hun dialogen bevatten meer variatie in volgordepatronen met daarin meer vragen en prompt antwoorden die studenten mogelijk stimuleren tot uitweiden en reflecteren (Nückles et al., 2009; Wirth, 2009). We veronderstellen daarom dat het programma heeft bijgedragen aan de competentie-ontwikkeling van docenten in het ondersteunen van het reflecteren van studenten.

In studie 5 analyseerden we welke feedbackcategorieën docenten gebruiken en hoe de studenten hierop reageren. Essentieel is docentfeedback die studenten stimuleert tot reflectie en diepgaand leren (Higgins, Hartley & Skelton, 2002; Quinton & Smallbone, 2010). De onderzoeksvragen zijn: *Wat is de relatie tussen docentfeedback en studentreflecties? Wat zijn de effecten van het trainingsprogramma op het feedbackrepertoire van docenten en de uitingen van individuele studenten in feedbackdialogen?*

Feedback is in vijf categorieën uitgewerkt: een betekenisvolle, twee didactische en twee correctieve categorieën (Chi, 1996; Chi et al., 2001; Chin; 2006; Shute 2008). Betekenisvolle feedback betreft samenvatten, checken, focussen, diagnostiseren en vooruitkijken; bij didactische feedback gaat het om uitleg geven en voordoen; bij correctieve feedback staan positieve waardering en negatieve waardering centraal. Reflecteren veronderstelt cognitieve, affectieve en regulatieve activiteiten (zie bijvoorbeeld Mansvelder-Longayroux, Beijaard & Verloop, 2007; Oosterbaan, Van der Schaaf, Baartman & Stokking, 2010). We hebben deze activiteiten in acht denkactiviteiten uitgewerkt: beschrijven, analyseren, structureren, evalueren, attribueren, verklaren, concluderen en intenties formuleren (gebaseerd op studie 1). We hebben als studentcategorieën toegevoegd inhoudelijke informatie (samenvatten en toelichten) en inhoudelijke waardering (positieve en negatieve).

In de nameting, vergeleken met de voormeting, checkten docenten in de controlegroep gemiddeld genomen meer en gaven ze minder uitleg. Voor docenten in de trainingsgroep gold dat ze gemiddeld genomen zich vooral meer hebben gefocust en meer positieve waardering hebben gegeven.

Het focussen riep veruit de meeste denkactiviteiten op, vooral verklaren en concluderen. Ook bleek dat de volgordes twee keer betekenisvolle feedback, betekenisvolle feedback en correctieve feedback, en twee keer correctieve feedback, denkactiviteiten opriepen. In de nameting gaven docenten in de trainingsgroep minder uitleg en checkten ze minder en hun studenten lichtten minder toe. Deze resultaten kunnen erop wijzen dat docenten en studenten elkaar beter begrijpen en dat studenten meer autonomie hebben gekregen. Hoewel de verschillen tussen beide metingen in het gebruik van de categorieën door de

trainingsgroep klein zijn, is onze voorzichtige conclusie dat de training in het geven van feedback lijkt bij te dragen aan veranderingen in het feedbackrepertoire van docenten tijdens feedbackdialogen die in potentie de reflectie van studenten kunnen ondersteunen.

Terugkijkend naar de centrale onderzoeksvraag: *Hoe kunnen docentcompetenties voor het ondersteunen van de ontwikkeling van reflectievaardigheden van studenten Verpleegkunde worden ontwikkeld door middel van een trainingsprogramma?*, concluderen we dat het uitgevoerde programma een combinatie is van drie van de vier moderne trainingsmethodes die we hebben besproken. Ten eerste zijn instructie, oefenen en feedback typerend voor een training & toepassing aanpak. Ten tweede passen discussie over de huidige praktijk en reflectie op alternatieven in een ontwikkeling & reflectie aanpak. Ten derde wijst informatie geven over en relevante kennis gebruiken op een op deskundigheid gebaseerde aanpak. Sociaal leren heeft minder centraal gestaan, omdat de tijd die beschikbaar was voor discussie en uitwisseling tussen de docenten beperkt was.

Het programma had nog twee andere specifieke kenmerken. Ten eerste vonden de feedbackdialogen tussen de docenten en hun studenten plaats in authentieke situaties. In de bijeenkomsten met de docenten tijdens het programma zijn videoclips van authentieke feedbackdialogen gebruikt. Ten tweede maakte de trainer bij het coachen van de docenten gebruik van video-interactie begeleiding en dit kan hun reflectie en groei in bekwaamheid hebben gestimuleerd. Studenten lijken op hun beurt gevoelig te zijn geweest voor de redeneervragen en aanmoedigingen van hun docenten. Deze vergemakkelijken en bevorderen kennelijk de ontwikkeling van zelfregulatie van studenten.

Een omstandigheid die het leren en de ontwikkeling van docenten en studenten zal hebben beperkt is waarschijnlijk institutioneel van aard geweest: de permanent turbulente situatie in de scholen.

Dankwoord

Elk woord telt: Onder leiding van promotor Karel Stokking en dagelijks begeleider Marieke van der Schaaf ben ik dit spel van woorden aangegaan. Beiden hebben mij met een kritisch oog en oor ondersteund in het wetenschappelijk onderzoek. Het resultaat is in dit proefschrift neergelegd. Hartelijk dank voor jullie wijze raad.

Mijn promotie-onderzoek maakte deel uit van een overkoepelend onderzoek naar drie vormen van professionele ontwikkeling van docenten door drie universiteiten. Aan de Technische Universiteit Eindhoven onderzocht Migchiel van Diggelen de variant ‘collegiale beoordeling en zelfbeoordeling’ en aan Universiteit Leiden boog Christel Verberg zich over de ‘onderhandelingsaanpak’. Aan de Universiteit Utrecht onderzocht ik de ‘expert benadering’. In Leiden richtte Dineke Tigelaar zich, eerst met Mirjam Bakker en daarna met Patrick Sins, vanuit het interuniversitaire onderzoek op formatieve beoordeling als onderdeel van activiteiten voor professionele ontwikkeling van docenten.

De vijf studies van mijn onderzoeksproject brachten met zich mee dat ik met veel mensen heb samengewerkt. Graag wil ik hen ook hier bedanken voor hun bijdrage: inhoudelijke experts uit het bedrijfsleven en van hogescholen en universiteiten en managers, docenten, werkbegeleiders en studenten van de middelbare beroepsopleidingen Verpleegkunde van ROC ASA in Amersfoort en Utrecht, ROC Aventus in Apeldoorn en Zutphen, Hoornbeeck College in Amersfoort, ROC Midden Nederland in Amersfoort en Utrecht, ROC Mondriaan in Den Haag, Noorderpoort College in Groningen, ROC Rijn IJssel in Arnhem en Saffier in Den Haag.

Tot slot over non-verbale communicatie...

“Mam, je zei altijd heel weinig, máár je geloof in mij was er altijd! Mijn dank voor het voorbeeld dat je mij gaf met je grote doorzettingsvermogen.”

Geloven in iemand doet wonderen. Dat heb ik ook in mijn eigen omgeving telkens mogen ervaren. Een half woord is dan soms genoeg.

Een speciaal woord richt ik tot mijn man Kees en onze dochters Judith, Petra, Simone en Marijke. Vanaf het begin hebben jullie mij met veel enthousiasme ondersteund bij mijn promotie-onderzoek. Dank jullie wel voor het vertrouwen en voor de ruimte die jullie mij hebben geboden om mij verder te ontwikkelen.

Curriculum vitae

Agaath Dekker-Groen was born in 1954 in Heerhugowaard, the Netherlands. She studied Educational Science and Technology at Technical University of Twente. In 1985 she graduated in Instructional Technology as the first graduated of this study (with honours). Successively, she worked as educationalist at the Educational Centre of the University of Twente, at the Freudenthal Institute of Utrecht University, and at the Centre Educational Service in Rotterdam, each for five years.

In 2000 she obtained the Didactic qualification for Adult and Senior secondary vocational education and became teacher and educationalist for health at Regional Education Centre Midden Nederland in Utrecht. In 2005 she was certified by the Open University in Heerlen as a coach for managers. In 2004 she also became a project advisor and in 2007 a research member as secondment to the Lectureship Teaching in multicultural schools of Utrecht University of Applied Sciences.

In September 2007 she started as a PhD student at Utrecht University and joined the Research Centre Learning in Interaction. Her research focused on teachers in senior secondary vocational education who support nursing students' reflection skills development. In the meanwhile she followed ICO master classes and courses and a course in video interaction guidance.