Negotiating Common Ground
Tools for Multidisciplinary Teams

PJ Beers

O U N

ico

This research was carried out in the context of the
Interuniversity Centre for Educational Research

NWO - MES

Society and the Electronic Highway

ISBN-10: 90-9020050-9
ISBN-13: 978-90-9020050-7

Cover design by Jeroen Storm and PJ Beers

Printed and bound by Datawyse, Maastricht, The Netherlands

Negotiating Common Ground: Tools for Multidisciplinary Teams

Proefschrift

ter verkrijging van de graad van doctor
aan de Open Universiteit Nederland
op gezag van de rector magnificus
prof. dr. ir. F. Mulder
ten overstaan van een door het
College voor promoties ingestelde commissie
in het openbaar te verdedigen

op vrijdag 25 november 2005 te Heerlen
om 15:30 uur precies

door
Pieter Jelle Beers

geboren op 7 oktober 1977 te Deventer

Promotores
prof.dr. P.A. Kirschner, Open Universiteit Nederland
prof.dr. H.P.A. Boshuizen, Open Universiteit Nederland
prof.dr. W.H. Gijselaers, Universiteit Maastricht

Beoordelingscommissie
prof.dr. R. Bromme, Westfälische Wilhelms-Universität
prof.dr. M.S.R. Segers, Universiteit Leiden
dr.ir. M.B.A. van Asselt, Universiteit Maastricht
prof.dr. E.J.R. Koper, Open Universiteit Nederland
dr. J.M. van Bruggen, Open Universiteit Nederland

Preface

So what to say... "They were four interesting years?" Yes they were. Working as an environmental health scientist in the field of educational technology certainly was a novel experience. And moving from a pioneering, three-year old, highly ambitious institute as a trainee, to work in an established institute with an established culture and over one hundred employees was nothing short of a culture shock. But after the first effects wore off I found that I had quite quickly become a member of a close-knit group of PhD-students and junior researchers.

I'd like to take this opportunity to thank some people. First of all my PhD-supervisors. Paul Kirschner, for giving me the opportunity to conduct a very interesting research project. Paul always kept up the enthusiasm for the project, and never failed to urge me on when I required urging. Also, he provided me with excellent examples of how to do a good presentation. I hope I'll be able to live up to that standard. Els Boshuizen provided a special role, acting as my daily supervisor. I will not mention all roles Els has played during my PhD trajectory, suffice it to say that they were many and varied, and worthy of a nomination for Supervisor of the Year. Els knows how to make a weak story strong, slowly but steadily moulding the arguments and claims into a coherent tale. I found working with Els a pleasure, especially when we were "writing together" and "thinking together," and asking the questions I needed to get me further. Furthermore, I'm extremely grateful for her support and patience throughout the project. I'd like thank Wim Gijselaers for his sharp eye for plans that needed change. Wim never failed to slow me down when I needed to be slowed down. Also, he showed me that writing an introduction to a scientific article is an art in itself, of which I still have a lot to learn.

Of the people that helped me doing my research I must first thank Piet Van den Bossche, working on a twin project at Maastricht University, for the inspiring discussions. Together with Sanne Akkerman, Ingrid Mulder, Jakko van der Pol, and Piet, I formed a small network of PhD-students on the topic of shared understanding, which proved full of critical enthusiasm, and the sense of urgency that can go with science. Thanks to Jan-Willem Strijbos for pointing out the importance of methodological rigour. I would like to thank Wim van der Vegt for his excellent job programming NTool, and Jochem Westendorp and Aukje Schurer for their assistance during the study. And Marjolein van Asselt, once more for pushing me in the right direction, and for all the trust and inspiration.

On a personal note, there are some people that I want to single out here for what they have meant to me the last four years. Tamara, Judith, Frans, Iwan, Bas, Jan-Willem and Lieke, it has been a pleasure being your colleague, but it's far more important that you have become friends to me. Chris, Maarten, Frank, Rutger, Jorg, and Margriet, thanks for the good times, and for being there when I needed you. Which brings me to the last people I want to thank. But certainly not the least: Tsjikke, Kees and Dirk. Thank you all!

PJ Beers

Contents

Chapter 1

Introduction

Society is increasingly confronted with complexity. Societal fields like technology, population composition, and economics are in continuous change, and the rate of this change is increasing (Sterman, 1994). This complexity is reflected in policy problems. Complex societal problems have a number of unstructured (Hisschemöller & Hoppe, 1995-96), or 'wicked' (Conklin & Weil, 1997; Rittel & Webber, 1973) qualities. They often have no correct solutions; they are often intertwined with other problems; they appear across different disciplines; they are characterised by their many uncertainties; their definition depends on the observer's perspective; and solutions often cannot be 'tried out' in advance of making a final decision (Hisschemöller & Hoppe, 1995-96; Rotmans, Kemp, & Van Asselt, 2001; Van Asselt, 2000). While structured problems may be solved using standardised techniques and procedures, societal complexity obviously can not be treated in the same way. As noted by Sterman (1994) and Lomi, Larsen, and Ginsberg (1997), societal complexity requires novel ways of problem conceptualisation, using knowledge from different scientific disciplines and societal perspectives, and doing justice to the inherently uncertain structures of complex problems.

The many-sidedness of complex societal problems, that is, the many uncertainties and value systems involved, and the subsequent need to integrate knowledge from different perspectives, underlines the need for *multidisciplinary* and multi-stakeholder approaches. Solving complex problems is therefore typically done in teams whose team members have a variety of backgrounds and/or points of view, education, expertise, or political/social orientation. In such teams, the team members can bring multiple perspectives to bear on the problem, which allows for the rich problem conceptualisation required to solve complex problems (Lomi et al., 1997; Vennix, 1996). Solving complex problems then becomes a matter of *sharing* and *integrating* the knowledge, expertise and points of view from these different perspectives and using it to construct novel problem solutions. We call this the collaborative construction of knowledge. However, an approach that integrates multiple perspectives is not a guarantee for better problem solutions, only a basic requirement.

This thesis aims to add to the growing body of knowledge and techniques of supporting complex problem solving. It is traditional in the sense that it takes an aspect of the process of complex problem solving and then constructs a theoretical analysis of it. This analysis is then used as a starting point for the design of a decision support principle. It is novel in the sense that it is one of the first examples of taking aspects of problem solving groups as the starting point.

The research question of this thesis is: How can we support the collaborative construction of knowledge in multidisciplinary teams so as to help the teams to better solve complex problems?

Traditions in Support for Complex Problem Solving

It should come as no surprise that designers and researchers in the field of decision support have taken general aspects of complex problems as starting points for decision support, especially since policy makers have a tendency to treat policy problems as more structured than they actually are (Hisschemöller & Hoppe, 1995-96). Different traditions in decision support have chosen different starting points. Some strands of decision support start from notions about *problem structure*, while others start from specific notions about *argumentation structure*. For instance, *system-dynamics* can be seen as a way of thinking that does justice to the structure of complex systems; that is, seeing a complex system as sets of interlinked cause-effect relations and feedbacks that behave non-linearly (Forrester, 1971; Senge, 1990), and thinking about them in terms of stocks and flows (Rotmans et al., 2001). Applying system-dynamics to complex problems can give insight in the dynamic consequences of decisions. It has revealed explanations for unexpected, ineffective, and sometimes even detrimental effects of some public policy programmes by laying bare the counter-intuitive behaviours of social systems (Forrester, 1971). System dynamics has been adapted in a wide variety of decision support methodologies such as group model building (Vennix, 1996), transition management (Rotmans et al., 2001), and soft systems methodology (Checkland, 1981).

Many examples in the area of *argumentation structuring* can be traced back to the seminal work of Toulmin on the construction of argument (Toulmin, 1958). Such approaches aim at identifying the rationale of decisions, and make argumentation fallacies more salient. Argumentation support can be said to augment reasoning (Buckingham Shum & Hammond, 1994). It has been adapted in a wide variety of support methodologies such as SODA (Strategic Options Development and Analysis, Eden & Ackermann, 2001), IBIS (the Issue-Based Information System, Kunz & Rittel, 1970/1979), and gIBIS (graphical IBIS, Conklin & Begeman, 1987)

Both of these traditions presume that the team members who collaboratively try to solve complex problems are able to effectively share and integrate their own knowledge with the knowledge of others. However, research from the cognitive sciences shows that this assumption does not hold. For example, experts have been shown to generally overestimate others' knowledge of their own expertise, and consequently they communicate in ways that are hard to follow for laypersons (Bromme, Rambow, & Nückles, 2001).

If we conceive of the multidisciplinary team as a collection of different experts who share their knowledge with laypersons in the area of the knowledge-sharer, one can easily imagine that understanding each other is not self-evident. Indeed, empirical studies have underlined the fact that multidisciplinary teams have limited shared understanding (Alpay, Giboin, & Dieng, 1998). The above-mentioned traditions fail to *explicitly* recognise the difficulties that groups may have in constructing the knowledge needed for solving complex problems. It is one thing to reckon with different perspectives, but it is something quite different to take knowledge from different perspectives and allow a group of people to integrate that knowledge.

In this thesis, each team member's internalisation of knowledge put forth by the other team members in multidisciplinary complex problem solving teams is seen as a necessary requirement for decision making. Hence it is taken as a starting point for decision support.

Complex Problem Solving as Collaborative Knowledge Construction

Complex problem solving is here conceived of as a collaborative process that starts with unshared disciplinary knowledge from individuals, and ends with the construction of new knowledge, which takes the form of a problem solution. In-between, a number of processes acts upon knowledge, taking it from one form to another. For example, the process of knowledge externalisation (see Figure 1.1) can be said to act upon unshared knowledge, transforming it into externalised knowledge. In the introduction, these processes will further be referred to as 'knowledge processes'. The overall question then is what processes mediate the transformation of unshared knowledge to newly constructed knowledge, and what intermediary forms knowledge takes in-between.

Research and theory on social learning (e.g., Salomon & Perkins, 1998; Sullivan Palincsar, 1998), linguistics (e.g., Baker, Hansen, Joiner, & Traum, 1999; Clark & Schaefer, 1989), cognitive psychology (e.g., Boshuizen & (Tabachneck-)Schijf, 1998; Bromme, 2000) and social constructivism (e.g., Jonassen, 2000; Ostwald, 1996) all offer views on knowledge processes. Naturally, these different bodies of research (some would go as far as calling them different research paradigms) can not be used in conjunction without proper care. However, the sole purpose of exploring theoretical possibilities may justify such an eclectic lookout for the purposes of this introduction.

Ostwald (1996), a researcher of constructivist orientation, puts great store in the processes of representing knowledge and interpreting knowledge. Both processes are actions by individuals on *external representations*. According to Ostwald, a shared representation is an object for a group to think with. Such an object can be a diagram, a paper, a presentation, but also plain human speech, in which case the external representation is of a particularly fleeting nature. From Ostwald's work it thus appears that individual knowledge can be externalised, that is, transformed into an external representation. Furthermore, knowledge can be interpreted by others, who construct their own knowledge of the external representation. The external representation itself becomes a shared artefact of a group, representing shared knowledge.

Whereas Ostwald (1996) has particular interest in the *external* representation, Boshuizen and (Tabachnek-)Schijf (1998), researchers of cognitivist orientation, take *internal* representations as a starting point for deducing the difficulties in knowledge construction and knowledge sharing. They define a representation as a specific format for storing information, and the means to operate on and with the stored information. Boshuizen and (Tabachnek-)Schijf note that information can be stored in multiple representations, even in the same individual, for example by storing a number as either Roman or an Arabic numeral; the representations offer the same information, but have radically different options for using it. Starting from difficulties individuals have in reasoning between different internal representations, they deduce some conclusions about the case in which multiple representations are distributed across different people, arguing that dissimilar representations can cause difficulties in communication due to misunderstandings, especially when a task requires *integration* of representations, such as is the case for solving complex problems, while no measures for co-ordination of the communication are taken.

Bromme's (2000) theoretical account of the *negotiation of common ground* nicely fits the work of Boshuizen and (Tabachnek-)Schijf (1998). Bromme, also of cognitivist background, theorises about the causes of misunderstanding in the case of multiple representations. He notes that people communicate on the basis of assumptions they hold about the others' perspectives. For example, a mathematician explaining a mathematics problem will probably be more elaborate when talking to laypersons than when talking to peers, will use other terminology, will use different representations on a piece of paper (to a mathematician a formula, to a layperson a graph for example, et cetera). Thus, one's perspective of the other affects the way we externalise our knowledge, and also our understanding of others' contributions. Unfortunately, our assumptions of others' perspectives have been shown to be inaccurate (Bromme et al., 2001), which explains the misunderstandings mentioned by Boshuizen and (Tabachnek-)Schijf. Here, Bromme introduces the concept of negotiation of common ground, iteratively making one's private understanding of the other explicit and providing feedback so as to reach common ground, which is a common cognitive frame of reference. Common ground, once it has been achieved, can act as a shared interface between multiple representations.

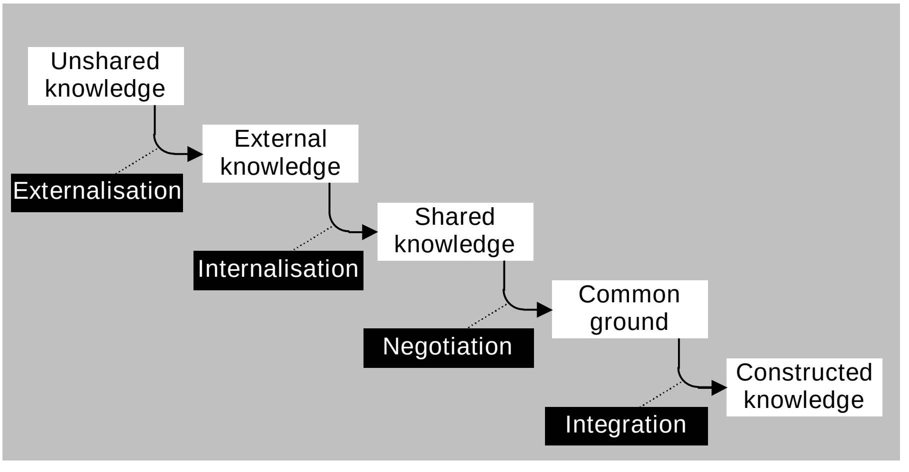

Figure 1.1: From unshared to constructed knowledge

Of all above processes, it seems that the negotiation of common ground is the one most involved with different perspectives in teams. In this thesis, our efforts have therefore concentrated on the support of negotiation processes.

The Studies in this Thesis

This thesis reports a series of experiments aimed at the iterative testing and development of NegotiationTool (NTool), an ICT-tool aimed at supporting negotiation of common ground. There are several ways to support knowledge processes. For instance, several systems for learner support make use of tutors (Problem-Based Learning, Barrows & Tamblyn, 1980) or facilitators (Selvin, 2002) to influence group processes. In Problem-Based Learning a tutor monitors group discussion and task content to make sure that students learn to co-operate well and that they construct knowledge with each other.

Other systems use ICT-tools to support groups. These tools are often grouped under de name of computer-supported collaborative learning (CSCL). CSCL-researchers use computers to achieve specific learning results. For instance, Suthers (2001) and Kanselaar et al. (2003) studied computer-support tools for argumentation. They used computers to require learners to be exceedingly explicit about their reasoning in order to increase the quality of their argumentation and in order to teach them how to reason in ways that are scientifically sound. (See Jonassen, 2000, for an interesting overview of such tools.)

Computer-support tools have a number of methodological advantages compared to facilitator-supported methods. The computer is not biased towards the participant. Furthermore, the computer keeps its behaviour constant whereas a human facilitator's behaviour can be affected by any of a number of factors (e.g., fitness, mood changes, etcetera). And finally, computer-support tools offer intrinsic storage of communication which is easy to access for the researcher. Our methodological considerations were informed by the field of CSCL, which is concerned with the development and testing of innovative learning environments. CSCL-researchers generally start with a collaborative learning problem they want to address, in our case the negotiation of common ground. Their next step is to use a theoretical framework as a basis for a *formalism*, a set of rules and guidelines for communication. Having done this, the CSCL-researcher is ready to enter the laboratory.

The actual development of a CSCL-environment (generally a computer program that can be used for synchronous and asynchronous communication), if done well will begin with simple face-to-face experiments to test the basic assumptions underlying the formalism and to gain insight in the ways the formalism may affect group collaboration (see Chapter 2, for our first face-to-face experiment). On the basis of these results, the next step is the adaptation and/or implementation of the formalism to an actual CSCL-environment, which then can be tested in the laboratory (see Chapter 3 for our first laboratory study with NTool). See Kirschner, Strijbos, Kreijns, and Beers (2004) for a discussion of this 'interaction design'.

In the most successful cases, that is, when all results are as expected, development can then be taken to the educational practical setting or the professional situation. In our case, results from the first laboratory study were somewhat mixed; our main hypotheses were confirmed (the stronger we applied the formalism, the greater its effects), but there were also some unexpected results that required further study (see Chapter 4), which took us back to a face-to-face setting.

Our final research effort took us to an educational setting. Although not the same as a professional setting, we considered this to be a lot more 'messy' than the laboratory, with regard to social processes, motivation, and educational level, and therefore a good opportunity for testing the robustness of our earlier findings.

This Thesis

The lion’s share of this thesis concerns reporting several experiments (Chapters 2 – 5) conducted with the aim to test and further develop the formalism, which has resulted in NTool, an ICT-tool for synchronous and asynchronous distributed communication with embedded support. Chapter 2 gives a broad view of the theory behind our support principle and reports on pilot-data to argue that the support principle can indeed influence negotiation processes. Chapter 3 builds on Chapter 2 by taking the decision-support from Chapter 2 and turning it in an ICT-tool, NTool. Chapter 3 reports on a laboratory study in which three different versions of NTool were tested. It turned out that the more stringently participants were made to act according to the formalism, the more common ground they negotiated. This is called *coercion*, using specific constraints on communication to keep participants from not using the support principle. However, we also encountered some unexpected results; one of the tested NTool versions resulted in a disproportionately large amount of regulation messages, suggesting something was wrong with the exact way coercion had been implemented.

In Chapter 4 the unexpected results from Chapter 3 are further examined in a second small-scale face-to-face study using additional measurement types so as to check for possible adverse effects of NTool. Chapter 5 reports on a second study with NTool. However this time the study was carried out in a secondary vocational education setting instead of the laboratory, to test how robust the findings obtained in the laboratory were. Chapter 6 critically reviews aspects of our research methodology, against the backdrop of the lack of methodological tradition in the field of computer-supported collaborative learning and working.

This thesis is not a book in the traditional sense but a collection of highly related articles. This is to warn the reader that in some cases s/he will come across some text or a figure s/he already has seen. Finally, throughout this thesis the word “we” is used instead of “I”. The reason for this is that research is not a solitary activity. “We” thus represents “the researchers” in this project.

Chapter 2

Common Ground, Complex Problems and Decision Making

Organisations increasingly have to deal with increasingly complex problems. They often make use of multidisciplinary teams to cope with such problems where different team members have different perspectives on the problem, different individual knowledge and skills, and different approaches to how to solve the problem. In order to solve those problems, team members have to share their existing knowledge and construct new knowledge. Theory suggests that negotiation of common ground can positively affect team decision making on the solution of complex problems by facilitating knowledge sharing across perspectives. In a small scale study, external representations supported by a specific negotiation ontology were used to facilitate negotiation by forcing participants to make their beliefs and values explicit. Results showed that the external representations supported the clarification of contributions made and increased group participation in discussions.

Different kinds of organisations—private business, public, and non-governmental alike—are confronted with increasing complexity in decision-making situations (e.g., Courtney, 2001; Rotmans, 1998; Sterman, 1994). Two examples of such complex, often societal, problems are:

- Whether to adopt and, by a positive decision, where to plan the trajectory of a high speed train. This problem is complicated by factors such as possible environmental damage, the necessity expropriation of private property, expected economic benefits and costs, effects of such a trajectory on roadways and truck transportation of goods, et cetera.
- How to reduce the number of high school drop-outs. This problem is complicated because it can be seen as a motivational problem, a pedagogical problem, and also as a legal problem, as well as the question of who needs to take responsibility (e.g., parents, city, state and national governments).

Research has shown that solving such problems is not easy. DeTombe (2002) notes that problem solvers tend to spend too little time on problem analysis, leaving them either unaware of complexity of the problem, or handling the wrong problem. Beers, Van Asselt, Vermunt, and Kirschner (2003) found that policy makers working on global sustainability issues had both information needs regarding related problems and regarding different cultural perspectives on the problem. Hasan and Gould (2001) found that neglecting certain problem perspectives in developing a university research performance indicator had undesired effects in terms of research output.

The many-sidedness of complex societal problems underlines the need for rich problem conceptualisations, and thus the need for multidisciplinary and multi-stakeholder approaches. However, forming heterogeneous teams to solve complex problems is no guarantee for a good solution since, especially in the case where teams are formed from different perspectives, team members can have difficulties in understanding each other and sharing knowledge. Bromme et al. (2001) found that people tend to make biased estimates about the knowledge of their discussion partners, which may result in ample explanations of what is widely known, or ignorance of misunderstanding. Bechky (2003) found that problems on the work floor caused misunderstandings between workers from different departments due to their different perspectives. Interestingly, she noted that in order to solve the work floor problems, different workers would first achieve some common ground (i.e., a common frame of reference) to bridge the differences in perspective and be able to share knowledge from their different viewpoints.

Much effort has been expended in the development of decision support-tools for the facilitation of complex problem solving. These tools generally aim at facilitating formal and informal communication, harvesting knowledge, and building knowledge repositories (Courtney, 2001). However, they do not facilitate bridging the gaps between different perspectives. As Hasan and Gould (2001) lament: “There has been an unfortunate tendency to view both the computerised information systems and the decision makers as comparable information processors" (p. 71). A different decision-support approach is needed to meet the requirements of facilitating complex problem solving.

This chapter approaches decision support as a group-processes phenomenon. It puts forward a framework for designing decision support for complex problem solving in multidisciplinary teams, and proposes a methodology for testing these designs. First some aspects of complexity and multidisciplinarity are dealt with. Then the framework is described, with an emphasis on theory about common ground (Bromme, 2000). It is argued that achieving common ground can be afforded by making individual team members' perspectives explicit to the others.

From this framework a set of *primitives*, basic building blocks (Dillenbourg, 2002) for the design of decision support, is derived. The overarching question discussed in this chapter is how to design tools for the externalisation of individual perspectives, and how to measure their effects.

The second part of this chapter reports on a small scale study aimed at making the process of grounding visible and whereby an attempt was made to measure common ground in six multidisciplinary groups.

Complexity

Complexity is related to the intricacy of systems. The more factors and relations within a system, and the more element interactivity between them, the more complex a system is (cf. Evans & Marciniak, 1987). In decision making, complex societal problems often cross disciplinary boundaries and involve many different stakeholders (Rotmans, 1998). For example, the problem of school drop-out can be conceptualised as pedagogical, social, economic, and legal in nature, just to name a few. Furthermore, the interests of a variety of actors are at stake, namely children, parents, teachers, politicians, civil servants, etcetera.

Such problems are not only complex, their nature also is inherently "wicked" (Rittel & Webber, 1973), which poses some difficulty to problem-solving. In such problems "solutions" usually cannot be tried out, so aspects like plausibility and acceptability of information play an important role in problem solving. Furthermore, one "solution" is often either incomplete, and/or unacceptable since it does not or cannot take into account the effects on, and the problems of all other stakeholders.

Lomi et al. (1997) argue that mono-disciplinary approaches to solving such problems generally lead to adaptive or incremental solutions, and, more importantly, fail to generate those innovative solutions necessary for coping with societal complexity. In sum, to deal with societal complexity, it is important to involve different perspectives on the problem, so as to prevent solving the wrong problem, or not solving the problem at all.

Multidisciplinarity

Team members, based on their professional and personal background, will each have their own *perspective*, their own way in which they coherently and consistently "interpret and make sense of the world" (cf. Douglas & Wildavsky, 1982; Van Asselt, 2000, p. 115). In multidisciplinary and multi-stakeholder teams, the members can bring their different *perspectives* to bear on the problem, resulting in *multiple representations* of the problem within one team. These different representations, when properly co-ordinated, can cater for a broad problem scope which assists teams in going beyond adaptive and incremental solutions (cf. Dillenbourg, Baker, Blaye, & O'Malley, 1995; Vennix, 1996). Or, as Vennix (1996) puts it, "the more different perspectives are taken into account, the smaller the chances of premature problem definition and 'solving the wrong problem'" (p. 1). However, different perspectives have also been shown to give rise to misunderstandings and disagreement among decision making-team members (e.g. Alpay et al., 1998; Boshuizen & (Tabachneck-)Schijf, 1998; Bromme & Nückles, 1998), threatening the decision-making process itself.

Members of multidisciplinary teams make use of language either not understood by those with other areas of expertise, or with a different meaning in another's field (Van Someren, Reimann, Boshuizen, & De Jong, 1998). People can make contributions which are evident from their own point of view, but not from the perspectives of other group members. This, in turn, can lead to either non-understanding or misunderstanding. If misunderstanding remains undetected and a team continues to work, multiple representations may come to equal 'multiple ignorances'. The team may not be able to construct a coherent problem representation at all, or only solve partial problems. It even runs the risk of unknowingly implementing contradictory solutions to the same problem.

Bromme's (2000) theory of cognitive interdisciplinarity holds that *common ground* is an important condition for knowledge sharing. We argue that multidisciplinary teams need the common ground required for sharing knowledge among different perspectives.

A Framework

The framework, developed by the researchers, attempts to bring complex problem solving and multidisciplinary teamwork together by focusing on the various stages which take knowledge from being something implicit in the mind of one person to becoming a team's explicit constructed knowledge.

The route from unshared knowledge in one participant's head to newly constructed knowledge in a team (see Figure 2.1) goes through three intermediate forms (i.e., external knowledge, shared knowledge, and common ground) via four processes (i.e., externalisation, internalisation, negotiation and integration).

Private knowledge is *externalised* when a team member makes his/her own, as yet unshared knowledge, explicit or tangible to others (Leontjev, 1981). This can be oral, written, symbolic, etcetera. Once a team member has made a contribution, the others can all try to *internalise* this contribution. They can consider aspects of the contributor such as background, current situation and possible views so as to better "understand" the contribution. Also, their own beliefs and assumptions play a role while they try to understand the contribution. A contribution is thus understood against the *presumed perspective of the other*, as well as against one's *own perspective* (Bromme, 2000).

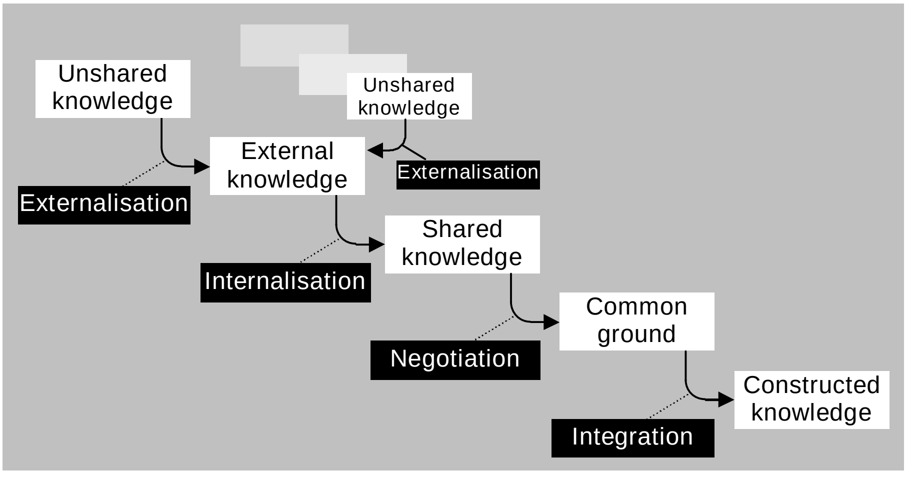

Figure 2.1. From unshared knowledge to constructed knowledge

Externalising and subsequently internalising each others' knowledge does not mean that the team members all have arrived at the same understanding. Representational differences result from interpreting a contribution in one's own perspective only (a graphical designer has a different understanding of, and use for the term "elegance" than a computer programmer, see below) or from minimising or rejecting its validity or plausibility due to differences in conviction or opinion (an environmentalist may reject the corporate utterance on principle). Negotiation has to take place in order for a contribution to be accepted and agreed upon by the whole team (e.g. Alpay et al., 1998; Bromme, 2000; Dillenbourg et al., 1995).

As an example, imagine the following. A computer programmer and a graphical user-interface designer are two members of a team designing a new user interface. At a certain point the programmer states that the chosen solution needs to be "elegant". The designer readily agrees, but to what? The programmer meant that the program needs to be as short as possible (her definition of elegance); the designer understood aesthetically pleasing (his definition of elegance). There is a problem here, but as yet it remains implicit and unseen, and no one can tell when it will surface, and how much damage it may cause in the interim. In other words, people may think they are on common ground while at the same time maintaining important representational differences. Such differences need to be *detected* before people can effectively start negotiating a shared representation to come to a solution.

Negotiation of common ground is conceived here as a dual concept. First there is *negotiation of meaning*, which, in the case of knowledge construction, leads to an agreement regarding meaning and understanding of a contribution. Negotiation of meaning concerns people making public to others their private understanding of some contribution, verifying whether and to what extent their own understanding is different from what others intended them to understand, receiving feedback on this, re-verifying, and so on. Negotiation of meaning is thus an iterative process that takes place until "the contributor and the partners mutually believe that the partners have understood what the contributor meant to a criterion sufficient for the current purpose" (Clark & Schaefer, 1989, p. 262, the grounding criterion).

Negotiation of position is the second part of negotiation. This concerns people making their private opinion about some contribution public to others, checking whether their own position is clear to others, and vice versa. It is through the process of internalising others' contributions, and subsequently providing feedback by word or action about those contributions based on one's own perspective, that common ground can be negotiated (Alpay et al., 1998; Baker et al., 1999). Common ground is never absolute or complete, but is an interactive and ongoing process in which assumed mutual beliefs and mutual knowledge are accumulated and updated (Clark & Brennan, 1991).

Starting from the common ground, new knowledge can be built by adding new relations and concepts to the common ground, via *integration*. Knowledge construction is based on the common ground the team has built, and will broaden and deepen the common ground because the common constructed knowledge becomes part of the common ground.

Primitives of Negotiation

Primitives can be seen as types of communication acts that are derived from a specific model of dialogue (Dillenbourg, 2002). Primitives of negotiation can thus be seen as deriving from a dialogue model of negotiation, in which each primitive serves as a basic building block of negotiation. To make this clear an exemplary account of negotiation is presented and then used for identifying primitives of negotiation.

Negotiation can start when someone has made a contribution (externalisation), and another has tried to pick it up (internalisation). When this has been done, a set of checks can be performed as to whether the discussion partners understand each other. The receiver compares what s/he thinks the contributor intended to say with his/her own understanding of the contribution, verifies this understanding (often with a question or rebuttal), and upon receiving feedback decides if there are no important differences between his/her own understanding and the contribution. The next step is deciding whether there is agreement. The receiver of a contribution can either agree with the contributor or hold an opposing opinion. Negotiation continues until all negotiators think they understand each other sufficiently, and either hold the same opinion or can sufficiently respect an opposing opinion. Note here that this does not mean that they assume the other's position. They *respect* it and can *agree to disagree*.

The above account of negotiation can be used to derive some 'basic building blocks' or 'primitives' of negotiation. Negotiation starts with a *contribution* of some sort such as a hypothesis or a position. A contribution does not stand on its own; rather it is based upon ideas and background of the contributor. Every contribution can therefore be underpinned by some sort of *clarification* by the contributor, which sheds light on the meaning of the contribution, or the opinion of the contributor. This clarification can remain implicit, for example, when the known background of the contributor sheds light on his/her contribution, but it can also be made explicit. Clarifications are needed because contributions are often not understood by the others in the way the contributor intended them to be. In other words, the meaning behind the message is not clear (Fischer, Nakakoji, & Ostwald, 1995).

Third, *verification* is needed for contributions to check the understanding of a contribution because people articulate and understand the contribution against their own background knowledge (Fischer et al., 1995). Verification and clarification are thus alike in nature, but differ with regard to the originator. A clarification is performed by the contributor, whereas the verification is performed by the one who is trying to understand the contribution.

A fourth element is *acceptance/rejection* of a contribution. This refers to whether one can judge a contribution as true (acceptance), based on the explanation given, or as untrue or unintelligible (rejection). The fifth and last primitive requires every negotiator to decide upon a *position* regarding the contribution. This includes the possibility of accepting a certain contribution, but disagreeing all the same, for example when neither person can prove the other wrong. In such cases, people can agree to disagree, and alternate representations that are equally legitimate can ensue.

External Representations for Supporting Negotiation

Representations of abstract concepts exist in many cases partly in our heads, in which case they are called internal representations, and/or partly in our environment and are called external representations (Alpay et al., 1998). People can externalise their internal representations to a certain extent, and capture them externally. External representations can take many forms. The most common are self-made, idiosyncratic notes, outlines, diagrams, flow charts and even mind maps, but there are also standardised ones such as mathematical and scientific notations.

Making an external representation for a group process requires both a *carrier* of the external representation, such as pen and paper or a software tool, and a *formalism*, such as a common language or representational technique, to guide knowledge externalisation (Suthers, 2001). This “formalism” is thus a set of objects and rules that guides making an external representation.

Formalisms can be tailored to specific activities. Suthers (2001) for example developed a formalism to enhance scientific discourse, which he implemented in a software tool called Belvédère®. This formalism consists of a small set of rules which (1) require evidence for every statement made by the group, (2) require opposing evidence for each group statement, and (3) prompt users to check whether this evidence supports statements other than the one it was given for. This formalism helps teams to distinguish between strong and weak statements, and to articulate uncertainties.

A formalism can influence both knowledge externalisation and knowledge internalisation (Figure 2.2). It can streamline problem solving processes if it limits problem space without excluding information and relevant knowledge. An associated risk is using the wrong formalism. Consider the problem of using the Belvédère® formalism, with its empirical scientific discourse-orientation, in fields such as law or design (cf. Verschuren, 1997). Using such a formalism can be a hindrance at best, and counterproductive at worst (Van Bruggen, 2003). In such cases the ontology of the formalism does not fit, or even conflicts with the ontology of the problem domain. A formalism needs to fit the activity it attempts to facilitate.

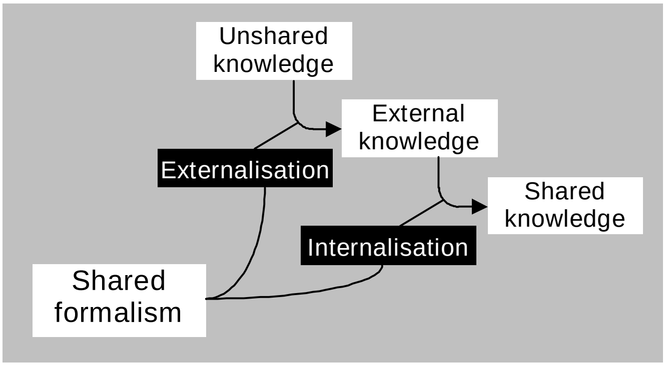

Figure 2.2. Formalisms focus knowledge externalisation and knowledge internalisation

A Formalism Based on Negotiation Primitives

Negotiation, as previously stated, can be broken down into a small number of primitives. The primitives can be seen as objects that result from actions by team members, and a formalism that aims to facilitate negotiation has to support these actions. In regular communication, the status of people's statements in terms of negotiation primitives remains implicit. Using a formalism designed to facilitate negotiation has to enable distinguishing between original contributions, clarifications, verifications, etcetera, and make the steps in negotiation explicit. By doing so, individual differences in understanding and opinion should more easily surface.

Having defined the primitives, the set of rules for negotiation are specified as follows: Every new issue added to a conversation is a *contribution* (rule 1), and is assumed not to be part of a team's common ground. To assist in the detection of differences between individual representations, every team member must *verify* to see whether their understanding of the contribution sufficiently matches what the contributor intended. The original contributor then has to add explicit *clarification*. Rule 2 is that all *contributions* have to be followed by *verifications* by the other team members, and Rule 3 is that all verifications require a *clarification*. Rule 2 and 3 can be iterated until common understanding of the contribution is reached.

The fourth rule is about *accepting* or *rejecting* a statement, based on one's judgement of whether it is right. The statement 1 + 1 = 10, for example, is true only if we understand (through rules 1 and 2) that the contributor is using the binary system. A contribution should be accepted as part of the common ground if it is true, or after it has been modified so that it has become true.

The fifth rule is used to add a value judgement to the contribution. People are required to explicitly state their own *position* regarding the statement. This allows for clarification / determination of perspective, and this in turn aids in the verification and clarification of further contributions. In the case of irresolvable disagreement about previously accepted statements, the fifth rule may result in multiple scenarios, each based on another value judgement (i.e., agree to disagree). Table 2.1 summarises these rules.

Table 2.1
Rules for a Formalism for the Facilitation of Negotiation

1.	Every new issue is termed a *contribution*
2.	Contributions require a *verification* by the other team members
3.	Each verification is responded to with *clarification* by the original contributor
4.	When all verifications are clarified, and no new verifications are performed, all team members state whether they *accept* or *reject* the statement
5.	All team members state their *position* about accepted statements

Using this formalism is expected to induce more negotiation of meaning and negotiation of position, because forcing people to make their private understandings and opinions public will make differences in understanding and opinion visible or salient (Bromme, 2000). This will be reflected in negotiation by the number of verifications and clarifications for every original contribution. By strengthening the negotiation process, we thus expect this formalism to increase the amount of common ground.

Study

A pen-and-paper version of the formalism for supporting negotiation was tested in which team members were required to both state their position concerning others' contributions in a face-to-face setting, and to explicitly verify and clarify their contributions on a flip-over against idiosyncratic representation and negotiation. The study was as much a test of the formalism as it was a test of analysis methods to determine

1. whether the formalism influences negotiation of common ground;
2. how participants experience negotiation and achieving common ground; and
3. how participants used the formalism.

The first aspect aims at determining the effects of the formalism on negotiation processes and common ground. The formalism was expected to make negotiation more explicit. Furthermore, groups using the formalism were expected to negotiate more thoroughly about each conversation topic than groups without the formalism. Finally, the formalism groups were expected to establish more common ground than the non-formalism groups, measured as overlap between individual representations, due to the increased negotiation.

The second aspect focussed on the validity of the research setting. Interview data were used to draw in-depth impressions of two topics in particular, namely participants' *thoughts about multiple perspectives and negotiation*, and their *awareness of grounding and perspectives* during collaboration. The data were used to test whether the present research setting corresponded sufficiently with a complex problem solving situation as perceived by the participants to offer leads for the formalism to be of use. Furthermore, it was intended to gain insight into whether participants were aware of the presence of multiple perspectives and the need for actively achieving and maintaining common ground.

The third aspect focussed on how participants actually used the formalism, and what they used it for.

Method

Six multidisciplinary groups (triads) of senior college students who were given a problem-solving task were studied. They were instructed to collaboratively carry out this task. Half of the groups were instructed to use the formalism for working with a white- or blackboard and flip-over to structure their collaboration (the formalism condition). The other groups could use these materials any way they wanted to (the idiosyncratic condition). Afterwards, all participants were interviewed.

Participants

Participants were students in their senior year from a business degree program of a Dutch University. Participants majored in such diverse fields as Accountancy, Marketing, Organisation Science, and Macro-economics. Six multidisciplinary teams were formed by dividing participants who majored in different fields into groups of three. These participants were assumed to have different perspectives due to educational differences and socialisation effects from their educational careers.

Materials

Task. A task was assigned to the groups, requiring them to arrive at an investment decision for a machine-producing company called “Thyssen & Krupp”. The assignment was derived from the (computer) simulation game “STEER The Economy” (Woltjer, 2003). Participants were provided with a large amount of company and market data covering such diverse fields as the company itself and its past decisions, competing companies, within-market developments, and overall macro-economic indicators. The abundance and variety of available data enabled complementary approaches from multiple economic perspectives. The participants received the following task description:

> You are employed as an organisation consultant to analyse and advise organisations. Thyssen & Krupp, a company which produces machines, contracts you to advise them on an investment decision. They require an analysis of their current situation and an investment strategy for the near future (number and type of machines ordered). All data regarding Thyssen & Krupp are available on your computer.

Formalism. Participants were supplied with a blackboard or whiteboard (henceforth ‘board’) and flip-over, and writing materials of different colours. Groups that used the formalisms received instruction on use of the board and flip-over. They were to be used as much as possible during discussion for writing down new topics for discussion, and for sharing opinions about those topics. Each participant used a specific colour for writing his/her contributions.

The flip-over was intended for *clarification* of contributions. Participants were to represent their own understanding of others' contributions on the flip-over. The original contributor could subsequently, also on the flip-over, represent what s/he really meant. This means that participants were not meant to discuss each other's opinions on the flip-over.

Once sure of understanding a contribution, participants could represent this on the board by initialling the original contribution. Finally, participants could represent their *opinions* about each other's ideas on the board, and also use the board to represent new ideas from the discussion.

In summary, the board was intended for recording ideas and opinions and the flip-over was intended to clarify the understanding of those ideas. Participants were instructed to begin with a short brainstorm on the board, and then continue on the flip-over with clarifications. After completing those two steps, the course of discussion was free. The groups that did not use the formalism could use the board and flip-over to their own discretion. These groups are referred to as the idiosyncratic groups.

Interview guideline. Participants were interviewed to gain insight in their thoughts about perspectives and negotiation, as well as their awareness of grounding and perspectives during collaboration. The interviews were conducted by three graduate assistants who had received training in basic conversational and interview techniques as part of their university education. Each interviewer conducted six individual interviews of a participant. The interview focussed on how the participants arrived at new ideas; how new ideas were introduced and exchanged during collaboration; whether ideas "landed" by the other team members; whether team members understood each other; and whether team members agreed with each other.

Procedure

Participants were given 45 minutes to explore the simulation and browse through all the different types of graphs and charts to get some experience with the information available in the program.

After this exploration, participants started working on the case. In order to promote the construction of an individual perspective, as well as to allow the researchers to determine participants' individual representations, participants first carried out the task individually (pre-test) and wrote down their solutions.

Next, participants had to collaboratively carry out the task in triads. They received a board and a flip-over to take notes. Half of the groups received the formalism instruction for using the board and flip-over; the other half could use them in an idiosyncratic way. All groups were instructed to use the flip-over for writing down their final solution to the problem task. Participants were allowed to bring their notes. The collaboration process was video-taped.

After the collaboration, participants were again asked to individually carry out the task (post-test).

The interviews were conducted as soon as possible after collaboration (always within 24 hours). All interviews were tape recorded and typed out by the interviewer who conducted them. Video recordings of the collaboration process were used during the interview to stimulate the participant's recall of their thoughts during collaboration. The participants watched the video-taped collaboration process and were instructed to report their thoughts during collaboration. They could stop the recording at any time. The interviewers also could stop the tape during the interview and were instructed to do so when the participant was silent for an extended period, had reacted fiercely to another on the tape, and vice versa, had spent an extended period searching for information or was neglected by the others. Figure 2.3 schematically represents the procedure for one triad.

Variables and Analysis

Negotiation, common ground, and participants' thoughts about negotiation of common ground were analysed. Negotiation was operationalised as *quality of negotiation*; *negotiation per conversation topic*; and *participation per conversation topic* during the collaboration. Common ground was measured by comparing individual representations before and after collaboration with respect to solutions and their justification. Thoughts about negotiation were measured by qualitative analysis of the interview data.

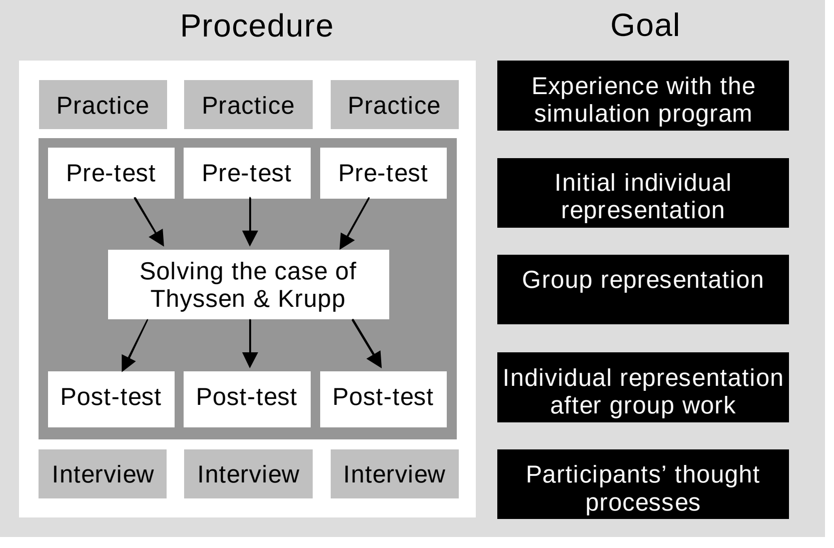

Figure 2.3. Experimental design

Quality of Negotiation

We developed a coding scheme for coding function and content of utterances during collaboration (cf., e.g. Fischer, Bruhn, Gräsel, & Mandl, 2002; Mulder, Swaak, & Kessels, 2002; Thomas, Bull, & Roger, 1982). All utterances were coded with regard to:

- Cognitive content: Utterances directly related to solving the problem.
- Regulative content: Utterances related to monitoring the problem solving process, and regulating the collaboration process. Talk about the formalism instruction was coded Regulation as well.
- Other content: Any utterance not in another category, or non-codeable.

Utterances with cognitive content were specifically coded for function, using subcategories based on the primitives, namely:

- Contribution: An utterance in which a new topic of conversation is introduced, that has not been discussed before.
- Elaboration: An utterance in which a contribution is expanded upon by either adding more information, agreeing, disagreeing, accepting, rejecting or summarising.
- Verification: An utterance in which, directly or indirectly, information is requested about the intended meaning of a contribution or elaboration.
- Clarification: An utterance in reaction to a verification or a perceived lack of understanding, in which the intended meaning of a contribution or elaboration is elucidated.

The coding scheme originally included separate categories for acceptance, rejection, agreement, and disagreement. However, very strict coding rules were needed to make these categories reliable, and with these rules the number of instances of the respective utterances was too small to be incorporated in the analyses.

Segmentation was done on the basis of utterances. However, acknowledgements (e.g., “Hmhmm,” “yeah, yeah,” etcetera) were ignored. Furthermore, if a speaker obviously changed the type of utterance in mid-speech, it was split into two segments. Segmentation and coding were done simultaneously.

A graduate assistant was trained for 25 hours on the use of the coding scheme and the video-coding software package The Observer® (Noldus, Trienes, Hendriksen, Jansen, & Jansen, 2000). Comparable data from an experiment with second-year students were used for training purposes. Comparing a sample of 25 minutes of video-data (9 % of the total amount of video-data) coded by the author and the graduate assistant resulted in a substantial (Landis & Koch, 1977) inter-rater reliability (Cohen’s kappa) of .68 (*SE* = .066). All data were coded by the graduate assistant, blind to which trials corresponded to which condition.

Verification and clarification were seen as indicative for explicit negotiation activities. The total number of contributions discussed was used as an indicator for the range of topics discussed. It was assumed that the wider the range of discussed topics, the better different perspectives were represented.

Negotiation per Conversation Topic

To measure the number of verifications and clarifications per conversation topic first those episodes in the discussion that dealt with one conversation topic were identified. The contributions that were identified earlier using the coding scheme served as a means to identify these episodes. A discussion episode generally started with a contribution, and ended when one of the participants would make a new contribution, and all of the discussion in between these contributions would deal with one conversation topic.

Note that in specific cases the discussion about one topic was interspersed with discussion of another, for example when the focus of discussion would shift between topics. In such cases all discussion about one topic would be added to one episode, even when it was interspersed with another episode. Negotiation per conversation topic was operationalised as the total number of verifications and clarifications per episode.

Participation per Conversation Topic

The participation per conversation topic was operationalised as the number of participants that made an utterance or an acknowledgement during an episode. Effectively this means that participation can be 1 (in case of a monologue to which no one reacts), 2, or 3 for each conversation topic, depending on the number of speakers during the corresponding episode.

Common Ground

Common ground was conceptualised as the extent to which individual representations overlapped each other with regard to content. To characterise the content of the individual representations the discussion content itself was first characterised (see Figure 2.4). The discussion episodes identified earlier served as a basis for characterising the discussion content.

Each episode was first numbered and summarised. For example, one of the episodes was summarised as follows: “The current low stock of Thyssen & Krupp is only a random indication because stock is fluctuating within the nine-month time of delivery for new machines.”

The next step involved characterising the content of the individual representations, both prior to (pre-test) and subsequent to collaboration (post-test), as well as characterising the content of the group representation. For every individual representation the topics represented and not represented were assessed. In Figure 2.4, episode number 7 is judged to be present in one of the initial individual representations (Jane’s) in the group representation, and in all of the post-tests. By repeating this procedure for each of the episodes in the discussion, it was determined where each conversation originated, whether it was present in the group representation, and whether participants used it in their post-tests.

Three different measures of common ground were used, based on different comparisons of individual and/or group representations. The first measure concerned the overlap of individual representations subsequent to collaboration (overlap after collaboration). The second measure concerned the difference between overlaps of pre-tests as compared to post-tests (change in overlap). The third measure constituted a comparison of the overlap of the post-tests to the group representation.

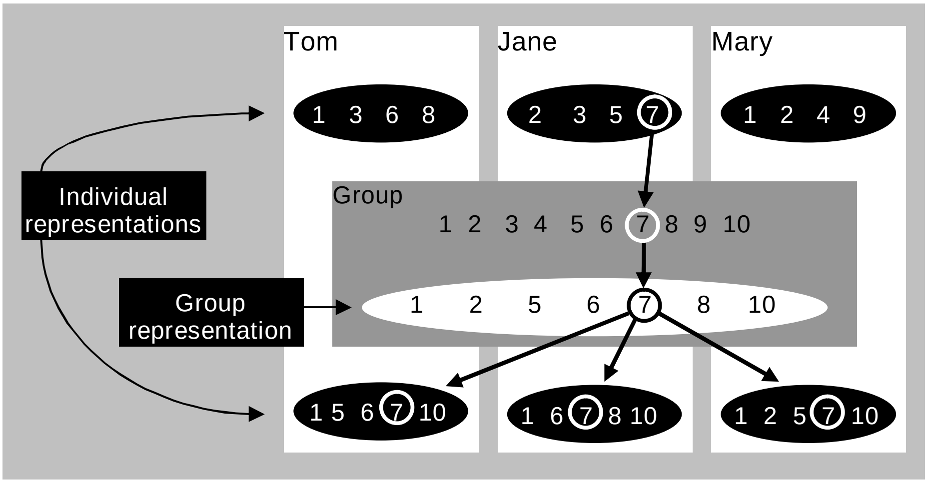

Figure 2.4. Analysis of common ground; numbers indicate topics

Qualitative Analysis of the Interview Data

The interviews were qualitatively analysed to gain insight in the participants' thoughts about multiple perspectives and negotiation, and their awareness of both grounding and multiple perspectives during collaboration. These topics are important to help judge the *validity of the research setting*, that is, whether there was enough initial misunderstanding between participants to sufficiently allow for negotiation processes, and whether the participants were able to learn from each other through negotiation (i.e., that different perspectives added value to solving the problem). It is also important to determine whether the participants were aware that different perspectives apply to the task, and that there is a need for achieving and maintaining common ground. The qualitative analysis was also used to gain insight in participants' use of the formalism, by focusing on the way the participants used the board and flip-over.

To answer the second and third research questions, all reported thoughts from the interview transcripts about any of the aforementioned aspects of negotiation were gathered and categorised. To assess differences between formalism and idiosyncratic groups it was determined whether some thoughts occurred only in the formalism groups and not in the idiosyncratic groups for each topic, and vice versa.

The qualitative analysis was carried out by the author.

Results

Quality of Negotiation

Table 2.2 shows the mean number of occurrences for each of the primitives for both the formalism and the idiosyncratic groups. Formalism groups worked longer than idiosyncratic groups. During that time, they discussed more contributions, and each of these contributions was negotiated more thoroughly, as shown by the amount of negotiation of meaning per contribution, than in the idiosyncratic groups. In general, the number of verifications and clarifications was higher in the formalism groups than in the idiosyncratic groups. Mann-Whitney testing showed that the difference in number of clarifications was marginally significant, $U(N = 6) = .500$, $p = .072$. Furthermore, in the formalism groups the mean number of participants per conversation episode was significantly higher than in the idiosyncratic groups, $\chi^2(2, N = 150) = 8.77$, $p < .05$. No other differences were found to be statistically significant, although all of the observed differences were sizeable and in the expected direction only. Seeing as how the sample was small and the research was explorative in nature, this was not unexpected. On the whole, the data suggest that the formalism groups negotiated more and more thoroughly than the idiosyncratic groups.

Table 2.2
Negotiation Primitives

	Formalism groups		Idiosyncratic groups	
	M	*SD*	*M*	*SD*
Time (s)	3181	394	2341	579
Contribution	27.0	7.8	23.0	4.4
Verification	17.0	5.2	10.7	3.5
Clarification#	23.3	7.5	14.7	1.2
Negotiation of meaning	40.3	12.2	25.3	2.5
Elaboration	197.3	57.6	151.7	18.8
Regulation	30.0	6.6	24.0	13.1
Other	43.0	14.2	32.7	21.2
Negotiation[a] per contribution	1.7	1.0	1.1	0.1
Participants per episode*	2.8	0.22	2.5	0.09

[a]Negotiation is meant here as negotiation of meaning, and consists of the sum of all verifications and clarifications.
#$p < .10$; *$p < .05$

Common Ground

Table 2.3 reveals that the participants in the formalism group discussed more topics than the participants in the idiosyncratic group. Also, the members of the formalism groups mentioned more different discussion topics in their individual problem representations after the problem-solving task. However, the idiosyncratic groups captured more discussion topics on their group external representation than the formalism groups ($M = 13.0$ vs. $M = 10.7$) as can be seen in Table 2.3. Furthermore, the number of discussion topics mentioned in the post-tests (overlap after collaboration) by all members is the same for both conditions ($M = 2.0$, number of discussion topics in three individual representations), which means that no difference in common ground was found. Adding pre-tests to the measurement of common ground (measuring change in overlap) suggests a bit more common ground in the idiosyncratic groups.

Table 2.3
Common Ground

	Formalism groups		Idiosyncratic groups	
	M	*SD*	*M*	*SD*
Total number of episodes	27.0	7.8	23.0	4.4
In one pre-test	12.0	3.6	11.7	2.9
In two pre-tests	2.7	1.5	2.7	2.9
In three pre-tests	1.0	1.0	.3	.6
External representation	10.7	2.3	13.0	7.8
Solution	6.3	5.1	4.0	1.7
In one post-test	7.3	3.1	6.3	1.2
In two post-tests	4.3	2.1	4.7	1.5
In three post-tests	2.0	1.7	2.0	1.7

When comparing the overlap between post-tests to what was in the external representation, the idiosyncratic groups, although they captured more topics on their external representation, did not subsequently show more overlap between individual post-tests.

It may be the case that the present analysis method was too crude to make the differences between the groups visible. For example, one person in one of the idiosyncratic groups decided to externally represent all her contributions, which heavily coloured the results. One of the formalism groups, on the other hand, decided to not write anything down before all agreeing upon it. Although this meant acting in the spirit of the formalism, it lead to an under-representation of contributions on the board.

Participants' Thoughts, Perspectives and Negotiation

The qualitative analysis yielded observations about multiple perspectives and negotiation and participants' awareness of grounding and perspectives. Interview excerpts are provided to give the reader an impression of the data that informed our observations. These excerpts are coded as follows; A, B and C are the formalism groups, and X, Y and Z are the idiosyncratic groups. The numbers 1, 2, and 3 designate the different team members. Each excerpt also is given a time stamp (format mm:ss), which designates the moment during collaboration at which the reported thought occurred. An excerpt coded "C2 12:34", for example, designates a thought from member number 2 of group C, which is a formalism group. This thought occurred after 12 minutes and 34 seconds of collaboration.

Multiple Perspectives and Negotiation

In the interview data, the presence of multiple perspectives is apparent through frequent misunderstandings among participants and through the experience of benefiting from each others' contributions in solving the task.

The interview data show that at a number of moments participants detected a misunderstanding on the part of one of their discussion partners. Sometimes they felt they were justified in this judgement, for example, when it was plain from discussions that one of the discussion partners was not able to follow the discussion at all. At other times though, participants merely assumed that one of their discussion partners was incorrect. In some of these cases these participants had no intention of questioning their own views, whereas later it could even become apparent that they themselves were ‘wrong’ and their discussion partners ‘right’:

> Y1 15:09 “‘Cause she said much, but half of it wasn’t relevant.”
> X1 09:43 “If you ask me, the one on the left, I think she just lost it.”
> B3 26:56 “But at that point I thought, no that’s really a stupid remark. And later on he even convinces me, if I recall correctly. But, er, at that point I immediately think no, that’s stupid.”

At times, participants would sense a major misunderstanding between them and their discussion partners without trying to repair this misunderstanding:

> B3 42:05 “This is really bad of me. Cause I nod while I haven’t heard a word of what she said.”

In the above misunderstandings, there was no attempt to repair them, based upon the interview data. Participants seemed to be content with their own, or their partners’, lack of understanding for the purpose of solving the case. Such misunderstandings, without ambition to overcome them, were observed in both the idiosyncratic and the formalism groups. The data suggest that the formalism does not support negotiation in cases where participants are not at all interested in trying to understand each other.

Despite the misunderstandings mentioned above, participants also reported that they had the opportunity to learn from each other during the problem solving process:

> B2 30:31 “I enjoyed seeing how you can build on one another until you reach a good answer.”

From the interview data it seems that participants felt they were able to achieve a better result collaboratively than they would have by themselves, because they gathered knowledge and insights from their discussion partners and in some cases collaboratively constructed the meaning of some piece of data in the computer system.

> C1 41:42 “At this point I’m finding out that through discussions you can ehm, gain far more insight in what’s happening.”

Participants in both the formalism and the idiosyncratic groups reported that they were able to profit from each other.

The above observations indicate that the research setting, i.e., a multi-perspective higher education problem solving situation, offered enough opportunity for negotiation activities, if misunderstandings become explicit. The setting appeared to allow for collaborative negotiation of meaning and knowledge construction, as the participants reported being able to benefit from each other's knowledge. From these observations it is apparent that the setting under study is valid for studying the framework.

Participants' Awareness of Grounding and Perspectives

Participants' awareness of grounding is apparent from the interview data in two ways. First, various interview excerpts show that participants consciously account for each others' economic expertises. Second, there are several instances that show that participants consciously clarify and verify their understanding, in various degrees of commitment to achieving and maintaining common ground.

From the interviews it seems that participants deemed the differences in background between them and their partners relevant to the topic at hand:

> Z1 18:41 "But I had heard that she had done Econometrics, was currently doing Economics, so she's probably looking at totally different things than we are."
> Z3 00:46 "And, uh, I don't know what they majored in, but uh, I think they know more about company management than I do."

On the other hand, sometimes perspectives were seen as a cause of ignorance:

> B3 23:57 "I also thought, he's studying Econometrics, so maybe he doesn't really get the picture."

In both idiosyncratic and formalism groups, participants were aware of such differences in perspective.

In sum, it appeared that the participants linked the perspective of their discussion partners to differences in problem representations, and that they thought that these different perspectives might add new points of view to the discussions. Both acknowledging the difference between understanding and agreement and acknowledging differences in perspective can be seen as a prerequisite for effectively using the formalism, because the rules of the formalism imply a certain awareness of these differences.

The interview data also suggested various, qualitatively different ways in which participants reported activities related to grounding. Three ways in which participants attended to their team partners' contributions were observed. First, they actively tried to increase participation of the discussion partners, especially in the case of someone who was very quiet:

> Y2 06:52 "I wanted to hear what she had to say. Because she had been quiet for some time already."
> B2 01:24 "So I let him finish, that speaks for itself, that you hear what someone's got to say."

A second example of grounding was being actively open to feedback from discussion partners, and being able to accept it. In various cases participants made contributions that their discussion partners deemed false. In such cases, being open to arguments against one's own position is seen as being open to clarifications from other team members:

> C2 34:43 “So maybe [I didn’t make] the best suggestion, but the feedback was very good. Both [team partners] gave me a reason why you could see it differently.”
> Y1 02:57 “I liked that. . . . . they explained to me how I arrived at a wrong interpretation.”

The third way takes commitment to grounding one step further. There were specific cases where participants reported an active effort to achieve common ground in cases of misunderstanding on the part of one of their discussion partners, even when this discussion partner did not actively request any further explanation or clarification and could have been easily ignored:

> B3 23:41 “And it’s probably from his eyes that I could tell, or his way of, I don’t know, his behaviour at that moment, and that’s why I thought, I have to ask if he understands it. Because at that moment I’m like, he doesn’t get it.”
> B2 26:28 “I tried to explain it to him in such a way that he might understand it as well.”
> C1 54:47 “We found a conclusion, of which I found that it was one hundred percent of what the three of us thought. . . . . What I found most important, was that we had a conclusion we all could support [rather] than a conclusion that was one hundred percent correct, but to which one of us did not agree.”

In conclusion, despite often misunderstanding each other, or having disagreements, participants also actively tried to establish common ground. This was apparent in attempts to involve silent team members in the discussion, in being open to feedback from others, and in actively trying to repair misunderstandings. This last type of commitment to grounding was only observed in the formalism groups. The other types grounding were observed in both formalism and idiosyncratic groups.

Use of the Formalism

In the interviews, the participants mentioned a variety of uses for the board. The most important use seems to be that it acts as a record for points that everyone agrees upon. When something is written down on the board, it seems to get the status of common ground. This was the case for both the formalism groups and the idiosyncratic groups:

> Y3 14:49 “I didn’t really agree with what he said. You can tell, ‘cause I didn’t write it down.”

In some cases participants attributed a certain status to a point of discussion if it was written on the board:

> C3 54:46 “So just confirming officially by writing it on the board.”

The board also was reported to help structure the discussion. Participants saw the use of the board as a mechanism to keep track of the various topics that had been discussed, and sometimes used it as a structuring tool by using specific space on the board for a specific part of the case (for example, the left side of the board for market analysis, and the right side for company analysis):

> C1 09:23 "And I thought if we don't write that down we're gonna lose structure, so."

Furthermore, in some cases participants who brought little to the discussion judged that they could make themselves useful by writing things down on the board. From the interview data it is also clear that this was seen as useful by the discussion partners:

> X3 05:12 "Then I'll start writing on the board. That way I can do something useful as well."

The board thus appeared to serve as a lasting record for knowledge that was part of common ground, as a means to structure the discussion, and even as an excuse for having relatively little to contribute in terms of content. Participants did not, however, report thoughts in which they linked use of the board to the formalism.

Conclusion and Discussion

This chapter reports on a framework for decision support of multidisciplinary teams, and an exploratory study to test its method.

With regard to negotiation, results indicate that the formalism groups spent more time on negotiation processes than those not given the formalism instructions (i.e., who used their own idiosyncratic representation method). Furthermore, members of the formalism groups participated in more of the discussion topics than those in the idiosyncratic groups. This, and the fact that the formalism groups discussed more conversation topics than the idiosyncratic groups suggests a more equal representation of different perspectives in the collaboration process than in the idiosyncratic group. The difference in total time might also be an effect of difficulties in following the formalism instructions. However, such difficulties are not represented in the number of contributions, verifications, clarifications or elaborations, since talk about the formalism instruction was coded regulation. With respect to the first research question, namely whether the formalism influences the way negotiations take place, the results lend credence to the hypothesis that the formalism is able to make negotiation more explicit.

With regard to common ground, the results did not indicate any differences or trends between the groups. This may be due to the crudeness of the common ground-measure used here. If the formalism did positively influence the extent of common ground, it was not apparent in the results.

The qualitative analysis of the interview data showed that the present research setting acted as a valid model of multidisciplinary problem-solving for studying the framework. Furthermore, it showed that the participants were, to varying degrees, committed to achieving and maintaining common ground. The results suggested that commitment to grounding was exercised most actively in the formalism groups. However, more research is needed to firmly anchor this statement.

From the interviews it was clear that the board served a variety of purposes during the collaboration. It did not become clear whether the participants were aware of the formalism as intended in this specific setting. Nonetheless, participants did discuss their formalism instructions during work, for example when they were deciding whether to use the board or the flip-over to record their ideas.

In the present study a pen-and-paper implementation of the formalism was used. This appeared to have certain drawbacks compared to an ICT-implementation. For example, one can only hope, and cannot enforce, that participants sufficiently follow instructions, and make use of all the materials that are part of these instructions. Had an ICT-implementation of the formalism been used, participants could have been "forced" into certain activities to make them adhere more closely to the formalism.

In conclusion, it appears to be justified to conclude that the formalism, as tested here, may facilitate negotiation and achieving common ground in decision-making teams with multiple representations. However, further research is needed to decisively test this. In future studies, a larger number of groups will be tested, using a more sophisticated (ICT-) implementation of the formalism. Scripting methods, for example, can 'force' or 'coerce' participants to adhere to the formalism more closely. Furthermore, future studies will test whether the variety of different perspectives affects the hypothesised effectiveness of the formalism.

Chapter 3

Computer Support for Knowledge Construction in Collaborative Learning Environments[1]

Organisations increasingly use multidisciplinary teams to construct solutions for complex problems. Research has shown that multidisciplinary teams do not guarantee good problem solutions. Common ground is seen as vital to team performance. In this chapter an ICT-tool to support complex problem solving is studied. A framework for knowledge construction inspired the design of computer support for knowledge construction. The basic support principle consisted of making individual perspectives explicit, which serves as a basis for negotiating common ground. This principle was embedded in a collaborative learning environment in three ways, which differed from each other in the extent to which users were coerced to adhere to the embedded support principles. Coercion, as expected, was correlated with negotiation of common ground; the more coercion, the more participants would negotiate the meaning of contributions to the ICT-tool, and the more common ground they would have. Self-report data suggested that Intermediate coercion resulted in the least common ground. This may have been caused by some disruption of group processes.

[1] This Chapter has been published as: Beers, P. J., Boshuizen, H. P. A., Kirschner, P. A., & Gijselaers, W. H. (2005). Computer support for knowledge construction in collaborative learning environments. *Computers in Human Behavior, 21*(4), 623-643.

Multidisciplinary teams are used in industry, government and education (Derry, Adams DuRussel, & O'Donnell, 1998), because they are regarded as a *sine qua non* for solving complex problems (Vennix, 1996). The main advantage of multidisciplinary teams is that the team members can bring different perspectives to bear on a problem. Multiple perspectives are expected, for example, to allow for rich problem analyses and solutions (see Lomi et al., 1997). Courtney (2001) argues that business organisations need to integrate different perspectives to ensure organisational sustainability. Hasan and Gould (2001) showed that ignoring certain perspectives on a complex problem can lead to unexpected adverse effects of the ultimate problem solution. And finally, Vennix (1996) notes that, “differences of viewpoint can be very productive” (p. 1). However, multidisciplinarity is not always an advantage. Sometimes individuals outperform multidisciplinary teams, even when it concerns the task of complex problem solving (Barron, 2003). The question is thus: What makes a multidisciplinary team successful?

Recent research by Barron (2003) in the domain of education confirms empirically the need for cognitive frames of reference. She showed that team performance is related to team interaction. She noted that the willingness to construct a shared problem space seemed to be essential for engaging multiple perspectives. High performing teams engaged solution proposals through discussion and acceptance, whereas low performing teams ignored and rejected proposals. According to Johnson and Johnson (1994), synthesis of multiple perspectives might result in better decisions and solutions to complex problems. Bromme (2000) argues that a team needs some *common ground*, a shared cognitive frame of reference, before it can attempt to synthesise perspectives. It seems that members of multidisciplinary teams need to find some kind of commonality between their different perspectives in order to benefit from them.

Many researchers have used ICT-tools to facilitate complex problem solving in teams. These tools use *formalisms*, which are constraints that structure conversation and discourse among collaborators with the aim to guide the exchange of knowledge and information. ICT-tools have been used to support complex reasoning, task-oriented activities, and collaborative knowledge building (Lipponen, Rahikainen, Lallimo, & Hakkarainen, 2003). Specific formalisms are tailored to facilitate specific aspects complex problem solving, and ICT-tools *coerce*[2] (Dillenbourg, 2002) people to follow the rules of such formalisms. To give some examples, ICT-tools used specific formalism to facilitate teams in diverse fields and topics as design activities (Buckingham Shum, MacLean, Bellotti, & Hammond, 1997), scientific reasoning (Suthers, 2001), and argumentation (Van Bruggen, 2003). Such tools have attained good results on cognitive aspects of group learning by focussing on task aspects.

---

2 Some dictionary definitions (Webster’s student Dictionary, 1996) of coercion hold that to coerce involves ‘to constrain or force to do something’. We wish to stress that this thesis uses to coerce in the sense of constraint, not force.

However, they have not explicitly addressed the problem of common ground in multidisciplinary teams.

In this chapter, we describe facilitating the negotiation of common ground. We report on NegotiationTool, a collaborative learning environment (CLE) with an embedded formalism to support negotiation processes. NegotiationTool coerces the users into exploring each other's perspectives to augment the negotiation of common ground. The optimal level of coercion is a trade-off between the impact aimed for (high coercion) and keeping the collaboration 'natural' to the users (low coercion) (Dillenbourg, 2002). Research has shown that a low level of coercion may lead to small effectiveness of a formalism, whereas high levels of coercion may disrupt collaboration to an extent that it starts to hamper collaboration.

First we describe our framework for supporting negotiation. From this framework we will then derive the design primitives for NegotiationTool, and describe three different versions of this tool, that differ with respect to the amount of coercion applied to the participants. The first research question is whether a grounding formalism facilitates the grounding process, and the second research question regards the relation of coercion and negotiation of common ground. We tested the effects NegotiationTool on the grounding process and common ground itself.

A Framework

In Barron's study (2003), members of successful teams engaged in each other's thinking, whereas members of low performing teams typically ignored each other's proposals. Performance depended on the negotiation of a shared problem space as a basis for the construction of complex problem solutions. Team members critically explored each other's thinking, and explicitly accepted, agreed, and subsequently documented contributions to the discussion, which ultimately resulted in better problem solutions. Barron (2003) produced very useful results for the study of problem solving teams. However, her research context, sixth-grade triads solving complex problems, may impose some constraints on generalising her results.

In our framework, we address both knowledge construction, to reflect on how individual knowledge becomes part of a solution to a complex problem, and group processes, to reflect on the team processes that take knowledge from being in the 'mind' of one learner to becoming a team's constructed knowledge. The framework is inspired by sources on social learning (e.g. Salomon & Perkins, 1998; Sullivan Palincsar, 1998), knowledge sharing (e.g. Boland & Tenkasi, 1995; Walsh, 1995), and grounding (Baker et al., 1999; Bromme, 2000; Clark & Brennan, 1991). It is an attempt to link the solution requirements in terms of constructed knowledge, and the group processes that underlie the construction of this knowledge.

The route from unshared knowledge in one participant's head to newly constructed knowledge in a team goes through three intermediate forms (i.e., external knowledge, shared knowledge, and common ground) via four processes, namely externalisation, internalisation, negotiation and integration (see Figure 3.1).

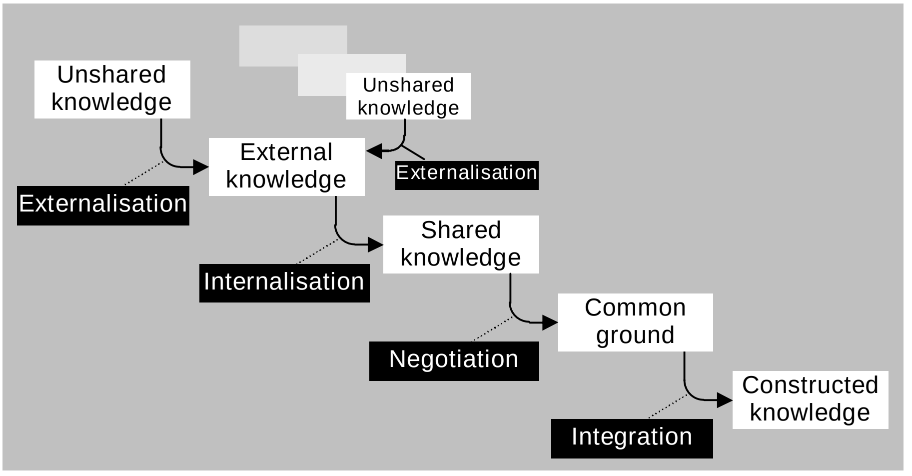

Figure 3.1. From unshared knowledge to constructed knowledge

Private knowledge is externalised when team members make their, as yet, unshared knowledge explicit or tangible to others (Leontjev, 1981), for example by making a contribution to a conversation. Once a team member has made such a contribution, the others can try to internalise it. While constructing their individual understanding, they can consider aspects of the contributor such as background, current situation, and views to better “understand” the contribution. Also, their own beliefs and assumptions play a role while they try to understand the contribution. A contribution is thus understood against the presumed perspective of the other, as well as against one’s own perspective (Bromme, 2000). Having *shared* a contribution with a team does not mean that the team members all have arrived at the same understanding. All kinds of representational differences result from interpreting a contribution in one’s own perspective only (a graphical designer has a different understanding of, and use for the term “elegance” than a computer programmer) or from minimising or rejecting its validity or plausibility due to differences in conviction or opinion.

A shared contribution is the starting point for negotiation of common ground. Common ground is a shared cognitive frame of reference (Bromme, 2000). It is through the process of internalising others’ contributions, and subsequently providing feedback based on one’s own perspective by word or action, that common ground can be negotiated (Alpay et al., 1998; Baker et al., 1999). Common ground is never absolute or complete, but is continually accumulated and updated (Clark & Brennan, 1991).

We conceive of negotiation of common ground as a dual concept. *Negotiation of meaning* leads to an agreement regarding meaning and understanding of a contribution. It concerns people making public to others their private understanding of some contribution, verifying whether and to what extent their own understanding is different from what others intended them to understand, receiving feedback on this, that is clarification, re-verifying, and so on, until “the contributor and the partners mutually believe that the partners have understood what the contributor meant to a criterion sufficient for the current purpose”

(Clark & Schaefer, 1989, p. 262, the grounding criterion). *Negotiation of position*, the second part of negotiation, concerns people making public to others their private opinion about a contribution, to check whether one's position is clear to others, and vice versa. One might debate, philosophically, whether opinions and truisms can be discerned from each other. Instead of getting into this debate, we want to point out that the difference between truth and opinion is assumed to be meaningful to the negotiators. Note that neither of these definitions imply the more common, generic use of the term negotiation, namely to discuss with an opposing or adversarial party until consensus or compromise is reached.

Starting from common ground, new knowledge can be built by adding new relations and concepts to common ground, via integration. Knowledge construction is based on the common ground the team has built, and will broaden and deepen the common ground because the common constructed knowledge becomes part of the common ground. With regard to problem solving, constructed knowledge represents the solution(s).

A Formalism to Support Negotiation

In this section we use the above framework to construct a formalism for the support of negotiation. The steps from unshared to constructed knowledge serve as a basis for the formalism. It consists of *primitives* of negotiation, and *rules* that prescribe the use of these primitives. Primitives can be seen as basic building blocks that model a specific type of dialogue (Dillenbourg, 2002). We couple these primitives with a set of rules, to mimic the negotiation process as explicitly as possible, which results in a formalism for negotiation. Note that this formalism models an ideal negotiation process; in regular communication, the status of people's statements in terms of negotiation primitives remains implicit. The formalism must enable distinguishing between original contributions, clarifications, verifications, et cetera, making the steps explicit. By doing so, individual differences in understanding and opinion should more easily surface.

First, negotiation starts with a *contribution* (Primitive 1) of some sort, such as a hypothesis or a position, which is assumed not to be part of a team's common ground (Rule 1). To assist in detecting differences between individual representations, every team member must *verify* (Primitive 2) their understanding of another's contribution (Rule 2) because people articulate and understand the contribution against their own background knowledge (Fischer et al., 1995). Third, a contribution needs to be *clarified* (clarification, Primitive 3), using the ideas upon which it was based. For example, the educational background or the political orientation of the contributor may shed light on the meaning of his/her contribution. Nevertheless, a clarification need not always be made by the original contributor, but may also be performed by another team member who feels knowledgeable. Rule 3 is that all verifications require a clarification. Together, Rules 2 and 3 can be iterated until common understanding of the contribution is reached. Note here that a correct clarification of a contribution one team member can be seen as a successful verification by another.

The fourth primitive is *acceptance/rejection* of a contribution, which refers to whether one can judge a contribution as true (acceptance), based on the explanation given, or judges it untrue, or unintelligible (rejection). For example, the statement 1 + 1 = 10, is true only if we understand (through Rules 1 and 2) that the contributor is using the binary system. A contribution should be accepted as part of the common ground if it is true, or after it has been modified so that it has become true. Rule 4 is that every contribution needs to be accepted or rejected by the team members. Finally, Rule 5 is that people must explicitly state their own *position* (Primitive 5) on the contribution. In the case of irresolvable disagreement about previously accepted statements, Rule 5 may result in multiple scenarios, each based on another position (i.e., agree to disagree). This also means that one may accept a certain contribution, but disagree all the same, for example when neither person can prove the other wrong. In such cases, people can agree to disagree, and alternate representations that are equally legitimate can ensue. Table 3.1 summarises these rules.

Table 3.1
Rules for a Formalism for the Facilitation of Negotiation

1.	Every new issue is termed a *contribution*
2.	Contributions require a *verification* by the other team members
3.	Each verification is responded to with *clarification* by the original contributor
4.	When all verifications are clarified, and no new verifications are performed, all team members state whether they *accept* or *reject* the statement
5.	All team members state their *position* about accepted statements

NegotiationTool

The formalism for supporting negotiation was implemented in an ICT-tool called the *NegotiationTool* (NTool). NTool is based on a newsgroup reader, featuring (a)synchronous, distributed, text-based discussions. To optimise the NTool for negotiations among multiple representations the formalism was implemented to structure the negotiation process in three ways, with increasing levels of coercion (cf. Dillenbourg, 2002).

Coercion refers to the degree of freedom participants are allowed in following a formalism. Coercion and formalism together constitute a collaboration script. The higher the coerciveness of a script, the more participants are required to adhere to its formalism. Scripting requires “subjects on most or all occasions to make a particular type of speech act in a specific context.” (Baker & Lund, 1997, p. 176). For Dillenbourg (2002) a “script is a set of instructions regarding to how the group members should interact, how they should collaborate and how they should solve the problem.” (p. 64). This means that a script can be aimed at either the interaction and collaboration level, for example by offering sentence openers or prescribing communicative acts (e.g., Baker & Lund, 1997; e.g., Barros & Verdejo, 1999; Soller, 2002) and/or the problem solving process, for example in problem-based learning. In such cases, scripting results in the use of distinct phases for discussion, with distinct purposes with regard to problem solving (Barrows & Tamblyn, 1980; Dillenbourg, 2002; O'Donnell & Dansereau, 1992).

A script that uses very little coercion leaves participants many degrees of freedom whereby usage of the formalism attains a high degree of idiosyncrasy. A script with a high level of coercion constrains the number of options participants have, thus guiding them along the lines of the formalism. In the study reported here, three different ICT-implementations of the formalism were implemented (they are dealt with in detail in the Method-section). One implementation had very little coercion, and was called the 'Idiosyncratic' version. One could compare this situation with giving a person a set of lines and symbols to be used in constructing a diagram, but leaving it up to her/him to decide which symbols and lines to use for what purpose. A second used medium coercion and was aimed at the problem-solving level (here termed 'Scripted', in appreciation of Dillenbourg's (2002) use of the word). The third used scripts aimed at the interaction and collaboration level, using high coercion and was called 'Stringent'. In each version, coercion was specifically aimed at verification and clarification primitives, that is, at the extent to which people were required to verify and clarify in specific circumstances.

NTool was expected to increase negotiation of both meaning and position because it forces people to make their private understandings and opinions public, making differences in understanding and opinion visible or salient (Bromme, 2000). More specifically, we hypothesised that (1) coercion would be correlated with negotiation, that is, the higher the level of coercion, the more negotiation. We expected this correlation to hold for both verifying and clarifying. Likewise, we hypothesised that (2) the amount of negotiation of meaning per contribution would be correlated with coercion as well. Differences were also expected with regard to common ground; (3) common ground was expected to be highest in the Stringent version and lowest in the Idiosyncratic version. All three hypotheses rest on the assumption that more coercion will make participants follow more closely an ideal model of negotiation, as laid down in the formalism. Nonetheless, differences caused by the different ways in which coercion was implemented were also studied exploratively, because scripts that are too coercive can be counterproductive if they disrupt collaboration (Dillenbourg, 2002).

Method

Participants

Participants were students in their senior year from the Maastricht University from the departments of Cultural Sciences, Economics and Business Administration, and Psychology. Seventeen multidisciplinary groups were formed by assigning participants from different degree programmes to teams of three. These participants were assumed to have different perspectives due to educational differences and socialisation effects from their educational careers.

Task

Participants were required to solve the “school drop-out” case (Kirschner, Van Bruggen, & Duffy, 2003). They received the following task description: “You have been asked by the government to advise the Minister of Education how to solve the high school drop-out problem. At the end of the session you are expected to come up with a viable solution that can be implemented as government policy.”

Formalism

Each group was supplied with NTool. Three different collaboration scripts were used.
Idiosyncratic. This version used only the primitives. On-screen information was presented about every contribution, and whether it needed yet to be verified or decided upon (agreeing or disagreeing). Furthermore, each participant was informed when he/she had not yet verified all contributions, and when he/she had not yet decided on all contributions.
Scripted. This version used the same primitives as the Idiosyncratic version, but the problem solving process was now divided into two distinct phases. Phase 1 was aimed at negotiation of meaning. Here participants could compose contributions, verifications, and clarifications. To end the first phase, all participants had to accept or reject all contributions. Participants were specifically informed that they were to refrain from stating opinions during this phase. Phase 2 was aimed at negotiation of opinion and ended when all contributions had been decided upon (i.e., there were no more contributions on the agenda). Participants were no longer allowed to compose new contributions. Using prompts, participants were informed in which phase they were.
Stringent. This version also used the same primitives, but allowed negotiation of only one contribution at one time. Furthermore, participants were not allowed to compose reject-, agree-, and disagree-messages before the contribution had been verified. Using prompts, participants were informed as to whether they had to verify or decide on a contribution.

Six groups used the Stringent formalism (high coercion group), five had the Scripted formalism (low coercion group) and the final six groups could use the NTool Idiosyncratically (no coercion condition).

Procedure

The procedure entailed two phases and an interview.
Practice phase. First the participants received a 20-minute tutorial on the ICT-environment that addressed the basics of NTool communications, and then proceeded to emphasise the rules of the formalism, and the way they constrained communication. To ensure that participants were proficient with NTool a practice case (solving the problem of road traffic safety) was used to enable them to gain experience with the NTool. Participants practised a total of 45 minutes.
Experimental phase. After a 15-minute coffee break, participants started working on the experimental (school drop-out) case. To promote the construction of an individual problem representation and to allow the researchers to determine what this representation was, participants first had to carry out the task

individually (pre-test, 20 minutes). Participants could take notes while working individually on the task. Next, they solved the problem collaboratively (90 minutes), and after that individually again (post-test, 20 minutes). All resulting individual problem representations and solutions, as well as the group problem representation and solution were recorded. In their post-test, participants were also asked to state the points on which they felt that they had differences in opinion with their team members, to account for agreeing to disagree.

Interviews. Three focus group interviews were held, one for each version of NTool. The main purpose of the interviews was to derive hypotheses to interpret the potential effects of the NTool, and to explore possible effects of coercion and the formalism.

Variables and Analysis

Analysis involved negotiation, common ground, and participants' perceptions of coerciveness. Two operationalisations for negotiation were used, namely *quality of negotiation* and *negotiation per conversation topic*. Negotiation was measured by analysis of the collaboration. Common ground was also measured in two ways. Firstly, by comparing individual representations before and after collaboration, and secondly, by questionnaire (Mulder, 1999). Participants' perceptions of coerciveness and effects of the formalism were measured by qualitative analysis of interview data.

Quality of negotiation. A coding scheme for coding function and content of messages during collaboration was developed (cf., e.g., Avouris, Dimitracopoulou, & Komis, 2003; Fischer et al., 2002; Mulder et al., 2002; Thomas et al., 1982). All messages were coded with regard to:

- Cognitive content - directly related to solving the problem.
- Regulative content - related to monitoring the problem solving process, regulating the collaboration process, which also entailed tool use.
- Other content - not in any other category or non-codeable.

Messages with cognitive content were specifically coded for function. The following subcategories were used to code negotiation:

- Contribution: A new topic of conversation that has not been discussed before is introduced.
- Verification: Information is directly or indirectly requested about the intended meaning of a contribution or elaboration.
- Clarification: A reaction to a verification or a perceived lack of understanding, in which the intended meaning of a contribution or elaboration is elucidated.
- Acceptance: A reaction to a contribution in which the contribution is judged intelligible and/or correct.
- Rejection: A reaction to a contribution in which the contribution is judged unintelligible and/or incorrect.
- Agreement: A reaction to a contribution in which the sender voices his/her agreement with the contribution.
- Disagreement: A reaction to a contribution in which the sender voices his/her disagreement with the contribution.

In many cases, messages did not fit any of the above subcategories, for example if people built on each other's communications, without explicitly negotiating meaning of, or position on a contribution. Such messages were coded Elaboration: A contribution is elaborated upon by adding information or summarising.

A research-assistant was trained for 10 hours to use the coding scheme (he had already received 25 hours of training earlier in a comparable coding scheme). The data from the practice phase were used for training purposes. Comparing one randomly selected experimental session coded by the author and the research-assistant resulted in a substantial (Landis & Koch, 1977) inter-rater reliability (Cohen's kappa) of .70 (*SE* = .034). All data were coded by the research-assistant.

Verification and clarification, in contrast to elaboration, were considered indicative for explicit negotiation activities. The total number of contributions discussed was used as an indicator for the range of topics discussed.

Negotiation per conversation topic. To measure the number of verifications and clarifications per conversation topic, episodes in the discussion that dealt with one conversation topic were first identified. The contributions identified with the coding scheme for negotiation were considered starting points for a new discussion episode. An episode generally started with a contribution and ended when one of the participants would make a new contribution, and all of the discussion in between these contributions dealt with one conversation topic. For each group, negotiation per conversation topic was then calculated by dividing the sum of all clarifications and verifications by the number of contributions.

Common ground. Common ground was operationalised as the extent to which the content of individual representations was present in individual representations. To characterise the content of the individual representations the discussion content itself was characterised (see Figure 3.2). The discussion episodes identified earlier served as a basis for characterising the discussion content.

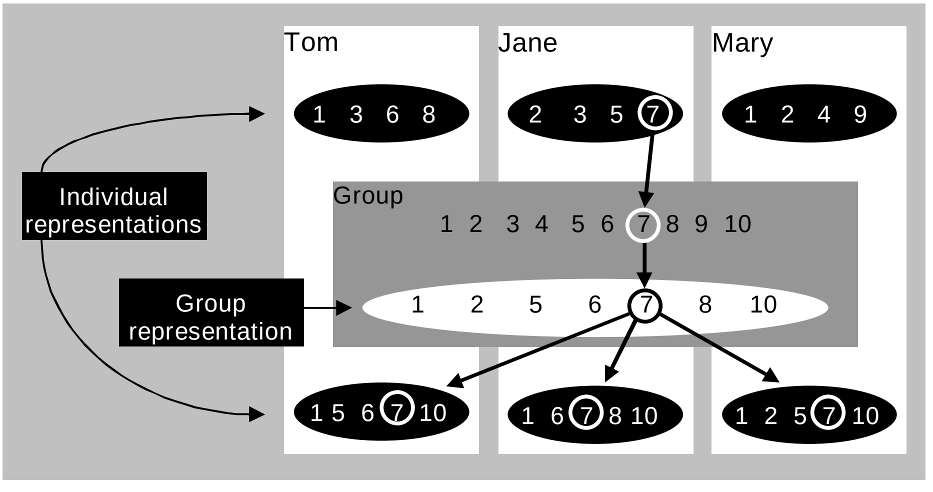

Figure 3.2. *Analysis of common ground; numbers indicate episodes*

Each episode was first numbered and summarised. The next step involved characterising the content of all individual representations, both initial (pre-test) and subsequent to collaboration (post-test), and the group representation. The summaries were used to identify the content within individual the representations. For every individual representation the topics that were and were not represented were assessed. For example, in Figure 3.2 episode number 7 is present in Jane's initial individual representation, in the group discussion, and in all post-tests. By repeating this procedure for each of the episodes in the discussion, the origin of each conversation topic, whether it was present in the group representation, and whether participants used it in their post-tests was determined. Using these data we computed, for each group, the mean number of pre-tests and post-tests that a contribution would end up in. This mean number of post-tests per contribution was used as a measure of common ground.

We also adapted some questions from Mulder's (1999) instrument for measuring understanding, which measures various cognitive and social aspects of common ground and shared understanding. Some of the questions from the original instrument were omitted because they assumed that participants would have multiple meetings instead of one. Questions referred to understanding of the problem definition ("How well did you understand the problem definition?"), shared understanding of the problem ("To what extent did you and your group members obtain the same understanding of the problem?"), social relations between the participant and his team members ("To what extent do you feel you know the other group members?"), social relations between the other team members ("To what extent do you feel the other group members know each other?"), and problem approach ("To what extent did you and your group members agree about the problem approach?"). Questions were posed in the form of 6-point Likert-scales.

Statistical analyses. Statistical testing was done using Kruskal-Wallis tests. Subsequent testing for the directional effects (all previously stated hypotheses) was done using Spearman's test for non-parametric correlation (i.e., verification, clarification, negotiation per contribution, and common ground). Subsequent testing for the other variables (e.g., contributions, regulations) was done for contrasts using the Mann-Whitney U test. Subsequent tests were only performed in case the Kruskal-Wallis test was significant at the .05-level. Testing of contrasts was also done for expected directional effects that were significantly different in the Kruskal-Wallis test, but turned out not to be correlated. All tests concerned group means (N = 17).

Interview. All participants were asked whether they wanted to join one of the interviews. Unfortunately, few participants did so, due to practical reasons of planning as well as low interest for the interviews. The Idiosyncratic focus group numbered 5 participants, the Scripted group 1, and the Stringent group 2.

The interview was conducted on the basis of a semi-structured interview guideline developed prior to the interview. The questions were directed at participants' perceptions of coerciveness in NTool. For example, participants were asked how they used NTool, what they liked and/or disliked in NTool, and how discussions in NTool differed from discussions in general. At the beginning of the interview, participants were told that all different opinions were of equal importance, and they were explicitly invited to voice dissenting views and opinions if they held such. The interviews were analysed by the author. The interviews were videotaped and analysed using The Observer® (Noldus et al., 2000), a software package designed for behavioural observation using video data. An open coding approach was used, which focussed on the effects of the formalism on participants' interactions, and the differences between the three versions of NTool, with special focus on coercion.

Results

Negotiation

We compared the different conditions with regard to number of codes (Table 3.2). Statistical analyses using Kruskal-Wallis tests revealed significant differences between the conditions for the number of contributions, $\chi^2(2, N = 17) = 8.85, p < .05$, number of verifications, $\chi^2(2, N = 17) = 7.08, p < .05$, number of clarifications, $\chi^2(2, N = 17) = 7.33, p < .05$, number of acceptance messages, $\chi^2(2, N = 17) = 10.58, p < .01$ and number of regulation messages, $\chi^2(2, N = 17) = 8.03, p < .05$.

Computing Spearman correlations revealed significant correlations between coercion, and verification, $r_s(N = 17) = .63, p < .01$, and between coercion and clarification, $r_s(N = 17) = .54, p < .05$. Post hoc contrasting of Idiosyncratic groups with Scripted and Stringent groups revealed a significantly higher number of contributions in the Idiosyncratic groups, $U(N = 17) = 4.00, p < .005$. Finally, contrasting of Scripted groups with Idiosyncratic and Stringent groups revealed significantly higher numbers of acceptance $U(N = 17) < .001, p < .005$ and regulation messages $U(N = 17) = 4.00, p < .01$ in the Scripted groups. In other words the Idiosyncratic groups made significantly more contributions, verification and clarification were significantly correlated with coercion, and the Scripted groups accepted significantly more statements, and regulated more.

Table 3.2
Mean Numbers of Negotiation Primitives

	Condition		
	Idiosyncratic	Scripted	Stringent
Contribution	8.0	5.4	5.0
Verification	8.8	10.2	16.7
Clarification	10.7	9.2	17.7
Elaboration	56.6	35.6	48.5
Acceptance	3.0	13.6	1.8
Rejection	1.2	4.6	1.7
Agreement	8.7	6.0	11.7
Disagreement	1.3	1.6	2.0
Regulation	30.7	106.0	43.7
Other	8.0	8.8	5.0
n	6	5	6

Kruskal-Wallis testing revealed that the amount of negotiation of meaning per contribution (see Table 3.3) differed significantly between the different versions of NTool, $\chi^2(2, N = 17) = 11.17$, $p < .005$. Coercion was found to be significantly correlated with negotiation per contribution, $r_s(N = 17) = .83$, $p < .0005$. These results indicate that contributions were most heavily negotiated in the Stringent groups and least in the Idiosyncratic groups.

Table 3.3
Negotiation of Meaning[a] per Contribution

	Condition		
	Idiosyncratic	Scripted	Stringent
M	2.37	3.51	7.50
n	6	5	6

[a] The sum of all verifications and clarifications.

Common Ground

No statistically significant differences were found with regard to pre-tests, $\chi^2(2, N = 17) = 1.78$, $p = .41$. The distribution of contributions across post-tests was significantly different between conditions, $\chi^2(2, N = 17) = 6.14$, $p < .05$. Subsequent Spearman correlation testing showed that the distribution of contributions across post-tests was significantly correlated with coercion, $r_s(N = 17) = .57$, $p < .05$. This means that the higher the coercion, the higher the number of post-tests a contribution would end up in.

Table 3.4
Common Ground

	Condition		
Mean number of…	Idiosyncratic	Scripted	Stringent
… pre-tests per episode	1.13	1.21	1.10
… post-tests per episode	1.97	2.00	2.39
n	6	5	6

Table 3.5 shows the self-report data for common ground. Kruskal-Wallis tests revealed significant differences for the extent to which the group held the same problem understanding, $\chi^2(2, N = 17) = 6.36$, $p < .05$ and the group understanding of the task approach, $\chi^2(2, N = 17) = 7.88$, $p < .05$. However, contrary to expectations, no correlations were found in the questionnaire data between coercion and common ground. Subsequent testing of the contrasts using Mann-Whitney tests showed that the Scripted version of NTool resulted in the lowest perception of common ground. Both the extent to which the group held the same problem understanding and the group understanding of the task approach were lowest in the Scripted groups, $U(N = 17) = 6.50$, $p < .01$, and $U(N = 17) = 6.00$, $p < .01$, respectively. Furthermore, group understanding of the task approach was highest in the Stringent groups, $U(N = 17) = 12.00$, $p < .05$.

Table 3.5

Questionnaire Data

	Condition		
To what extent...	Idiosyncratic	Scripted	Stringent
... did you understand the problem definition?	4.89	4.73	5.39
... did you and your group members obtain the same understanding of the problem?	4.72	3.80	4.67
... do you feel you know the other group members?	3.39	2.80	3.11
... do you feel the other group members know each other?	3.44	2.67	3.06
... did you and your group members agree about the problem approach?	4.50	3.87	4.89
n	6	5	6

Note. Judgements were made on 6-point scales (the higher the number the larger the extent).

Participants' Perceptions of the Formalism and Coercion

Interview excerpts were coded for condition (Id = Idiosyncratic, Sc = Scripted, St = Stringent) and for interviewee (interviewees are numbered).

Effects of the formalism. The formalism may have caused the participants to refrain from immediately giving their opinions. Some interviewees noted that they were tempted to immediately give their opinion about new contributions:

> Id1: "When I saw a contribution, I first checked whether I had an opinion about it. And if so, then you're generally tempted to Agree or Disagree."

In this regard, one interviewee's observations about the various uses of verifications were relevant as well:

> Sc1: "One use [of verifications] was a normative judgement about contributions, . . . [a second use was to indicate] "I just don't think it's correct," and [3] just really asking for pure clarification."

This may have influenced whether contributions were accepted, because using verifications could change participants' opinions:

> Id1: "Through verificationing [sic] you get that people may come with knowledge you know nothing about. . . . while at first you may think: 'disagree', because then you don't know yet. With a clarification someone can explain something you didn't know, a theory, or a field."

However, others had less trouble not immediately giving their opinions:

> Id5: "In our group people did not have the need to immediately state their opinions. We'd first discuss some, . . . and only after that people would share their opinions."

Furthermore, allowing participants to immediately agree or disagree, as in the Idiosyncratic version of NTool, did not mean that they would indeed do that:

> Id4: "If you agree, then it's not a problem. But if you disagree; . . . you can't do that right away."
> Id3: "You don't disagree right away, because it's rude. You also want to explain why, and give the contributor an opportunity to react."

This remark shows that participants posted verifications instead of disagreements, even when they did in fact disagree.

From the preceding, it appears that the formalism may have affected the acceptance of some contributions. It may have caused participants to refrain from immediately giving their opinion, which afforded changing one's opinion through verification.

About coercion. Some interviewees using the Stringent version made some observations about coercion:

> St2: "If someone posts a verification, then the contributor should be the one to post the clarification, so you have to wait, because he can't answer two verifications at the same time."

Interviewees mentioned that they needed to be able to signal when they came up with a new point during the middle of discussing another point:

> St2: "You don't know if the others have noticed that you posted a new contribution. .... In a regular discussion [in contrast to discussions in NTool] you can say that [you want to raise a new point], and then the other can still propose to return to your point later, but at least you've been able to make known that you have a point." St1: "Especially because you'll have lost your point again when the discussion's concluded."

From these statements we gather that the Stringent version may have caused some disruption because it limited discussions to one contribution only, whereas participants needed to raise other contributions at that time, or had to wait until their team mates had finished their verifications and clarifications.

Other interviewees reported that no coercion at all, as in the Idiosyncratic version, resulted in a lack of closure of discussion topics. Various interviewees remarked that they needed summaries to keep track of the discussion. Some complained about a lack of closure, a lack of being "on the same wavelength":

> Id4: "Some topics were concluded, others weren't. . . . . I think towards the end you need the participants to be obliged to agree or disagree."

One interviewee remarked the following about the phases in the Scripted version of NTool:

> Sc1: "We felt that accept and reject were a bit redundant."

Furthermore, some interviewees noted that the NTool formalism could be improved by adding more facilities for having argumentative discussions:

> Id3: "There isn't really a place where you can convince the others that your idea is the better."

This need was also apparent from the fact that verifications were sometimes used for normative purposes:

> Sc1: "The verifications shifted from being factual to being normative. When we'd hold different opinions we'd accept that, so agree and disagree was more of a formal closure of the discussion."

The interview data suggest that to some extent argumentative discussions were taking place using verifications and clarifications. Participants may have been able to predict beforehand who would agree and who would disagree. From these statements it can be concluded that the current distinction between meaning making and position taking in the Scripted version may have had little use, especially when negotiation of position had effectively taken place in verification and clarification messages. Nonetheless, the same interviewee noted some unique uses for the reject and accept messages:

> Sc1: "You accept something because you feel that it's right. . . . you can disagree with it. . . . We used reject for contributions that we did not have enough time for, for some contributions that belonged in the chat, and one that was utter nonsense."

It can be concluded that every version of NTool has its specific advantages and drawbacks. The interview data do not favour one version above all others.

Conclusions and Discussion

The present research studied the relationship between negotiation, the negotiation formalism, and coercion, with the ultimate goal being to design an ICT environment that facilitates knowledge construction. The main approach was the design of a formalism for the facilitation of common ground, which appears to be a prerequisite for knowledge construction. Three versions of NTool, an ICT-tool for group discussion with a formalism for support of negotiation, were studied. The Idiosyncratic, Scripted and Stringent versions of NTool differed with regard to the extent to which they coerced participants to hold to the formalism. Coercion was expected to be positively correlated with negotiation, negotiation per contribution, and common ground.

The results showed that the three versions of NTool differed with regard to negotiation, negotiation of meaning per contribution, and common ground. Subsequent testing revealed that coercion, as we hypothesised, was significantly correlated with negotiation and negotiation per contribution. It appears that

NTool does affect the negotiation of common ground, and that it does so increasingly with more coercion. With regard to common ground, the analysis of individual post-discussion representations suggests a significant correlation between coercion and common ground, as hypothesised. However, the questionnaire data about common ground revealed that the Scripted version of NTool resulted in less common ground than the other versions, as perceived by the participants.

Further analyses revealed some unexpected results. Discussions in the Scripted version of NTool featured significantly more acceptance of contributions, and regulation. Both effects may have been caused by the specific way coercion was implemented in this version. Regarding acceptance, the Scripted version used a distinct "acceptance" message-type, to be able to distinguish between the meaning-making and position-taking phases of discussion, whereas in the other versions acceptance was implicit in agreeing or disagreeing, and therefore not used in a separate message type. Regarding regulation, taking the discussion from the meaning-making phase to the position-taking phase may have caused some difficulties. For example, it may have been difficult for the participants to keep track of messages they had not verified yet, causing them to be unable to post agreement and disagreement-messages. This may have confused participants at times.

Also, analyses showed that the Idiosyncratic discussions showed significantly more contributions than the other versions. This may mean that the range of topics was widest in the Idiosyncratic version, which could suggest a trade-off between topic range and common ground. However, it may also be the case that participants in the Scripted and Stringent versions, knowing that they had less opportunity to post contributions, chose to word their contributions more broadly, in which case fewer contributions would still cover the same topic range. Further research may shed some light on these explanations.

Disruption of collaboration (Dillenbourg, 2002), which can be caused by over-scripting collaboration, may explain some of the unexpected results. The need for more regulation in the Scripted version may have caused participants perceive the least common ground, even though the analysis of the post-tests suggested otherwise. The other two versions of NTool did not seem to have any such adverse effects. In sum, NTool seems to influence both negotiation of common ground and common ground itself, and seems to do so increasingly as coercion increases. In the case of the Scripted version, the specific way coercion was implemented may have caused some disruption of collaboration. Both the Idiosyncratic and Stringent version did not seem to influence collaboration in a disruptive way.

The interview data suggested a possible mechanism for the way the formalism affects negotiation. The formalism may have restrained participants from immediately stating their opinion, in which case they may have verified their understanding instead. Subsequent clarification may have changed others' opinions from disagreeing to agreeing. The interview data did not offer any contrary mechanism, where verification may have led someone who initially agreed to disagree. In sum, it may be the case that shedding light on contributions through verification and clarification, instead of immediately taking a position, increases the chances that contributions are accepted, and become part of common ground.

The interview data also suggested more mechanisms that may have caused some disruption. Interview data showed that the distinction between meaning making and position taking (the Scripted version) may have been redundant since argumentative discussions already took place during the meaning-making phase. Furthermore, the interview data suggested that the Idiosyncratic version of NTool lacked some opportunities for closing off discussion topics. Participants experienced the agreements and disagreements as a formal way of closing the discussion topics. This may explain the possible difference in common ground between the Idiosyncratic and Stringent versions because only the participants in the Stringent version were coerced to provide post agreement and disagreement messages, thus effectively closing off discussion topics. Finally, in the Stringent version participants could not raise a point at the time they may wanted to raise it (i.e., a prior contribution must be closed before a new one can be made).

These results do not show that NTool was overly disruptive. The fact that some of the above mechanisms may have caused some inconveniences does not necessarily mean that communication was disrupted. A number of the issues put forward in the interviews are inherent to ICT-tools in general. Signs that can be transmitted through body language in a face-to-face situation are not available for communication in ICT-tools. Research on ICT-tools for collaborating groups has shown that such tools generally feature low participation, diverging discussions, and mixed results regarding social and context-oriented communication (Lipponen et al., 2003). However, NTool has show that increasing coercion can result in more negotiation per contribution, which suggests less divergence of discussions. Regarding disruption, NTool does not seem to distinguish itself negatively from the average ICT-tool.

The results are promising with regard to the facilitation of the grounding process. In her study, Barron (2003) showed that interaction is important for problem solving. In her study, engaging in each other’s thinking was related to better solutions. This study has shown that ICT-tools can be used to facilitate such interactions, by using a formalism for negotiation, and coercing the user into following it. However, more research is required to test our ultimate aim of facilitating complex problem solving. The present study does argue a relation between common ground and the quality of problem solutions, but does not explicitly measure it. Furthermore, the experiment took place in a single 90 minute laboratory session. It remains to be seen how effective NTool is when employed in a more authentic setting, like an educational setting or a professional project team. In future experiments, we plan to take both limitations of the current study into account, by employing NTool in an educational practical setting, and by designing specific measurements for solution quality. Overall, it can be concluded that NTool and its underlying framework affect negotiation of common ground, and that adding some coercion increases this effect, without being harmful to collaboration.

Chapter 4

Measuring Effects of ICT-Tools

Many (ICT-)tools use specific support techniques and instructions for facilitating the externalisation of knowledge so as to produce specific learning outcomes. This knowledge externalisation has been theorised to decrease working memory load, which frees up working memory capacity needed by ICT-tools. However, some research has shown that knowledge externalisation may increase cognitive load. Furthermore, the techniques used to support the instruction may obscure the effects of the instruction itself. This chapter presents the results of two studies. The first is a study on why tools should not only be studied in terms of their specific intended outcomes, but also in terms of their effects on working memory, and the cognitive mechanisms needed to achieve the intended outcomes. The second uses cognitive load measurements and stimulated recall interviews to obtain a more comprehensive view of the effects of learning tools. Results suggest that traditional outcome measures need to be complemented with quantitative and qualitative measures of cognitive processes to substantiate conclusions about intended effects of ICT-tools.

Educators and researchers increasingly use (ICT-) tools to augment learning in a great variety of tasks (Jonassen, 2000; Norman, 1993). For example, Buckingham Shum et al. (1997) studied supporting group design processes, Suthers (2001) studied facilitation of scientific reasoning, and Van Bruggen (2003) studied support of argumentation. All of these researchers have convincingly shown that they achieved the effects they intended with their ICT-tools, that is, learners' task performance was positively affected by the ICT-tools. However, these results beg the question of how they were achieved, and at what cost.

Generally these tools share two characteristics. They make use of *externalisation of knowledge*, and they use specific *instructions* for the externalisation of that knowledge to support a specific aspect of learning. These instructions encourage learners to engage in novel ways of representing knowledge that afford learning (Jonassen, 2000). The ultimate goal is that learners learn to apply, or transfer, these representation methods in contexts without the ICT-tool.

It has been argued (Scaife & Rogers, 1996) that external representation of knowledge provides an off-loading effect to learners, that is, representing knowledge externally lowers working memory load because information that is externally represented does not have to be kept in that memory. This frees up working memory resources, which subsequently affords processing for the instructed activities. However, Van Bruggen, Kirschner and Jochems (2002) found that this need not necessarily be the case. If the specific instruction for knowledge externalisation is poorly designed, for example when it does not fit the task it is intended to support or when the use of the instruction requires training, external representations may increase cognitive load due to the effort needed to apply the instruction in action. The instruction inherent to the use of these tools may thus increase cognitive load, and, in the worst case, could even counteract learning processes. The key to successful ICT-support, thus, seems to reside in the way the user of the ICT-support is instructed to use it for knowledge externalization.

Several measures can be taken to ensure that the instruction given to the user does not cause extraneous (not facilitative to the learning process) working memory load, while still resulting in germane (intended and helpful) effects on collaboration. The instruction given to the users should be plain and simple enough to be used without much training. Furthermore, the execution of the instruction during knowledge externalisation can be supported by using techniques such as coercion (i.e., the extent to which users are constrained to follow the instruction) or sentence openers (which act as advance organisers for the externalisation of knowledge). Such support acts as a *performance constraint* in the sense that it makes actions unavailable that are not relevant to the instructed activities (Van Merriënboer & Sweller, 2005).

By its very nature, performance constraints implicitly structure the learner's actions without giving him/her insight into the reasons why this structure is necessary. Hence, if a task is supported too well it may prevent learners developing a deeper understanding of the actions that are allowed/required, and thus prevent transfer to other situations. This means that a good balance is needed between instruction and performance constraints (at different stages of the learning process), and that a research methodology is applied that takes cognitive load and working memory processes into account, so as to show how, and at what cost, the instruction results in the intended effect.

The goal of this chapter is to study whether and how cognitive load measurements and interviews can complement traditional outcome measures (i.e., intended effects) to inform the design of support and instruction in ICT-tools. First, an earlier study with NTool (Beers, Boshuizen, Kirschner, & Gijselaers, 2005) is revisited to illustrate the tension that can occur between performance constraints and instruction. NTool is an ICT-tool that supports complex problem solving in multidisciplinary teams by increasing a team's common ground, a cognitive frame of reference shared by all team members (Bromme, 2000). Some of the unexpected results in experiments with NTool point in the direction of unwanted, or extraneous, side-effects of the NTool support principle, which may point to defective performance constraints.

Next, we report on an exploratory study without performance constraints that builds on these outcomes, where cognitive load measurements (Sweller, Van Merriënboer, & Paas, 1998) and stimulated recall interviews were used to study both possible beneficial and adverse effects of the instruction. The cognitive load measurements were used to study whether there were differences in working memory load. The interview data were used to get an impression of the actual cognitive processes of the participants during collaboration, so as to assess whether they were germane or extraneous to the process, and whether they gave insight into the effect mechanisms of the instruction.

The NTool Support Principle for Complex Problem Solving

The NTool support principle for complex problem solving is based on three notions. First, complex (societal) problems are generally solved in multidisciplinary teams (DeTombe, 2002; Vennix, 1996). Second, problem solving teams need some common ground; a shared cognitive frame of reference (Bromme, 2000) to be able to construct a shared problem representation (Ostwald, 1996). Finally, expert problem solvers spend relatively more time on problem representation than novices, which shows that the problem-representation phase is highly important for problem solving (Lesgold et al., 1988). The negotiation of common ground can be seen as an activity that is intrinsic to solving complex problems, because common ground is needed to afford the sharing of knowledge and the subsequent construction of a shared problem representation in multidisciplinary teams.

NTool aims at supporting multidisciplinary teams by improving the negotiation of common ground. It does so by requiring problem solvers to explicitly verify their understanding of each other's contributions to a conversation and to explicitly articulate their positions on those contributions. The researchers expected that this would allow easy recognition of content areas where the participants lacked common ground, which in turn would lead to more negotiation of common ground. Ultimately, the intended learning effect of NTool is that the user learns how to successfully negotiate common ground. In terms of outcome, we thus expect more explicit negotiation of common ground and more common ground with the use of NTool in multidisciplinary teams that solve complex problems.

The NTool support principle consists of five rules for communication that mimic the process of negotiation (see Table 4.1). Participants are required to follow these rules by having to choose from a set number of message types during on-line communication. New conversation topics are introduced using a contribution message, and verified and clarified using verification and clarification messages. Furthermore, participants can use agree- and disagree-messages to make their position known to their team mates, and they can post rejections to messages that are unintelligible or objectively incorrect in the eyes of the team members.

Table 4.1
The NTool Support Principle

1.	Every new issue is termed a *contribution*
2.	Contributions require a *verification* by the other team members
3.	Each verification is responded to with *clarification* by the original contributor
4.	When all verifications are clarified, and no new verifications are performed, all team members state whether they *accept* or *reject* the statement
5.	All team members state their *position* about accepted statements

Study 1

The main aim of Study 1 (Beers et al., 2005) was the continued development of NTool. Three versions of NTool were tested. Instruction was given about how to use NTool in communication. Furthermore, different performance constraints were used to support the use of NTool. The performance constraints consisted of different levels of coercion, based on the hypothesis that the fewer degrees of freedom the user is left with when using NTool (coercion, Dillenbourg, 2002), that is, the stronger the performance constraint, the stronger the effects of NTool will be.

Method

NTool was tested in a laboratory setting in which 17 3-person multidisciplinary teams solved the complex societal problem of high school drop-out using NTool (N = 51). Participants received instructions about the formalism and about the specific type of performance constraint they were to use during problem solving before working on the task. Three different versions of NTool were used, which supported the formalism in three different ways:
Idiosyncratic. This version used explicit message types for verifying and clarifying contributions and for stating one’s opinion. Support consisted of on-screen descriptions of the various message types and their uses. Furthermore, prompts informed participants about contributions that had not yet been verified.

Scripted. This version used the same message types as the Idiosyncratic version, and also the same on-screen descriptions and prompts. However, the Scripted version also prevented the posting of certain messages at certain times. More specifically, it prohibited participants from posting messages in which they stated their opinions before all contributions had been verified. This resulted in a two-phase structure of the discussion. To end the first phase, participants were required to have verified and accepted all contribution messages. Phase 2 was aimed at sharing positions on the various contributions. Using prompts, participants were informed of which phase they were in.

Stringent. This version also used the same message types as the Idiosyncratic version, but applied even more constraints on collaboration than the scripted version did. The Stringent version allowed discussion about only one contribution at one time. Furthermore, participants were not allowed to compose reject-, agree-, and disagree-messages before the contribution had been verified. Using prompts, participants were informed as to whether they had to verify or decide on a contribution.

Six groups used the stringent version, five used the scripted version and the final six groups could use the NTool idiosyncratically. Groups were given 90 minutes to complete the task.

The instruction described how to use NTool. It consisted of a description the various message types and their uses in communication. The NTool support principle was embedded in these descriptions. The exact instructions differed only with regard to the descriptions of the performance constraints. There were no differences with regard to explanation of the support principle.

Analysis focussed on the number of messages that reflected the negotiation of common ground, and the number of messages with regulative content. Negotiation of common ground was defined as directly or indirectly requesting information about, or giving clarification on, the intended meaning of a contribution. Regulation was defined as regulating the collaboration process, which also entailed tool use, and monitoring the problem solving process. Negotiation and regulation were measured using a coding scheme developed by Beers et al. (2005) that, among other things, distinguishes between new contributions, verification of one's own understanding of a contribution, clarifying another's understanding of a contribution, and regulatory activities.

Results and Discussion

NTool was shown to be increasingly effective with increasing stringency of the embedded support principles; there was a significant association between coercion and the negotiation of common ground, $r_s(N = 17) = 0.51$, $p < .05$. This means that the fewer degrees of freedom that the participants were allowed, and thus the more the execution of the instruction was supported, the higher the effects on the negotiation of common ground. However, the results also showed that the medium coercion groups required significantly more regulation than the other groups, $U(N = 17) = 4.00$, $p < .01$ (see Table 4.2). This unexpected result is the more interesting one for the purpose of this chapter.

The high number of regulation messages in the scripted groups may have been *extraneous* to the task because the other versions influenced collaboration as expected, but did not result in additional regulation. The instructions, thus, caused an increase in activities that did not benefit the instruction's purposes. It stands to reason that this was caused by differences in performance constraints between the three versions of NTool and/or deficiencies in the instruction itself. From a design perspective, this difference is highly important; if it turns out that the performance constraints are poorly designed, the problem can be overcome by improving their design. If, however, the instruction turns out to be deficient, the question must be raised as to whether the goal of the ICT-tool can be reached at all.

Table 4.2
Negotiation of Common Ground and Regulation

	Coercion		
	Low	Medium	High
Negotiation of common ground[a]	33.7	45.2	51.6
Regulative utterances[b]	30.7	106.0	43.7
n	6	5	6

[a] Negotiation was associated with coercion.
[b] Regulation was higher in the medium coercion groups than in the high and low coercion groups.

In the case of the Idiosyncratic version, hardly any performance constraints were implemented, so that neither positive nor adverse effects were to be expected on regulation. Maintaining a discussion did not appear to be difficult with this version, as shown by the relatively low number of regulation messages. However, this does not mean that the instruction was good, because the formalism did not require any regulative activity for the discussion to be maintained. In other words, the low number of regulative activities can also be the consequence of neglect of the instruction on the part of the participants.

In the case of the Stringent version, NTool constrained collaboration quite strongly. Nonetheless, this did not put extra regulative demands on the participants. In this case, it seems that although the support restricted many options during collaboration, they may have been quite easy to learn because the process of discussing one topic reoccurred several times during one discussion. For every discussed topic, the sequence of communicative acts was kept the same, which enabled the participants to learn about the performance constraints during discussion if they had not acquired enough knowledge about them during the instruction prior to collaboration. It can thus be concluded that the performance constraints were designed well enough to prevent additional regulative activities. However, this is still no evidence for a good instruction.

For both the Stringent and the Idiosyncratic version, either the lack or the abundance of constraints may have allowed a discussion that demanded little regulative activity. However, the results so far do not preclude that the same results would have been found without instruction.

In the case of the Scripted version, NTool was not so constraining, but the participants had to keep the instruction in working memory for the entire duration of the experiment to be able to navigate the Scripted version. The effect was that participants required a great deal of regulation during the whole session to keep their discussion going. This could have been prevented with either stronger support or with better instruction about the support and the formalism beforehand.

In sum, notwithstanding the intended effects in study 1, it is unclear whether the results should be attributed to the instruction itself, the performance constraints, or a combination of these causes. Did participants in the Idiosyncratic and Stringent groups need the instruction at all? Did participants in the Scripted groups require more regulation due to badly designed scaffolds or due to deficient instruction? Apparently, outcome measures alone do not give full insight into the effects of NTool. To distinguish between the effects of the support mechanisms and the effects of the instruction, a study with only instruction, and no performance constraints whatsoever was performed.

Study 2

In Study 2 a face-to-face setting without ICT-support was used to rule out any effects of performance constraints on grounding, and to enable studying effects of instruction only, with the use of new types of measurement. A pen-and-paper approximation of the NTool support principle was used as a face-to-face carrier of the instruction. Study 2 can thus be regarded as piloting a new measurement strategy on the basis of Study 1, which was a more extensive laboratory study. Cognitive load measurements and stimulated recall interviews were used to gain insight in the working memory effects of the NTool instruction and to distinguish the intended positive (i.e., germane) effects of the instruction from its extraneous effects.

Cognitive load theory (Sweller et al., 1998) couples insights on working memory to the design of instruction. It holds that the human cognitive architecture consists of a limited working memory which interacts with an unlimited long-term memory. Working memory processes information to construct and automate long-term memory schemas. For learning, two types of cognitive load are important. *Germane load* is caused by working memory processes that lead to schema construction and automation, whereas *extraneous load* is caused by trying to understand the instruction aimed at generating germane processes, but that is not beneficial to the learning itself. Because of working memory limitations, it is important that extraneous load is minimised, and that total load does not exceed working memory limitations (Sweller et al., 1998).

In the case of ICT-tools, we assume that the working memory processes that are important to the intended learning effects and those important for completing the task can be seen as constituting germane working memory load, whereas those processes that constitute thoughts about the instruction that do not lead to the intended learning outcomes, or are not beneficial to the task, can be seen as extraneous. This means that the instruction needs to be simple so as to minimise extraneous load, and effective so as to increase germane load.

The main aim of Study 2 was to distinguish the effects of instruction from the effects of performance constraints and to gain insight in the nature and mechanisms of the instruction effects. The main research question was: Can cognitive load measurements and stimulated recall interviews complement intended outcome measures to study the effects of ICT-tool instruction?

Research focused on (1) the extraneous and germane effects of the tool in terms of cognitive load, (2) the intended outcomes of the tool (i.e., negotiation, common ground), and (3) the possible mechanisms for the effect of instruction. Cognitive load was studied quantitatively via a questionnaire (Paas, Tuovinen, Tabbers, & Van Gerven, 2003). Furthermore, stimulated recall interviews were held to gain qualitative insight in the actual working memory processes, and instruction effect mechanisms.

The instruction was expected to result in more cognitive load during collaboration, either due to an increase in germane load as a result of increased grounding activities or due to an increase in extraneous load due to increased participant attention to the instruction. The instruction was expected to result in more negotiation of common ground and more common ground. The interview data were used to qualitatively identify the actual processing activities and to distinguish whether they were germane or extraneous to the task.

Method

As in Study 1, participants worked in 3-person multidisciplinary teams solving the complex problem of high school drop-out. The same overall procedure as in Study 1 was used, but with additional measures for cognitive load. The main difference between the studies was the implementation of the NTool support principle (using ICT in Study 1 versus a pen-and-paper approximation in Study 2).

Participants

Participants were 12 undergraduate students from the fields of psychology, economics, and cultural sciences from Maastricht University. Four multidisciplinary teams were formed by assigning one student from each discipline to a triad.

Materials

The pen-and-paper version of the NTool support principle was used with a whiteboard and four coloured whiteboard markers (black, blue, red, and green) for creating an external representation during group collaboration, given to all teams. Two teams received specific instruction for using the whiteboard and the markers (Instruction condition); the other teams could use their whiteboard and markers any way they wanted (Idiosyncratic condition). The instruction required participants to write new contributions on the whiteboard (i.e., externalisation of concepts), and to react on others' new contributions by giving their own perspective on them. Participants were assigned personal coloured whiteboard markers to allow easy recognition of contributors. The execution of the instruction was not supported during collaboration.

Cognitive load measurement was done through self-report of invested mental effort on a symmetrical scale ranging from 1 (very, very low mental effort) to 9 (very, very high mental effort) (Paas et al., 2003). Mental effort refers to the cognitive capacity that is actually allocated to solve the problem and can be considered to reflect the actual cognitive load (Sweller et al., 1998).

Procedure

Each team was given 30 minutes to collaboratively analyse the drop-out problem so as to come up with a solution. The team collaborations were videotaped. Participants were required to write down their individual perception of the problem and their solution(s) both prior to (pre-test) and after (post-test) the team collaboration. Cognitive load was measured three times, namely after the pre-test, after the team collaboration, and after the post-test. Within two hours after the post-test, open and stimulated recall interviews were carried out in which participants were asked to recall their thoughts during collaboration.

Quantitative Analyses

The first measure of cognitive load, carried out after the pre-test, served as a baseline measure. No differences were expected. The second measure was carried out after group collaboration. Participants in the instruction groups were expected to report higher cognitive load, either due to more negotiation (i.e., germane cognitive load) or due to using the instruction (i.e., extraneous cognitive load). The third measure was carried out after the post-test. Again, the instruction groups were expected to report higher cognitive load. However, this time no specific instructions were given, so any difference had to be caused by differences in negotiation processes during collaboration. In other words, higher cognitive load after the post-test was assumed to reflect germane load.

Negotiation was again measured with the coding scheme developed by Beers et al. (2005) to distinguish between new contributions, verification of one's own understanding of a contribution, and clarification of another's understanding of a contribution.

Common ground was measured by comparing content overlap in pre-tests and post-tests. Content was operationalised as contributions from the group discussion that appeared in the pre-tests and post-tests. For each contribution it was determined in how many pre-tests and post-tests it showed up, after which the group mean for the average contribution was used as an operationalisation for overlap between individual representations. The higher the change in overlap between pre-test and post-test, the more common ground there was.

Qualitative Analysis of the Interview Data

The interviews were qualitatively analysed to gain insight in the participants' actual thought processes and how they related to germane and extraneous load. Analysis focussed on participants' thoughts about grounding processes, about knowledge construction, about the instruction, and their explicit thoughts about effort involved in following instructions.

In the case of grounding, thoughts about understanding each other, about each other's positions, and about perspectives on the task were seen as indicative for grounding activities, and therefore germane to the task. Thoughts about co-construction of knowledge were seen as intrinsic to the task, but not specific to grounding activities. Thoughts about the use of the white-board were seen as germane to task if they were linked to grounding activities and co-construction of knowledge. Other thoughts about the instruction and the use of the whiteboard were seen as extraneous to the task.

All reported thoughts from the interview transcripts about negotiation processes and instruction were gathered and categorised. The qualitative analysis was carried out by the author.

Results

Cognitive Load

No differences in cognitive load after the pre-tests were found (see Table 4.3). Contrary to expectations, no differences were found after group collaboration. However, after the post-tests, participants in the Instructed groups reported significantly more cognitive load, $U(N = 12) = 7.00$, $p < .05$, one-tailed, suggesting that these participants required more processing in the post-test than the participants in the Idiosyncratic groups.

Table 4.3
Cognitive Load

	Condition	
Cognitive Load measured after the...	Idiosyncratic	Instructed
... Pre-test	5.67	5.50
... Group collaboration	5.83	6.00
... Post-test[a]	5.00	6.17
n	6	6

[a] Cognitive load after the post-test was higher in the instructed groups than in the idiosyncratic groups.

Negotiation and Common Ground

No statistically significant differences were found with regard to negotiation and common ground. Contributions were most heavily negotiated in the Idiosyncratic groups (see Table 4.4), Instructed groups made more contributions.

Table 4.4
Negotiation and Other Utterances

	Condition	
	Idiosyncratic	Instructed
New contributions	8.0	13.0
Negotiation of common ground	33.0	13.5
Other task-oriented utterances	124.0	91.5
Regulative utterances	25.0	13.5
Other utterances	8.5	14.0
n	2	2

The Idiosyncratic groups achieved the most common ground after collaboration, as shown by the content overlap in post-tests, but they also started with some unexpected overlap in pre-tests, which the Instructed groups did not have (see Table 4.5). Differences in common ground were not statistically significant.

Table 4.5
Common Ground

	Condition	
Content overlap[a] in the	Idiosyncratic	Instructed
… pre-tests	1.35	1.00
… post-tests	2.01	1.41
n	2	2

[a]A content overlap of 1 is the minimum score, meaning that conversation topics were mentioned by only one person. A content overlap of 3 means that all conversation topics present pre-tests and post-tests were mentioned by all team members.

Interview Data

Grounding processes and knowledge construction. Interview data reveal that participants are aware of instances of agreement and disagreement, and also of mutual understanding and misunderstanding. In other words, they acknowledge the status of a contribution to the discussion:

> S1E[3]: "It then crosses my mind that she makes points that I have thought of myself, so I thought that's OK, so I also agreed with her."
> I1C: "But... yeah, I got that feeling, I don't think you got it, and then that turns out to be right."
> S2P: "I thought that there might be possibilities, but also that it would be difficult, the way she put it."

These excerpts show that participants are aware of the distinction between understanding and agreeing, as also shown by the following excerpt, which reflects agreement to disagree:

3 Interview excerpts are coded as follows; the first letter signifies experimental condition ("S" for instructed groups and "I" for idiosyncratic groups); the digit represents the group number (starting at "1" for each experimental condition); and the last letter signifies educational background ("C" for Cultural Sciences, "E" for Economics and Business Administration, and "P" for Psychology).

> S2C: "I think I understand what she thinks about this, and I also understand that the positions she's taking and I am taking, uh, differ."

Awareness of agreement to disagree was only mentioned by participants from the Instructed groups.

The interview data also give insight in the way participants react to perceived (mis)understanding and (dis)agreement. For instance, participants report thoughts about the background of others' contributions to the discussion, and others' educational and even philosophical backgrounds are mentioned as well. From these excerpts it seems that participants actively attribute contributions to the contributor's background.

> I1C: "With regard to the problem we're working on, how . . . um . . . why does she mention this."
> S2C: "At this moment, I'm trying to understand what the philosophical view of the other positions . . . is."
> I2P: "The difference in background that appears quite quickly. . . that he clearly has a background in economics."

Several qualitatively different reactions to differences in understanding and position become apparent from the interview data, ranging from outright neglect of mutual misunderstanding to accepting something on the basis of another's expertise:

> I1E: "I got what she meant, but well, you don't completely go like, what's your view then and bla, bla, bla, that doesn't matter too much."
> I1C: ". . . and then you hope they'll nod, like yeah, that's right what you're saying or, or, we understand, and that didn't happen."

These excerpts show that possible misunderstandings are sometimes detected, and not actively addressed. Another reaction was waiting, leaving the contributor time to explain his/her intentions, as opposed to immediately giving one's primary response:

> S1C: "I was like, wow, that's radical, and then I thought, well, maybe it has something."
> S2C: ". . . assuming that you've understood something, that... that, that what the other says is something you recognise, I'm very cautious about that."

These excerpts show that participants sometimes consciously wait for the contributor's clarification to see whether their primary (negative) reaction towards a contribution is justified. It seems that withholding one's reaction for a short time may allow for understanding that otherwise would not have emerged. Such active attention to mutual understanding was also mentioned on a more general note:

> S2C: “At that moment I thought, OK, we’re really doing our best to understand each other and get somewhere.”
> I1P: “I actually said that just to see if I had understood.”

Finally, in both conditions, the interview data show that participants actively build on each other’s knowledge, and are also capable of revising their own and each other’s ideas:

> I2P: “here I don’t wholly agree . . . so that’s why I’m putting it down just a bit differently.”
> S2C: “He takes my criticism seriously. And that he’s willing to consider it, then also continue thinking along that line.”
> S1P: “And this was more of building on each other.”
> S1E: “I thought that’s a smart move that she more or less linked her solutions and the points I was contributing at that point.”

Effects and use of whiteboard and instruction. The interview data suggest that using the whiteboard and coloured markers helped structure the discussion and keep track of individual contributions:

> S2E: “I thought it contributed to structure, the writing down.”
> I1C: “. . .to keep track of what we had, ... and keeping track of what needs to be discussed and things we might forget.”
> S2P: “I saw quite a bit of blue and red appear on the board, but hardly any green, so...”

Some interview excerpts suggest ways in which the whiteboard may have contributed to the negotiation of common ground. These excerpts show that the whiteboard takes away the necessity of immediately understanding something or agreeing on something and reacting, because the participants can refer to what is on the whiteboard later in the discussion. In other words, if something is on the whiteboard, it is easier to withhold a primary response.

> S2C: “In a discussion you have to react immediately if you want to support or challenge someone. And with a whiteboard, you can, it’s on there. So you can reconsider.”
> I1E: “Sometimes when you didn’t completely understand or had missed what someone had said then you could also see it on the whiteboard.”

Furthermore, some interviewees linked using the whiteboard to the negotiation of meaning. From these excerpts it seems that, on the one hand, the requirement to write something down requires participants to be clear and on the other hand, that the presence of something on the whiteboard facilitates access to the meaning of others’ contributions.

> S2C: “. . . and maybe that the whiteboard requires you to be a bit more explicit.”

> S1C: “I liked it this way [with the whiteboard] because this way you can see what someone really means.”

From further interview excerpts it seems that the co-presence of different contributions on the whiteboard may facilitate co-construction of knowledge:

> I1C: “And then we looked at those things that the three of has hadn’t thought of individually, but at which we arrived as a group by reading each others notes.”
> S1E: “Like I just said, you see ‘culture’ and ‘motivation’ put down [on the board] and then you combine those.”

The instructed groups were told to share individual perspectives on the whiteboard:

> S2C: “In principle this [sharing and writing down opinions] goes well, if only you’ve had a small discussion before. That you first state your opinion, like this and this is what I mean.”
> S1C: “I think it [following the instruction] worked out, if someone said something, then another would build on that or say whether or not they agreed.”
> S1P: “I thought it went quite well. Everybody had their own opinion, the only thing is, well, we actually agreed very much on many things.”

From the interview data, it can be concluded that writing down opinions on the whiteboard did not pose a problem to the participants. It also seems that the participants valued talking about those opinions so as to clarify what they meant. These excerpts, thus, suggest that requiring groups to share individual perspectives may encourage them to negotiate the meaning of individual contributions before they get written down. However, one interviewee mentioned that the whiteboard was not used as intended:

> S1E: “. . . because after that [the beginning] we only used the whiteboard to summarise our discussion.”

The interview data also suggest some areas of difficulty encountered when using the whiteboard:

> I1C: “It can slow down a discussion, because you keep, at least in my case because I was taking the notes, busy in your head with how to write something down.”
> S2E: “You talk faster than you write.”

It appears that being required to both take notes and discuss something at the same time can sometimes be taxing.

Discussion

Differences in cognitive load were found after the individual post-test, but, unexpectedly, not after group collaboration. The latter result may be explained by the interview data which showed that participants did not find using the whiteboard this way effortful, so even if discussions were different, this may have caused the same cognitive load.

In line with the expectations, cognitive load was highest when participants solved the problem individually after the group work. The instruction may have elicited more knowledge in the group work that was new for individual participants. After the group work, they may have been busy processing the others' contributions while producing an individual solution the task. This explanation points at germane effects of the instruction.

The findings on negotiation and common ground were not in the expected direction, but were statistically inconclusive. Nevertheless, the fact that negotiation was lower in the Instructed groups, while they made more contributions, may be explained by the instruction itself, that is, always stating one's perspective which may have resulted in elicitation of further contributions instead of negotiation of existing contributions. The pen-and-paper instruction did not really explicitly lay out rules that mimicked negotiation, as was the case in Study 1.

The stimulated recall data suggest a number of mechanisms for the use of the whiteboard to affect grounding processes. First, the presence of contributions on a whiteboard allows one to postpone one's reaction and wait for further clarification instead of immediately stating one's opinions. Second, the co-presence of different contributions on the whiteboard appears to allow co-construction of knowledge. Finally, the instruction to publicly write down ones perspective appears to allow the (spoken) negotiation of meaning of individual contributions, prior to actually writing them down.

General Discussion

This chapter aimed to show how cognitive load measurements and interview data might complement traditional outcome measures to study the effects of instruction itself, in the absence of scaffolding.

Study 1 illustrated that ICT-tools may cause unwanted side effects that are extraneous to the goals of those tools in terms of extra effort on the part of the participants. Although all intended effects (i.e., negotiation of common ground) were achieved as expected (i.e., more coercion resulted in more negotiation), the results did not preclude that the effects were produced by performance constraints alone.

It was hypothesised that cognitive load measurements and qualitative interview data might give insights in the effects of ICT-tool instruction. Study 2 showed that cognitive load measurements indeed can yield information about germane and extraneous cognitive load effects of ICT-tools. No differences were found after group collaboration, but the higher cognitive load in the Instructed groups after the individual post-test suggests some germane effects of the instruction. With regard to adverse effects of NTool, it can be concluded that the instruction for using the whiteboard and the markers in Study 2 does not significantly increase extraneous load compared to the groups without the instruction, and that it may have had some germane effects on common ground. In sum, the cognitive load measurements in Study 2 suggest an increase in germane load, which may have been caused by increased grounding activities although these conclusions can not be substantiated by the negotiation and common ground measurements.

Study 2 also exemplified how qualitative interview data may give insight in the effect of instruction. Notwithstanding the absence of statistically significant differences in outcome measures, the interview data suggest how the instruction might result in germane processing by allowing co-construction of knowledge and negotiation activities prior to writing down one's own perspective. This facilitates postponing one's primary reaction, which could in turn allow for consideration of another's contributions that otherwise does not take place. Such consideration is in line with the theorised effect mechanism of the NTool support principle, namely supporting the verification of another's contribution. The interview data also indicated where the instruction might have some extraneous effect, that is in dividing participants' attention between writing things down on the whiteboard and keeping up with the discussion.

As differences between the conditions were small, and as the number of groups in this study was low, more research is needed to strengthen these hypotheses. The cognitive load measurements and the stimulated recall interview lend credence to a possible germane effect of the NTool support principle, which suggests that the intended NTool effects in Study 1 can be attributed, at least in part, to instruction alone. The necessity for regulation in the Scripted (i.e., medium coercion) groups may thus be attributed mainly to defective performance constraints, and not to deficient instruction. More research on a larger scale is needed, using both intended outcome measures and other measures, to substantiate this conclusion.

In sum, it appears that both cognitive load measurements and qualitative interview data gave insights in the NTool instructions that could not have been achieved through outcome measures alone. For designers of ICT-tools, the use of outcome measures can lend credence to conclusions regarding the effectiveness of those tools. However, a thorough study of the effects of ICT-tools is needed to employ complementary measures of cognitive processes to substantiate those conclusions.

Chapter 5

Support of Grounding in a Classroom

Multidisciplinary teams are often employed to solve complex problems, but research has shown that using such teams does not guarantee arriving at good solutions. Good team-solutions require team members to achieve common ground, that is to say a shared cognitive frame of reference. In this chapter NTool, an ICT-tool based upon making individual perspectives explicit to other team members is studied. Two versions of the tool that differed in the extent to which users were coerced to adhere to embedded support principles were used in a secondary vocational education setting. Coercion, as expected, increased negotiation of common ground in both settings. However, results were contradictory with regard to the amount of common ground achieved. Overall, it can be concluded that NTool and its underlying framework affect negotiation of common ground, and that adding some coercion increases this effect. However, one should be careful with the specific task and audience before implementing NTool.

Professional organisations expect that working in multidisciplinary teams will improve problem solving. Expectations are especially high in the case of complex problem solving, because multidisciplinary teams supposedly can employ multiple problem perspectives. Indeed, research has shown that including multiple perspectives may lead to richer solutions for complex problems (Lomi et al., 1997), and that neglecting relevant perspectives can lead to solving the wrong problem, and in some cases even aggravate the problem (Hasan & Gould, 2001; Vennix, 1996). Research has also shown that individual team members have to engage each other's thinking in order for these expectations to hold (Van den Bossche, Gijselaers, Segers, & Woltjer, 2005). That means that the team members need to achieve a common cognitive frame of reference, or common ground (Bromme et al., 2001; Clark & Brennan, 1991), in order to reap the benefits of multiple problem perspectives.

The support of complex problem solving has received increasing attention from researchers and developers in the field of ICT. Specific ICT-tools have been designed to support problem solving activities like group design (Buckingham Shum et al., 1997), scientific reasoning (Suthers, 2001), and argumentation (Van Bruggen, 2003). As such they all address aspects of the structure of complex problems (based on a problem ontology). For example, Buckingham Shum et al. set out from the notion that any design problem can be structured in terms of a design question, options for answering the questions, and criteria that have to be fulfilled in order for the options to answer the question in actuality (the QOC-approach to group design). However, none of these tools have addressed the structure of the groups that solve these complex problems, and specifically the common ground they need. The present chapter presents NTool, an ICT-tool that addresses the group aspect of multiple perspectives by facilitating the negotiation of common ground.

NTool is an on-line communication tool with embedded support of grounding processes. NTool is the first to address the grounding process at a more basic level. It does so by making users explicate their private understanding of each other's contributions, an important requirement for the negotiation process. Like other ICT-tools, NTool uses specific communication rules (a formalism) and constraints (coercion) to attain this facilitation. Coercion is a means to make participants adhere more closely to the formalism so as to increase its effectiveness. However, increasing coercion also holds the risk of disrupting communication processes, when it constrains communication so much that it effectively prevents the users from carrying out their task (Dillenbourg, 2002).

Recent research by Beers et al. (2005) explored the influence of coercion on the effectiveness of NTool (low coercion, medium coercion and high coercion conditions). NTool's influence on both the grounding process and the amount of common ground was shown to increase with coercion. In other words, the more coercively the NTool formalism was applied, the more it facilitated negotiation of common ground, and the more common ground the teams achieved. However, NTool also appeared to have some adverse effects.

The medium coercion version required disproportionately more regulation than the low and high coercion versions, indicating that the participants had difficulties using this version that did not arise with the other versions. Furthermore, these experiments were done under highly regulated circumstances, with highly motivated participants, so that adverse effects on social processes may have been small. In sum, the previous laboratory studies may have obscured some effects of NTool. In this chapter we report on a study in a practical educational setting. The goal of the present study is twofold namely 1) to replicate the results of earlier studies in a setting that was more ecologically valid while controlling for social aspects of the team, and 2) to study possible adverse effects of NTool in terms of both cognitive load effects and social aspects of NTool itself.

Negotiation of Common Ground

Theory on negotiation of common ground originated in linguistics (Clark & Schaefer, 1989) and cognitivism (Bromme, 2000). The linguistic approach describes how negotiation of common ground occurs in conversation, whereas the cognitivist approach focuses on the way new knowledge is processed, the role that previous knowledge plays in this process, and how individual perspectives affect this. The combination of the two links the content of a learning process to the way it is communicated between people.

This chapter combines linguistic (Clark & Brennan, 1991) and cognitive (Bromme, 2000) approaches to the negotiation of common ground. In this conceptualisation, the grounding process starts when team members contribute their, as yet, unshared knowledge, so that others can try to comprehend that knowledge. At this point, a number of biases come into play that cause differences between the intended meaning of a contribution, and the contribution as it is understood. While constructing their own individual understanding, the other team members use their knowledge of aspects like the contributor's background and views held, and the current situation, to better "understand" the contribution. Also, their own beliefs and assumptions play a role while trying to understand a contribution. A contribution is thus always understood against the *presumed perspective* of the other and one's own perspective (Bromme, 2000). Therefore, having shared a contribution with a team does not mean that the team members all have acquired the same understanding. Representational differences can result from interpreting a contribution in one's own perspective only or from minimising or rejecting its validity or plausibility due to differences in conviction or opinion. The negotiation of common ground then is the iterative minimisation of these representational differences, through providing feedback based on one's own perspective by word or action (Alpay et al., 1998; Baker et al., 1999).

In the present chapter, negotiation of common ground is conceived of as a dual concept. The first aspect is *negotiation of meaning* which leads to an agreement regarding meaning and understanding of a contribution. This entails making one's private understanding of some contribution public to others, who in turn verify whether and to what extent their own understanding of the contribution is the same as or is different from what others intended, receiving feedback on this (clarification), re-verifying, and so on, until "the contributor and the partners mutually believe that the partners have understood what the contributor meant to a criterion sufficient for the current purpose" (Clark & Schaefer, 1989, p. 262, the grounding criterion). Negotiation of position, the second aspect, concerns people making their private opinion about a contribution public to others, checking whether one's position is clear to others, and vice versa.

The above steps serve as a basis for a formalism for the support of negotiation. The formalism consists of negotiation *primitives*, basic building blocks that model a specific type of dialogue (Dillenbourg, 2002), and *rules* that prescribe the use of these primitives. Table 5.1 summarises these rules. Together these rules and primitives mimic the negotiation process as explicitly as possible. Note that this formalism models an ideal negotiation process; in regular communication, the status of people's statements in terms of negotiation primitives often remains implicit. The formalism must enable the user to more easily distinguish between original contributions, clarifications, verifications, etcetera, thus making the negotiation process more explicit. This way, individual differences in understanding and opinion between users should more easily surface.

Table 5.1
Rules for the NTool Support Principle

1.	Every new issue is termed a *contribution*
2.	Contributions require a *verification* by the other team members
3.	Each verification is responded to with *clarification* by the original contributor
4.	When all verifications are clarified, and no new verifications are performed, all team members state whether they *accept* or *reject* the statement
5.	All team members state their *position* about accepted statements

First, negotiation starts with a *contribution* (Primitive 1) such as a hypothesis or a position, which is assumed not to be part of a team's common ground (Rule 1). To detect differences between individual representations, team members must *verify* (Primitive 2) their understanding of the contribution (Rule 2) because people articulate and understand a contribution against their own background knowledge (Fischer et al., 1995). Third, a contribution needs to be elucidated (*clarification*, Primitive 3), using the ideas upon which it was based. For example, the educational background or the political orientation of the contributor may shed light on the meaning of a contribution. A clarification need not always be made by the original contributor, but may also be performed by another team member who feels knowledgeable. Rule 3 is that all verifications require a clarification. Together, Rules 2 and 3 can be iterated until common understanding of the contribution is reached.

The fourth primitive is *acceptance/rejection* of a contribution, whether one can judge a contribution as true (acceptance), or untrue or unintelligible (rejection), based on the explanation given and one's own prior knowledge. For example, the statement 1 + 1 = 10, is true only if we understand (through Rules 1 and 2) that the contributor is using the binary system. A contribution should be accepted as part of the common ground if it is true, or after it has been modified so that it has become true. Rule 4 is that every contribution needs to be accepted or rejected by the team members. Finally, Rule 5 is that people must explicitly state their own position (*position*, Primitive 5) on the contribution. In the case of irresolvable disagreement about previously accepted statements, Rule 5 may result in multiple scenarios, each based on another position (i.e., agree to disagree). This means that one may accept a certain contribution, but disagree all the same, for example when neither person can prove the other wrong. In such cases, people can agree to disagree, and alternate representations that are equally legitimate can ensue.

Although the NTool approach to negotiation is primarily cognitive, research has shown that several team beliefs may have an important influence on the grounding process (Van den Bossche, Gijselaers, & Segers, 2004). One such factor is psychological safety, defined by Edmondson (1999) as "a shared belief held by members of a team that the team is safe for interpersonal risk taking" (p. 350). Her studies show that psychological safety affects team performance by augmenting a number of team learning behaviours, some of which are seeking feedback and sharing information. Chang and Lee (2001) reported associations between psychological safety and reflective thinking. Van den Bossche et al (2004) reported similar results for the influence of psychological safety. His results showed that psychological safety may facilitate the construction of shared knowledge, through affecting team learning behaviour. In the present study psychological safety is treated as an important covariate for the effect of NTool on negotiation.

The Negotiation Tool

NTool is based on a newsgroup reader for asynchronous, distributed, text-based discussions. To optimise NTool for negotiation of multiple representations, the formalism was implemented to structure the negotiation process in two ways with different levels of coercion (cf. Dillenbourg, 2002).

Coercion, a form of scripting, is defined as the degree of freedom participants have in following a formalism. Coercion and formalism together constitute a collaboration script. The higher the coerciveness of a script, the more the participants are required to adhere to the formalism. Scripting requires "subjects on most or all occasions to make a particular type of speech act in a specific context." (Baker & Lund, 1997 p. 176). A script that uses very little coercion leaves participants many degrees of freedom such that usage of the formalism attains a high degree of idiosyncrasy. A script with a high level of coercion constrains the number of options participants have, thus guiding them along the lines of the formalism.

In the present study, two different ICT-implementations of the formalism were implemented (see Methods). One implementation had very little coercion and was called the Idiosyncratic version. This situation resembles giving a person a set of lines and symbols to be used in constructing a diagram, but leaving it up to her/him to decide which symbols and lines are used for what purpose. The other implementation used scripts aimed at interaction and collaboration (high coercion) and was called Stringent. In each implementation, coercion was aimed at the verification and clarification primitives, that is, the extent to which people were required to verify and clarify contributions in specific circumstances.

Dillenbourg (2002) argues that there are trade-offs between coercion and team processes. On the one hand, increasing coercion may increase the effectiveness of the formalism in question. However, too much coercion may disturb interactions and increase cognitive load. In this study, these effects are measured in three ways. Measurements of the intended effects of NTool (negotiation of common ground and common ground itself) were complemented with measurements of possible adverse effects of coercion. The latter were tested both in terms of social aspects of NTool as perceived by the participants, and in terms of cognitive load.

With regard to social aspects of NTool, we use work by Kreijns (2004) on social presence and sociability in ICT-learning environments. Social presence can be defined as “the ability of learners to project themselves socially and emotionally” in a learning environment (Rourke, Anderson, Garrison, & Archer, 1999). Social presence is thought to support critical thinking processes in learning groups (Garrison, Anderson, & Archer, 2000). In the case of NTool, social presence is thus dependent on the participants’ ability to recognise and use NTool’s affordances for social interaction. Social presence is closely related to sociability, defined by Kreijns as “the extent to which the CSCL environment is able to facilitate the emergence of a social space” (p. 7), that is, a “human network of social relationships between group members” (p. 7). In the present study, both sociability and social presence are measured.

With regard to cognitive load effects, we build on the work of Van Bruggen et al. (2002) on the effects of external representations on cognitive load. It has been argued (Scaife & Rogers, 1996) that representing knowledge externally lowers working memory load, which subsequently can be used for other processing activities. However, Van Bruggen et al. (2002) found that poorly designed instruction for knowledge externalisation may increase cognitive load, due to the effort needed to apply the instruction in action.

Cognitive load theory (Sweller et al., 1998) distinguishes between g*ermane cognitive load*, which is caused by working memory processes that lead to schema construction and automation, and *extraneous load*, which is caused by understanding the instruction aimed at generating germane processes. Because of working memory limitations, it is important that extraneous load is minimised, and that total load does not exceed working memory limitations (Sweller et al., 1998). With regard to NTool, this means that high coercion could lead to an increase in extraneous cognitive load, because high coercion requires the participants to allot working memory resources to understanding the way coercion is implemented, but also in an increase in germane cognitive load, due to increased working memory allotted to verifying and clarifying contributions. The present study compares cognitive load measurements with differences in negotiation and common ground between Stringent and Idiosyncratic versions of NTool to study germane and extraneous cognitive load effects of coercion.

Hypotheses

NTool was expected to increase the negotiation of common ground because it forced team members to make their private understandings and opinions public, making differences in understanding and opinion visible or salient (Bromme, 2000). We hypothesised that (1) the higher the level of coercion, the more negotiation would occur. Likewise, we hypothesised that (2) common ground would be highest in the Stringent version and lowest in the Idiosyncratic version. Both hypotheses presume that more coercion will make participants follow more closely an ideal model of negotiation, as laid down in the formalism.

Furthermore, the Stringent version was expected to result in more cognitive load during collaboration, either due to an increase in germane load as a result of increased grounding activities or due to an increase in extraneous load due to increased participant attention to the instruction. Also, we expected this difference to hold after collaboration, while participants performed an individual task similar to the collaboration task. In this case however, a difference was expected to be due to reorganisation of schemas caused by the acquisition of new, relevant knowledge from the other group members.

Finally, psychological safety was expected to correlate positively with negotiation of common ground.

Method

Participants

Participants were 66 second year students (age 17.9 years, SD = 1.22) of senior secondary vocational education from three different education programmes: High-tech Metal-Electrics, Infrastructure, and Architecture. Participants were assigned to 22 three-person multidisciplinary teams.

Materials

Task

Participants were assigned the task to make a functional design of floating housing as a remedy for sea-level rise. The task was designed in collaboration with executives from the educational institution to ensure that the difficulty level was appropriate for the level of expertise of the participants, and that the task was interesting to all different educational programmes. The task was split in two parts. In the first part (“Floating Houses”) the participants were to design a floating house. In the second part (“Amersfoort-by-the-Sea”) additional information was given to keep the participants going.

NTool

Each team was supplied with three computers running NTool, one for each participant. Two different collaboration scripts were used.
Idiosyncratic. This version used all primitives. On-screen information was presented about every contribution, and whether it needed yet to be verified or

decided upon (agreeing or disagreeing). Furthermore, each participant was informed when he/she had not yet verified all contributions, and when he/she had not yet decided on all contributions. Participants were free to choose to what extent they verified and rejected, agreed, or disagreed upon the contributions in the discussion. No coercive rules were used to require this.

Stringent. This version used the same primitives as the idiosyncratic version, but used coercion to allow negotiation of only one contribution at one time. Furthermore, participants were not allowed to compose reject-, agree-, and disagree-messages before the contribution had been verified. Using prompts, participants were informed as to whether they had to verify or decide on a contribution.

Eleven groups used the Stringent script and eleven groups used the NTool Idiosyncratically.

Questionnaires. All questionnaires used in this study were adopted from other researchers who had been able to use them reliably. The reliabilities reported here were computed using the data from the current study. Group means of psychological safety (Cochran's $\alpha = .69$) were used as a covariate, to control for differences in common ground and negotiation (Van den Bossche et al., 2004). Social aspects of NTool were measured using scales (Kreijns, 2004) for sociability ($\alpha = .86$) and social presence ($\alpha = .86$). Cognitive load measurement was done through self-report of invested mental effort on a symmetrical scale ranging from 1 (very, very low mental effort) to 9 (very, very high mental effort) (Paas et al., 2003). Mental effort refers to the cognitive capacity that is actually allocated to solve the problem and can be considered to reflect the actual cognitive load (Sweller et al., 1998).

Procedure

Practice phase. First the participants received a tutorial that addressed the basics of NTool and then proceeded to emphasise the rules of the formalism, and the way they constrained communication. To ensure that participants were proficient with the NTool a practice case (solving a road traffic safety problem) was used to enable them to gain experience with the NTool. The practice phase lasted 75 minutes.

Experimental phase. The experimental phase was divided in two identical sessions (a morning session and an afternoon session), with a 75-minute lunch-break in between. After a 15-minute coffee break, participants started working on the experimental (sea-level rise) case. To promote the construction of an individual problem representation, as well as to allow the researchers to determine what this representation was, participants first had to carry out the task individually (pre-test, 15 minutes). Participants could take notes while working individually on the task. Next, they solved the problem collaboratively (60 minutes), and after that individually again (post-test, 15 minutes). All resulting individual problem representations and solutions, as well as the group discussion were recorded. In their post-test, participants were also asked to state the points on which they felt that they had differences in opinion with their team members, to account for agreeing to disagree.

Cognitive load was measured after each pre-test, each collaboration task, and each post-test.

Analysis

Negotiation. Negotiation was operationalised in two ways, that is, the number of explicit negotiation activities, and the amount of negotiation per conversation topic. A coding scheme for coding function and content of messages during collaboration was developed (cf. e.g., Avouris et al., 2003; Fischer et al., 2002; Mulder et al., 2002; Thomas et al., 1982). All messages were coded with regard to:

- Cognitive content: Directly related to solving the problem.
- Regulative content: Related of the task.
- Other content: Not in any other category.
- Nonsense: Uninterpretable messages.

Messages with cognitive content were specifically coded for function. The following subcategories were used to code negotiation:

- Contribution: A new topic of conversation that has not been discussed before is introduced.
- Verification: Information is directly or indirectly requested about the intended meaning of a contribution or elaboration.
- Clarification: A reaction to a verification or a perceived lack of understanding, in which the intended meaning of a contribution or elaboration is elucidated.
- Acceptance: A reaction to a contribution in which the contribution is judged intelligible and/or correct.
- Rejection: A reaction to a contribution in which the contribution is judged unintelligible and/or incorrect.
- Agreement: A reaction to a contribution in which the sender voices his/her agreement with the contribution.
- Disagreement: A reaction to a contribution in which the sender voices his/her disagreement with the contribution.

In many cases, messages did not fit any of the above subcategories, for example if people built on each other's communications, without explicitly negotiating meaning of, or position on, a contribution. Such messages were coded Elaboration: A contribution is elaborated upon by adding information or summarising. Verification and clarification, in contrast to elaboration, were considered indicative for explicit negotiation activities. The total number of contributions discussed was used as an indicator for the range of topics discussed. Furthermore, messages with regulative content that addressed the aimed to monitor the problem solving-process were (coded Monitoring) were distinguished from those that only regulated the conversation (coded Regulation).

A research-assistant was trained for 40 hours to use the coding scheme. Comparing three randomly selected experimental session coded by the author and the research-assistant resulted in a substantial (Landis & Koch, 1977) inter-rater reliability (Cohen's kappa) of .73 (*SE* = .024). All data were coded by the research-assistant.

To measure the number of verifications and clarifications per conversation topic, episodes in the discussion that dealt with one conversation topic were first identified. The contributions identified with the coding scheme for negotiation were considered starting points for a new discussion episode. An episode generally started with a contribution and ended when one of the participants would make a new contribution, and all of the discussion in between these contributions dealt with one conversation topic. For each group, negotiation per conversation topic was then calculated by dividing the sum of all clarifications and verifications by the number of contributions.

Common Ground. Common ground was conceptualised as the degree of overlap in individual representations after collaboration, in terms of contributions. For each of the episodes identified earlier, we assessed its presence in the various individual representations of the participants. The mean number of individual representations that an average episode would end up in was used as a measure for common ground.

Cognitive Load. Pre-tests were expected to be the same for both versions of NTool. Group tests were expected to differ between conditions; the Stringent version of NTool was expected to result in more cognitive load, either due to an increase in germane load as a result of increased grounding activities, or due to an increase in extraneous load due to required attention to the NTool formalism. Post-tests were also expected to differ between conditions, participants in the Stringent groups were expected to report an increase in cognitive load because they have acquired more knowledge from the other participants due to increased grounding activities.

Statistical analyses. Negotiation and common ground were analysed using repeated measures ANCOVA. Psychological safety was included as a covariate. In the case of significant main effects from condition, but non-significant main or interaction effects of psychological safety, analysis was repeated without psychological safety. All analyses were performed with SPSS version 11. Due to participant drop-out after the first experimental phase, the number of groups used in the statistical analyses was lower than 22. Data for 9 groups in the Idiosyncratic, and 5 in the Stringent conditions were eligible for statistical analysis. Significant effects of phase I on phase II (effects of time) were not considered relevant to our hypotheses, and are not reported here.

Cognitive load measurements were analysed using nested repeated measures ANOVA. The first level consisted of cognitive load measurements after the pre-test, after collaboration, and after the post-test. As these measurements were done repeatedly, once in the morning and once in the afternoon, session was included as a second level in the analysis. As we expected possible differences after collaboration and after the post-test, but not after the pre-test, we expected to find an interaction between condition and cognitive load. We expected possible main effects of condition on cognitive load after collaboration and after the post-test.

Table 5.2
Negotiation and Common Ground

	Condition							
	Idiosyncratic (*n* = 9)				Stringent (*n* = 5)			
	Morning		Afternoon		Morning		Afternoon	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Contribution	6.22	1.99	4.89	1.90	7.20	2.28	3.80	1.64
Verification	1.22	1.30	1.67	1.80	4.80	2.28	2.00	2.83
Clarification	1.89	1.97	.78	.67	3.60	1.67	2.40	3.21
Elaboration	27.44	16.36	27.11	14.43	46.00	23.05	20.40	18.15
Acceptance	1.56	1.13	1.33	1.32	2.00	1.58	2.60	2.41
Rejection	.44	1.01	.00	.00	.80	1.10	.20	.45
Agreement	3.33	2.78	4.44	2.96	3.80	3.27	2.80	2.39
Disagreement	.44	.73	.44	.73	1.40	.89	.80	.84
Regulation	24.44	17.51	22.78	19.85	43.00	20.24	44.13	16.36
Monitoring	2.78	3.35	2.22	2.33	3.88	2.80	2.00	1.85
Other	27.67	31.67	48.11	36.81	71.63	42.61	109.88	54.96
Nonsense	1.22	2.54	22.22	62.99	2.00	2.43	4.63	3.89
Negotiation per Contribution	.52	.51	.54	.50	1.31	.86	1.03	.72
Common Ground	1.90	.65	2.02	.61	1.41	.39	1.55	.41

Results

Negotiation and Common Ground

Repeated measures ANOVA tests revealed a significant interaction between session time and coercion on verification $F(1, 11) = 8.12$, $p < .05$, and also main effects of both psychological safety, $F(1, 11) = 5.83$, $p < .05$, and condition, $F(1, 11) = 5.61$, $p < .05$, on verification. In both sessions, Stringent teams made more verifications than Idiosyncratic teams. Inspection of Table 5.2 shows that this effect is significantly stronger in the morning sessions than in the afternoon sessions. Main effects of psychological safety and coercion on clarification were observed as well, $F(1, 11) = 6.67$, $p < .05$ and $F(1, 11) = 6.17$, $p < .05$, respectively. The Stringent groups featured more clarifications than the Idiosyncratic groups, and psychological safety positively affected both verification and clarification. Furthermore, there was a significant positive main effect of psychological safety on agreement, $F(1, 11) = 7.35$, $p < .05$. A non-significant, but notable effect was found for coercion on regulation, $F(1, 11) = 4.56$, $p = .06$, which is mentioned here because it may assist in explaining the unexpected results. Regulation was highest in the Stringent groups. Finally, there was a significant main effect of coercion on ‘other’ communication, $F(1, 11) = 5.60$, $p < .05$; eliminating psychological safety from the model still resulted in a significant main effect of coercion $F(1, 12) = 6.05$, $p < .05$, which means that ‘other’ communication occurred most in the Stringent groups.

We found significant main effects of both psychological safety and coercion on negotiation of meaning per contribution, $F(1, 11) = 9.34$, $p < .05$ and $F(1, 11) = 8.17$, $p < .05$ respectively. Negotiation per contribution was highest in the Stringent groups, and increased with psychological safety. Finally, there was a significant main effect of coercion on common ground, $F(1, 11) = 9.78$, $p < .01$, but not in the expected direction. Eliminating psychological safety from the model still resulted in a significant main effect of coercion $F(1, 12) = 8.46$, $p < .05$. Common ground was highest in the Idiosyncratic groups.

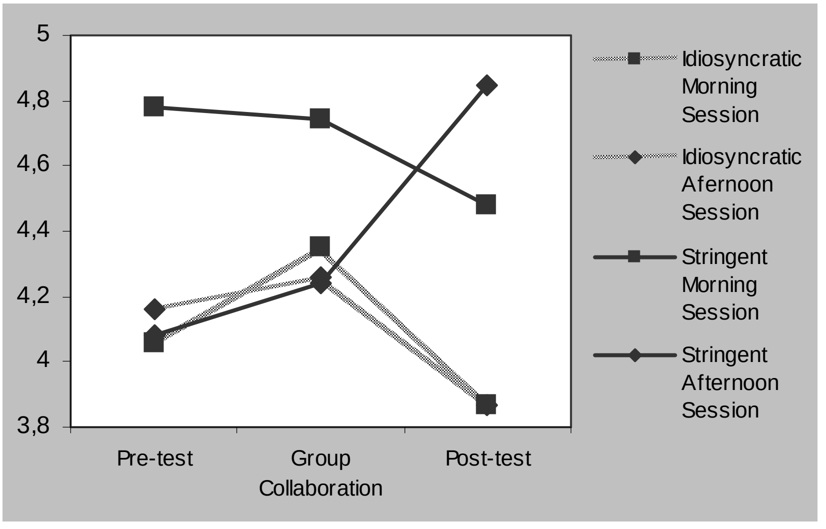

Figure 5.1. Cognitive Load Measurements

Cognitive Load

ANOVA of nested repeated measures ANOVA (see Table 5.3) revealed significant interactions of session time, cognitive load, and coercion, $F(1, 52) = 7.67$, $p < .01$, and of cognitive load and session, $F(1, 52) = 4.54$, $p < .05$, which shows that there were significant differences between the three cognitive load measurements (pre-test, group work, post-test), and that these differences were specific for session time (morning or afternoon) and coercion. Judging from Figure 5.1, it seems that cognitive load measurements in the Idiosyncratic groups follow the same pattern in both the morning and afternoon session, being lowest during the post-test. The same holds for the Stringent morning session, although cognitive load in this session is higher during all tasks (pre-test, group collaboration, and post-test) than in either of the Idiosyncratic sessions. However, In the afternoon Stringent session cognitive load on the post-test is higher than cognitive load after the pre-test and after group collaboration, whereas in the Stringent morning session as well as in both Idiosyncratic sessions it is lower. Furthermore, cognitive load measurements in the Stringent morning session seem to be higher than cognitive load measurements from the other sessions. Finally, we found a marginally significant main effect of coercion on cognitive load, $F(1, 52) = 3,67$, $p = .06$, suggesting that cognitive load was highest in the Stringent groups.

Table 5.3
Cognitive Load

	Condition					
	Idiosyncratic			Stringent		
Cognitive Load measured after....	*M*	*SD*	*n*	*M*	*SD*	*n*
... the practice task	3.36	1.56	33	3.88	1.85	33
... the morning Pre-test	4.06	1.55	31	4.78	1.25	27
... the morning Group collaboration	4.35	1.64	31	4.74	1.29	27
... the morning Post-test	3.87	1.38	31	4.48	1.05	27
... the afternoon Pre-test	4.16	1.37	31	4.08	1.94	25
... the afternoon Group collaboration	4.26	1.24	31	4.24	1.48	25
... the afternoon Post-test	3.87	1.36	30	4.85	1.46	26

Note. Due to missing values, some of the degrees of freedom for the *F*-statistics in the text may not add up to $n - 1$.

In sum, it appears that (1) in all morning sessions cognitive load was lower during the post-test than during the pre-test or group collaboration; that (2) in the afternoon sessions cognitive load during the post-test was higher in the Stringent groups than in the Idiosyncratic groups; and that (3) cognitive load overall was higher in the Stringent groups than in the Idiosyncratic groups.

Social Aspects of NTool

No significant differences were found with respect to social aspects of NTool (see Table 5.4).

Table 5.4
Social Aspects of NTool

	Condition					
	Idiosyncratic			Stringent		
	M	*SD*	*n*	*M*	*SD*	*n*
Sociability	3.29	.75	31	3.42	.60	26
Social Presence	1.91	.88	30	2.28	.74	26

Discussion

This chapter reported on the relationship between negotiation, the negotiation formalism, and coercion. The results showed that, as expected, the Stringent version of NTool increased negotiation activities, both with regard to the number of verifications and clarifications, and the amount of negotiation per contribution, although the effect of coercion on verification was significantly stronger in the morning sessions than in the afternoon sessions. This means that, as in our previous study with NTool, high coercion increases negotiation activities during interaction. However, contrary to our expectations, and also to previous findings, common ground was highest in the Idiosyncratic groups. It seems that high coercion did have the expected influence on interaction, but that this change in interaction did not result in associated effects on common ground.

Disruption of collaboration (Dillenbourg, 2002), which can be caused by over-scripting collaboration, may explain some of the results. Two unexpected differences in this study that did not occur in our previous study are quite telling in this respect. The first is the marginally significant difference in regulation activities and the second is the difference in ‘other’ communication. Both figures are highest in the Stringent groups. The difference in regulation might show that using NTool was quite taxing for the Stringent version of NTool. The difference in ‘other’ communication may signal a lack of motivation for the task at hand, that may have been a side-effect of the need for regulation.

Furthermore, the nature of the specific participant population is circumstantial to these results. In a previous study of the influence of coercion on the effectiveness of NTool (Beers et al., 2005) senior college students did not show more regulation with high coercion. Secondary vocational education institutions, however, generally draw a student population that is in many respects different from senior college students. On average, the participants of the current study were about four years younger, and therefore probably had less domain knowledge, and were less serious than the college students. Also, they had little prior experiences working in groups. This may explain why these participants needed more regulation. The high occurrence of ‘other’ communication with the high-coercion NTool may in fact signal that the participants were distracted or frustrated. Explanations involving cognitive overloading and a lack of social presence and sociability of the Stringent NTool can both safely be ruled out. The cognitive load measurements do not suggest increased cognitive load during group work, and no significant differences were observed with regard to sociability and social presence.

We expected and found significant differences in cognitive load; it appeared that cognitive load was lower during the post-test than during the pre-test or group collaboration, except during the afternoon sessions of the Stringent groups, in which cognitive load during the post-test was higher. Generally, cognitive load overall was higher in the Stringent groups than in the Idiosyncratic groups. This means that the results of cognitive load seem to contradict the results of common ground, because increased cognitive load after the post-test was expected to result from processing caused by grounding processes, whereas the amount of common ground is actually lower in the Stringent groups. Two explanations might account for these results. First, common ground being highest in the Idiosyncratic groups may actually indicate that the participants from Idiosyncratic groups may have performed processing during collaboration that participants from the Stringent groups could not do before the post-tests. Overall, these results would then indicate that NTool indeed affected negotiation as expected, but that this influence was limited to surface aspects of interaction, and did not co-occur with increased processing of other's knowledge. In other words, the participants may have been busy wording their messages as verifications and clarifications, instead of being busy actually verifying their understanding, and clarifying their intentions. If this was the case, it can be concluded that implementation of NTool in secondary vocational education needs to be accompanied with good training so as to minimise any communication difficulties arising from coercion.

Alternatively, it could be the case that participants from the Stringent groups did indeed pick up the contributions from the others, but needed to reorganise their own knowledge as a result of that. Reorganisation of one's own knowledge would account for higher cognitive load in the Stringent groups, while it also shows why they would limit themselves more to their own knowledge in their post-tests; they may have lacked the time to both properly reorganise their own knowledge *and* integrate the knowledge of the others. This would support the hypothesis that other knowledge can only be integrated *after* one's own knowledge is reorganised so as to enable an interface between one's own perspective and the others' knowledge (Boshuizen & Tabachneck-Schijf, 1998). Mentioning less of the others' contributions in one's post-test would accord with this explanation, as deeper processing would initially confront participants with a lack of understanding that would keep them from mentioning their colleagues' contributions. Educational differences could account for the fact that this was not found in previous laboratory studies with senior year college students (Beers et al., 2005).

With regard to psychological safety, our results are similar to the results of other researchers (Chang & Lee, 2001; Edmondson, 1999; Van den Bossche et al., 2004). Psychological safety positively affected both verification and clarification, as well as negotiation per contribution. This means that a sense of being safe within a team affects the extend to which people explicitly verify their understanding. Possibly, making a verification gives a sense of risk-taking because the one doing feels s/he reveals a lack of knowledge. Making a verification thus can constitute taking a personal risk.

The results are promising with regard to the facilitation of the grounding process, but they also indicate limitations in the applicability of such facilitation. In her study, Barron (2003) showed that interaction is important for problem solving, and that engaging in each other's thinking was related to better solutions. The present study has shown that ICT-tools can be used to facilitate such interactions, by using a formalism for negotiation, and coercing the user into following it. Furthermore, it has shown that NTool can be adapted for different populations, ranging from students of secondary vocational education to senior level college students. However, the results obtained in the laboratory (Beers et al., 2005) were more promising than the current results. The ultimate implementation of a tool like NTool should therefore carefully be weighed against the expected benefits, and the capacities of the intended audience. More research is needed to obtain guidelines for tweaking NTool so that it is better adapted to secondary vocational education.

More research is required to test our ultimate aim of facilitating complex problem solving. The present study does argue a relation between common ground and the quality of problem solutions, but does not explicitly measure it. Overall, it can be concluded that NTool and its underlying framework affect negotiation of common ground, and that adding some coercion increases this effect. However, one should be careful with the specific task and audience before implementing NTool.

Chapter 6

The Analysis of Negotiation of Common Ground in CSCL

The recent growth in CSCL research has given rise to a plethora of analysis methods, all with specific analysis goals, specific units of analysis, and made for specific types of data (chat, threaded discussions, etcetera). This chapter describes the development of a new analysis method, with the ultimate aim of drawing general guidelines for content analysis in CSCL. Special attention is paid to choices made and changes in those choices through the course of development of the analysis scheme, and to its underlying assumptions. The analysis scheme reported on was developed as part of a research project about "Knowledge sharing and knowledge building in expert teams with ICT". This project involves the development of support for negotiation of common ground. The results show that the main challenge in analysing negotiation involved achieving reliability without compromising the original operationalisation of the research question, and that results with the same coding scheme across different research settings can only be meaningfully compared with utmost care.

In the course of carrying out this research, a number of methodological questions arose while dealing with the reliability and validity of analysis methods and codes on the one hand, and comparing results across different studies on the other. This chapter attempts to discuss these questions and present our answers to them. The intended audience of this chapter is the community of researchers within the field of Computer-Supported Collaborative Learning (CSCL), as our use of content analysis methods is derived from it, and our experiences bear most relevance to this field.

The recent growth in CSCL research has given rise to a plethora of methods for the analysis of communication processes. However, these methods are usually poorly documented, even though their development is often fraught with numerous choices and problems. Every CSCL analysis method uses its own specific unit of analysis, has its own specific analysis goals, and is made for a specific data type (chat, threaded discussions, etcetera). This chapter draws on literature to describe some of the challenges in CSCL-analysis, and uses the development of one specific analysis scheme to explore how these challenges occur in practice.

Research on CSCL is characterised by two tendencies, namely gaining insight into how learning takes place on the one hand, and furthering the design and development of CSCL-environments on the other. With regard to the former, CSCL-researchers tend to complement measures of learning outcomes with measures of the communication process (content analysis) to gain insight in the way learning takes place (Chi, 1997; Strijbos, Martens, Prins, & Jochems, in press). The difficulty with doing content analyses of communication stems from a lack of guidelines for performing them validly and reliably (Rourke & Anderson, 2004; Strijbos et al., in press). This is to say that current know-how of content analysis remains largely undescribed and implicit, as scientific publications generally are too short to fully explain the procedures employed during analysis. The main goal of this Chapter is to make some of these procedures explicit.

The development of ICT-tools for learning is concerned with a great variety of settings which differ with respect to synchronicity, task length (which may range from an afternoon to a couple of months), task interdependence (the extent to which learners are dependent on their peers for executing the task), etcetera. Due to this variety, development often concerns a step-by-step procedure, covering the full range of face-to-face prototypes to practical (distributed, asynchronous) implementation (Kirschner et al., 2004; Nieveen, 1999). The associated analyses must enable meaningful comparisons across different studies and contexts so as to facilitate development and to enable drawing conclusions about the suitability for different learning contexts.

The challenge of analysis in CSCL thus concerns developing content analysis methods that are both structured enough to achieve reliability and validity (with regard to the former tendency), and flexible enough to allow meaningful comparisons between different experimental and developmental contexts (with regard to the latter tendency). The issue of validity and reliability concerns the question of how to arrive at a ‘good’ (i.e., valid, reliable) coding scheme for content analysis, starting from theory and observation. This issue has received quite some attention in recent literature. For example, Rourke, Anderson, Garrison, and Archer (2001) presented a set of guidelines to achieve validity by identifying fundamental issues such as objectivity, reliability, replicability, and systematic consistency. Furthermore, Strijbos (2004) showed how different treatments of segmenting CSCL-data can influence reliability by radically separating the segmentation procedure from the actual coding of the data. Nonetheless, little is known about the actual validity and reliability problems that emerge in practice throughout the development of a new analysis coding scheme.

The same holds true for the issue of comparing results in different research contexts such as face-to-face communication and computer-mediated communication. Literature shows that more traditional analysis methods using word counts and/or message counts as communication measures differ fundamentally between face-to-face and computer-mediated contexts (Hillman, 1999). Other researchers have focussed on meaning-oriented units of data-analysis (Suthers, Girardeau, & Hundhausen, 2003), but always with many precautions before actually starting a comparison. However, at present there are no known guidelines or accepted practices for comparing data from face-to-face with computer-mediated communication. Still, the typical developmental trajectory of CSCL often necessitates such comparisons.

This chapter reports on the experiences with content analysis as part of a research project about "Knowledge sharing and knowledge building in expert teams with ICT". The project involves developing support for negotiation of common ground. The associated analysis problem was how to qualitatively measure negotiation of common ground in communication processes in both face-to-face situations and computer-mediated threaded discussions in a quantifiable manner, and doing this both reliably and validly. Specifically, we encountered some difficulties in achieving reliability for our coding scheme, and we found that we could not use the raw results from our content analyses for comparisons between our studies.

The purpose of this chapter is to describe a practical example of content-analysis, and to draw some general guidelines for such analyses. The main questions are: "How can negotiation be validly operationalised in terms of the communication?", and: "How can the resulting coding scheme be reliably applied to different research contexts without compromising validity?"

First we describe some challenges of content analysis in more detail, both in terms of validity and reliability and meaningful comparisons between studies. Then we discuss how these issues were dealt with in our own studies. Finally, some guidelines for CSCL-analysis are presented in the Discussion.

Validity and Reliability in Content Analysis

As stated earlier, CSCL-research is classifiable into two types, namely research for determining the amount and type of learning and research for developing environments. In the first case, CSCL-environments generally use specific rules and constraints for communication (a formalism), either as a result of planning on the part of the developers (i.e., we want certain communication), or as an artefact of the system or software used (e.g., a text-based system does not allow for diagrams or flowcharts). A formalism is often based on educational theories that describe how it can be used to solve a learning problem. For instance, different formalisms have been developed to support such diverse processes as group design (Buckingham Shum et al., 1997), scientific reasoning (Suthers, 2001), and argumentation (Van Bruggen, 2003). Each specific educational goal requires an associated analysis method, as the intended effects of the formalisms will influence communication in specific ways. This also means that existing analysis methods cannot be used for analysing data from new CSCL-environments. In other words, new CSCL development-projects will generally require new coding schemes for analysis.

Several researchers (e.g. Chi, 1997; Rourke & Anderson, 2004) have described which steps they took in developing a CSCL *coding scheme* (see Table 6.1). Development starts with identifying the main analysis goal, which in our case concerns the negotiation of common ground, and defining the codes that validly represent the construct under analysis. The approach differs for exploratory studies, in which case an open coding approach is used to derive codes from the data (cf. analysis techniques like phenomenography, Marton, 1981; grounded theory, Strauss, 1987), and hypothesis-testing studies, which use theory as a basis for a coding scheme. The next step involves identifying the aspects of communication that typify each code. Segmentation of data is an important aspect of this, because the grain-size of data segments needs to be in line with the level at which the typifying communication aspects occur. For example, if one would want to code instances of successful negotiation episodes, it is of no use to take the sentence or a 'turn' as a segment, because a successful negotiation episode occurs across a number of conversation turns taken by different participants. In this example, the sentence and the turn level are too fine-grained to identify the construct of negotiation episodes. In practice, some researchers completely separate segmentation from coding (Strijbos et al., in press) whereas others use the code definitions only and leave it to the coders to identify appropriately sized segments.

Table 6.1
Steps for the Development of a Coding Scheme

1.	Definition of the codes
2.	Identification of code-specific communication characteristics and segmentation
3.	Development of coding rules to achieve reliability

The codes and their identifiers being identified, the next step is making *rules* for the actual administration of the codes. Although theoretically a segmented data set with a description of the codes and identifiers should result in a reliable coding scheme, the reality of CSCL analysis is often different. A widely used operationalisation of reliability is Cohen's kappa, which represents the degree of agreement between two coders after correction for chance agreement. Opinions vary about what a 'good' reliability is. According to Landis and Koch (1977), a kappa between .60 and .79 represents 'substantial' reliability, and a kappa over .80 is almost perfect. According to Capozzoli, McSweeney, and Sinha (1999), a kappa between .40 and .75 represents 'fair to good' agreement, and one above .75 'excellent'. Furthermore, some authors argue for some lenience as to a criterion for 'good' reliability, especially in the case of coding highly latent content, because aiming for reliability can sometimes hinder validity (Krippendorf, 1980).

Achieving reliability greatly depends on the type of construct that is being coded (Rourke et al., 2001). If the construct concerns manifest content, which can be coded with little need for subjective interpretation, achieving high reliability is relatively easy. On the other hand, coding latent content, which cannot be read from surface aspects of the data and which requires subjective interpretation, will require elaborate guidelines for scoring and interpretation. In other words, the extent to which data are subjectively ambiguous with respect to the codes influences the reliability of the coding scheme.

A coding scheme can be considered complete when it validly represents the construct of interest, and when coders can reliably apply the codes in practice. The challenge for the researcher thus concerns the development of such rules for the administration of codes such that they lead to reliable results, without compromising the original meanings of the code definitions.

Meaningful Comparisons Between Studies and Contexts

The second type of research involves the use of prototyping approaches (Nieveen, 1999) or interaction design (Kirschner et al., 2004) to the design of CSCL-environments, ranging from pen-and-paper versions for use in face-to-face situations, to full working prototypes for practical experimentation. The main advantage to such an approach is that findings from pilot-studies can be used to inform future design decisions. For example, version 1.0 of Belvedere, a CSCL-environment that scaffolds scientific argumentation in the classroom, differs substantially from version 2.1 (Suthers, 1999), the latter version being a complete redesign of the former. Suthers found that discussions in version 1.0 centred on how to use the formalism instead of actually using it. He used this finding in further development, resulting in a more simple design of version 2.1. For analysis, this means that a coding scheme must yield findings that can be compared across different studies, even if they differ as much as face-to-face and computer-mediated communication do. In other words, a coding scheme for CSCL-analysis must allow for a meaningful comparison of results across different studies.

Perhaps the most salient aspect of comparing CSCL-analyses is the difference between face-to-face data and computer-mediated discussion. As Hillman (1999) puts it: "due to the differing nature of face-to-face and computer-mediated interactions, coding systems used for synchronous communication are inappropriate for analysing asynchronous communication" (p. 37-38). Time-constraints are much stricter for face-to-face communication than for asynchronous communication, which means that the potential for communication is much higher in the asynchronous case. Furthermore, turn-taking in face-to-face communication generally follows the initiation-response-feedback model, whereas message sequences in computer-mediated communication are far less structured through time (Hillman). This means that one has to be extremely careful with comparisons of raw counts of codes between studies. In his article, Hillman proposes that this can be done by using *ratios* instead of counts of codes and using the unit of meaning for segmentation, an approach also used by Van Bruggen (2003). However, this only works on the assumption that the types of communication do not differ between face-to-face and computer-mediated communication. Since using a computer-tool will generally prompt communication about the use of such a tool (see for example Mulder et al., 2002, who use a specific coding category for such communication) while this will not be the case in the face-to-face situation, this assumption cannot hold.

A good example of such care is shown in the research of Suthers et al. (2003), who focus on the use and occurrence of *deixis* (the gestural referral to external representations in communication) in face-to-face and computer-mediated communication. They start out with restricting their analysis to on-task communication, and proceed with excluding communication that consisted of reciting of task materials. Furthermore, they corroborate their content analysis with qualitative analysis of the communication, to exemplify deixis. However, this careful comparison of face-to-face and computer-mediated communication does not attend to issues specific to the comparison of the effects of an intervention in different contexts.

The aforementioned shows that the actual comparability of a coding scheme across settings depends on the scheme itself, for example the segmentation procedure, and also on the way other analyses are used to corroborate the findings with a coding scheme.

In the case of our own coding scheme, we did not know of any valid coding schemes for the analysis of negotiation of common ground. No existing coding schemes for analysing negotiation were found, thus we needed to develop a new one. In the following sections we will describe how we dealt with issues of reliability, validity, and comparability of multiple analyses throughout developing our own coding scheme.

Validity and Reliability

Definition of the Codes

The analysis goal was to identify those aspects of communication that dealt with the negotiation of common ground. To that end we used the same theoretical framework that was used for the formalism. This strategy, using the same theory for both the design of CSCL and the analysis of its effects, is also described by Chi (1997). The main advantage of this approach is the difference between theory and formalism. A formalism can be seen as a set of rules and objects that together model conversation (Dillenbourg, 2002). Thus the definition of a CSCL-formalism requires the translation of (educational) theory to actual communication. In the case of negotiation of common ground, this approach to defining the codes is fairly straightforward, as theory of negotiation originated in linguistics (Clark & Schaefer, 1989) and cognitivism (Bromme, 2000). The linguistic approach describes what negotiation of common ground looks like in conversation, whereas the cognitivist approach focuses on the way new knowledge is processed, and the role that previous knowledge plays in this process. The combination of the two links the content of a learning process to the way it is communicated between people. Consequently, the step from a theoretical framework of negotiation of common ground to a formalism for the support of negotiation is quite obvious.

The above procedure resulted in the following codes for cognitive conversation content (directly related to solving the problem):

- Contribution: A new topic of conversation that has not been discussed before is introduced.
- Verification: Information is directly or indirectly requested about the intended meaning of a contribution or elaboration.
- Clarification: A reaction to a verification or a perceived lack of understanding, in which the intended meaning of a contribution or elaboration is elucidated.
- Acceptance: A reaction to a contribution in which the contribution is judged intelligible and/or correct.
- Rejection: A reaction to a contribution in which the contribution is judged unintelligible and/or incorrect.
- Agreement: A reaction to a contribution in which the sender voices his/her agreement with the contribution.
- Disagreement: A reaction to a contribution in which the sender voices his/her disagreement with the contribution.

Note that these subcategories are actually descriptions of the *function* of a turn in conversation. This combination of content-coding and functional coding has been detailed by Fischer et al. (2002).

Identification of Communicative Identifiers and Segmentation

Identification of negotiation was made possible using the above code definitions for cognitive conversation content. However, in many cases the communication did not fit any of the above subcategories. This necessitated the development of further codes. First of all, people building on each other (cognitive content) without explicitly negotiating common ground was coded *Elaboration*: A contribution is elaborated upon by adding information or summarising.

An example of an elaboration and the associated judgements to be made is the following: “That’s a good idea, but of course there are exceptions.” In the example there is no apparent questioning of one’s own understanding (i.e., it is not a verification) or clarifying of an original contribution (i.e., it is not a clarification), there’s no explicit agreeing or disagreeing (of course “good idea” does show a positive attitude, but “exceptions” shows that there are conditions to be met before the speaker can wholly agree), or explicit acceptance or rejection. However, it still represents task content, and furthermore, the “exceptions” mentioned add information to the discussion that was not mentioned before without introducing a new conversation topic which leaves us only the *Elaboration* category.

Furthermore, cognitive content communication was distinguished from *task regulation* and *other* communication. These categories were not in our original theoretical framework, but came from experiences while applying the content codes. It occurred to the coders that large amounts of communication were neglected with the original codes, and that these roughly covered conversation with the aim of regulating the task and 'other' communication. This procedure coincides with an open-coding approach to qualitative research. Two regulative categories were used; messages with regulative content that addressed monitoring the problem solving-process (e.g., "I think we have not yet arrived at a good problem definition;" coded Monitoring), and messages that only regulated the conversation (e.g., "Could you make a note of that on the board please;" coded Regulation, also including the regulation of using the NTool, where appropriate).

The segmentation procedure was not separated from the actual coding. The code definitions indicated the use of the conversation turn (in case of face-to-face communication) or the message (in case of computer-mediated communication). Only when a turn in conversation, or a message in NTool, had a clear break in terms of content would a conversation turn be split in two segments.

Rules for Administration

Using only the codes derived from theory was not reliable; additional rules for coding were needed to deal with ambiguous communication. We used two different strategies to arrive at acceptable reliability. A set of decision rules was created for ambiguous cases. The rules for coding "Yes," and "No," in face-to-face communication are a good example of this strategy (see Table 6.2).

The main problem of the meaning of "Yes," and "No," lies in the fact that they can be used in so many different ways, with so many different meanings. A rough approximation in writing of this phenomenon would be the difference between "Yes!" and "Yes…". Especially in the case of the codes agreement, disagreement, acceptance, and rejection, the words "Yes," and "No" were important sources disagreement between coders. We decided to circumvent this issue by only coding entirely explicit agreements, disagreements, acceptances, and rejections, and using a specific set of rules for coding "Yes," and "No."

Table 6.2
Rules for Reliably Coding "Yes" and "No"

1.	Clarification, when "Yes" or "No" immediately follows a Verification in the form of a closed question
2.	Acceptance or Rejection, when "Yes" or "No" immediately follows a Clarification or Elaboration about the meaning of something that has not been part of a Contribution before, and in which confirmation is asked.
3.	Agreement or Disagreement, when "Yes" or "No" is the answer to a question about another's opinion.
4.	Elaboration, if in any of the above cases "Yes" or "No" is followed by a conditional statement.
5.	Elaboration when in doubt.

Unfortunately, this resulted in very low numbers of agreement, disagreement, acceptance and rejection. This means that the actual numbers coded were lower than would have been justified on the basis of the code definitions alone. This is a clear example of the strain that can occur between validity and reliability.

Also, the above strategy alone did not result in good reliability. A second, somewhat more radical measure was recoding several categories into one. In the first study, codes for acceptance, rejection, agreement, and disagreement were recoded as elaboration, which resulted in an inter-coder reliability (Cohen's kappa) of .68. The analysis of the second face-to-face experiment, with a different assistant, included codes for agreement and acceptance (kappa = .71), but not for rejection and disagreements because the number of disagreements and rejections was too low for computing reliability statistics.

The analysis scheme was also used twice for coding computer-mediated communication, with an inter-coder reliability of .70, including all content codes and rules used earlier, except segmentation. Coding may have been easier for electronic communication since electronic communication is naturally segmented into a unit (the message) that is meaningful for the coding scheme.

Step-by-step Development of CSCL and Meaningful Comparisons

NTool development started with a number of unreported face-to-face pilot-studies, in which no analysis of negotiation took place. Next it was tested in four studies in which the results of the earlier studies were used to improve and reinforce the design of NTool. We aimed to ensure comparability between studies by using the same coding scheme each time, with the same type of segments. Still, we judged comparison of counts of codes, such as proposed by Hillman (1999), between studies to be too risky, due to three issues.

The first issue involved comparing *different face-to-face studies*. Reliability figures were different for the studies, and the codes that were used in the final analysis (not during coding itself) differed as well due to recoding for reliability (see above). To a lesser extent, this problem also played a role in the computer-mediated communication studies; in the first study monitoring and regulation were recoded into one to achieve reliability, whereas this was not needed in the second computer-mediated communication study. The latter example is not as severe a problem as the former, because the regulatory codes did not represent the core analysis aim of the study. Nonetheless, the point needs to be made that even when using the same coding scheme across studies, the actual coding procedure may differ.

The second issue was the use of *different coders across studies*. Three of the four studies were coded by a different assistant due to reasons of availability of the assistants. This may or may not have been a problem, because some precautions were taken to enhance comparability. All coders were trained in the use of the coding schema by the author and all reliability measures were computed by comparing coding efforts of an assistant and the author. However, no comparisons between different assistants were (or could be) made, which means that there were no data of a direct comparison between the actual coders of the study data.

Finally, notwithstanding the precautions for comparability taken, the type of comparison went further than the types reported above. In the aforementioned example of comparing the use of deixis in face-to-face and computer-mediated communication, the actual comparison involves a type of communication in different contexts. In the case of the present study, the comparison is of *the effect of an intervention* on a type of communication, repeatedly done in different contexts.

These three issues were deemed so problematic that we abandoned doing direct comparisons of code counts between studies. Instead we focussed on a higher level of aggregation, the comparison of statistical test results. This comparison enabled us to test our main hypotheses and to predict the results of consecutive studies. Furthermore, the unexpected significant differences in earlier studies were used in the design of later studies. In addition, we used an outcome measure of common ground to corroborate the results on negotiation, and in three of the studies semi-structured interviews were used the gain insight in the processes that might underlie the significant differences that were found in negotiation.

In sum, it seems that even when researchers take many precautions to enable meaningful comparisons between studies in terms of code counts, the practice of hiring and training coders and the high costs in terms of time and money may still result in differences between studies that prevent such comparisons. However, the alternative of using significant differences for comparing studies, and using additional measures to corroborate those differences still enables the meaningful comparison of different studies.

Discussion

The main questions in this chapter were: “How can negotiation be validly operationalised in terms of the communication?”, and: “How can the resulting coding scheme be reliably applied to different research contexts without compromising validity?” Regarding the first question, literature emphasises the tensions between validity and reliability, especially in the case of coding latent variables. In our experience, the first step – developing the codes – was fairly straightforward for two reasons. First, using the same theory for both the design of the CSCL-environment and the coding scheme, as proposed by Chi (1997), made it easy to translate between theory and actual communication. Second, the specific theories we used already contained some information about their occurrence in communication, being partly derived from linguistics.

The second step – identification of code-specific communication characteristics and segmentation – showed that our initial code definitions were a good starting point, but also that a lot of the communication did not fit any of our predefined categories. Developing new codes for a number of cases enabled coding all communication. We did not encounter specific segmentation problems. Apparently, the conversation turn as a segmentation criterion sufficiently fit the grain-size of the codes for it to be unproblematic.

The third step – the development of rules for the administration of codes – was indispensable for achieving reliability, but it also endangered validity. The main problem with reliability and validity probably stemmed from the very latent character of the codes, which means that data were likely to be ambiguous for the coders, even if the coders understood and agreed about the definitions of the codes. Nonetheless, the ultimate validity and reliability of the resulting coding schemes and rules for administration were sufficient to draw conclusions with regard to the research questions within the project.

With regard to the second question, literature tends to emphasise caution with the comparison of code counts across studies with the same coding scheme. Our own analyses stress this point, as they suffer from a multitude of small differences that complicate such comparisons, particularly with regard to reliability. Reliability between studies suffered due to different coders across studies and differences due to recoding. These differences show that even the use of a reliable set of rules for the administration of codes, conclusions with regard to reliability cannot be easily generalised to different coders and studies.

A general point needs to be made about the use of coding schemes developed by others. Earlier in this chapter we claimed that new CSCL-development projects generally require new analysis schemes. Based on the above, it could be argued that using rules for administration developed by others as well as their code definitions can enable the use of others' coding schemes, that is, in the rare case where an existing coding scheme is applicable. However, rules for administration of codes are generally absent from scientific articles that employ content analysis (Strijbos et al., in press). Furthermore, the present results show that each coder has his/her own capacities of understanding a coding scheme, and that individual differences may require different rules for the administration of codes. Thus, the availability of others' rules for the administration might increase the adoptability of others' coding schemes, but it is not a guarantee for reliability or comparability between studies.

As an alternative, we compared studies on the basis of statistical test outcomes only. It appears from our studies that this strategy enables meaningful comparisons between studies, without crossing the borders of what is reliable, but it may also be overly conservative. This brings up the question of other ways for comparing data from different studies with the same coding scheme, that neither compromise reliability, nor are too conservative. The question of comparison then becomes somewhat like that of a meta-study, in which different studies are compared as well. Although meta-studies differ from CSCL-research in that they use a wide variety of coding schemes, and do not require comparisons within a project, it may be possible to adopt some of their methodologies for comparing different studies.

In general, CSCL-analysis must take both validity of the coding scheme and the associated measurement strategy into account. Furthermore, it is important to distinguish between these two aspects of analysis, as the one may compromise the other. Aiming for reliability can render a coding scheme invalid if one applies a measurement strategy without proper care for validity. Furthermore, results with the same coding scheme across different research settings can only be meaningfully compared with utmost care.

Chapter 7

Discussion

The research question of this thesis was: How can we support the collaborative construction of knowledge in multidisciplinary teams so as to help the teams to better solve complex problems?

We took the need for multidisciplinary teams to integrate multiple perspectives for complex problem solving as a starting point for the design of NTool, an ICT-tool to support problem solving by multidisciplinary teams. Based on an integration of theory from linguistic (Clark & Schaefer, 1989) and cognitive (Bromme, 2000) sciences on common ground we designed a formalism to support the negotiation of common ground.

We hypothesised that (1) encouraging people to make their individual perspectives tangible to their fellow team members would facilitate the negotiation of common ground, and also (2) would result in more common ground. The formalism aimed to achieve this by requiring discussion partners to explicitly verify their own understanding of the contributions of other team members. Furthermore, it was assumed that the more coercively the support mechanism was applied, the more effective it would be in terms of increased grounding activity in communication and the amount of common ground a group would achieve.

Table 7.1
Series of Experiments and Experimental Conditions

	Formalism		NTool		
	None	Face-to-Face	Idio-syncratic	Scripted	Stringent
Face-to-face pilot	X	X			
NTool laboratory study			X	X	X
Cognitive load study	X	X			
Practical educational study			X		X

The main hypotheses were tested in four consecutive experiments (see Table 7.1), starting with a *face-to-face pilot* in which a pen-and-paper implementation of the formalism was used. Next a *laboratory study* with NTool, an ICT-implementation of the formalism, was conducted to test the effects of coercion (low coercion via idiosyncratic use of NTool, medium coercion via scripted use of NTool, and high coercion via stringent use of NTool) on formalism effectiveness. The third study again was again face-to-face in which we tested the usefulness of *cognitive load measurements* and *stimulated recall* as an addition to our methodology. In the fourth and final study NTool was used in a *practical educational* setting within a secondary vocational education school.

The various studies shared some methodological aspects. In all studies participants from three different disciplinary backgrounds were divided into teams of three and were given the task of solving a complex problem. Group discussions, be it with NTool or face-to-face, were recorded and subsequently coded for negotiation of common ground. Furthermore, both before (the pre-test) and after (the post-test) the collaborative task, participants were asked to solve the same problem individually. The content overlap between the post-tests within one team was used as a measurement of common ground.

The results from the face-to-face pilot strengthened the hypothesis that the formalism could facilitate negotiation of common ground, although results with regard to common ground itself were not statistically significant. Furthermore, they showed that the formalism positively affected the extent to which participants were committed to negotiating common ground. The NTool laboratory study showed that NTool does affect both the negotiation of common ground and the amount of common ground achieved after collaboration, and that it does so increasingly with more coercion, as hypothesised. However, some unexpected questionnaire results about common ground suggested that the Scripted version of NTool resulted in less common ground than the other versions, as perceived by the participants. The specific way coercion was implemented in the Scripted version may have disrupted communication to some extent, which was shown by excessive regulation activities with this version. The other versions did not seem to influence collaboration in a disruptive way.

Building on the unexpected results from the laboratory study, the cognitive load study used cognitive load measures and stimulated recall interviews to gain insight in overall effects of the NTool instruction, and to distinguish them from the effects caused only by coercion. Cognitive load after group collaboration did not differ with or without instruction, suggesting that cognitive load with the instruction remained within working memory capacity. Cognitive load after working individually again was significantly higher with instruction than without instruction, which suggests germane effects of the instruction. Furthermore, the interview data suggested that the instruction might result in germane processing by allowing co-construction of knowledge and negotiation activities, and in postponing one's primary reaction, which could in turn allow for consideration of another's contributions that otherwise does not take place. The interview data thus lend credence to the hypothesis that encouraging people to make their individual perspectives tangible to their fellow team members would facilitate the negotiation of common ground.

Finally, the practical education study tested the Idiosyncratic and Stringent versions of NTool with second year vocational education students. Again, the Stringent version, that is, high coercion, of NTool resulted in the most negotiation activities. However, contrary to our expectations, and also contrary to previous findings, common ground was highest in the Idiosyncratic groups. In this study, the Stringent version of NTool may have been more taxing than the

Idiosyncratic version, as shown by increased ‘other’ communication and increased regulation. Finally, cognitive load was higher in the Stringent groups than in the Idiosyncratic groups, and psychological safety was shown to positively affect negotiation. This means that a sense of being safe within a team affects the extend to which people explicitly verify their understanding.

Overall, it can be concluded that NTool does affect negotiation of common ground and also common ground itself. The actual use of NTool has to be carefully tailored to the intended audience.

Negotiation and Common Ground

Our experiments showed that the formalism affected negotiation activity, and that negotiation increased with coercion. This confirms the first hypothesis. Results of common ground differed between studies; in the NTool laboratory-study common ground *increased* with coercion, while common ground *decreased* with coercion in the practical educational study. The latter result is especially interesting because in the same study negotiation increased with coercion.

The difference between the NTool laboratory-study and the practical educational study can be explained in a number of ways. The first explanation lies in the difference in participants. In the NTool laboratory-study participants were senior year university students with both ample domain knowledge and experience in working in groups. The participants in the practical educational study were considerably younger, less knowledgeable, and less experienced. Furthermore, some data suggested that participant motivation in this study may have been rather low. It could be argued that participants did try to carry out the instructions they were given, resulting in the increased appearance of negotiation processes in communication, but without an associated significant increase of cognitive processes of comparing another's perspective with one's own perspective. In other words, surface aspects of communication would then be affected by NTool, but not the underlying cognitive processes. However, this does not explain why groups with less common ground reported more cognitive load during the post-test.

Another explanation might be that participants needed to reorganise their own knowledge before they could integrated the others' knowledge. It is known that new knowledge can trigger reorganisation of prior knowledge, especially when the new knowledge is not readily understandable in terms of the prior knowledge (Boshuizen & Tabachneck-Schijf, 1998). If NTool did in fact increase negotiation processes and affected the processing of others' knowledge during collaboration, it may have been the case that the teams that negotiated most also needed to reorganise their prior knowledge more. As there was no additional time to do this during discussion, this reorganisation must then take place during the post-test (solving the problem individually after having done so collaboratively), which would explain why these teams also reported a higher cognitive load during the post-test, especially in the afternoon session.

This explanation is more far-reaching because it implies a different relation between negotiation and common ground. The former explanation and the other studies reported here all implicitly assumed a positive relation between negotiation processes and common ground, albeit excluding negotiation as a superficial aspect of communication only. Below, we will elaborate on this.

In sum, our studies have shown that NTool affects negotiation; the more coercively the formalism is applied, the more negotiation of common ground will take place. Furthermore, there is evidence that NTool can increase common ground, although this effect may be bound to specific user conditions, like motivation and sophistication of prior knowledge.

An NTool Effect Mechanism

The main hypothesis carried an assumption about a mechanism for NTool to affect communication, being that explicit verification of understanding would make differences in perspective explicit, which in turn would offer a starting point for reaching common ground, but it was not clear on how NTool could achieve this. We performed explorative interviews to study the actual mechanism through which the formalism affected grounding processes and to compare it with the hypothesised mechanism. This was done for both face-to-face studies and the NTool laboratory study.

The interview analyses from the face-to-face studies suggested that the formalism allowed or promoted co-construction of knowledge and negotiation activities by delaying participants' primary reactions. This delay would then allow for additional consideration of another's contributions. Analyses from the NTool laboratory study are in line with these findings. Participants reported that they initially were inclined to give their opinion in reaction to a new contribution. NTool restrained participants from doing so, which may have led to additional consideration and verification of contributions. Subsequent clarification may have changed others' opinions from disagreeing to accepting and sometimes agreeing upon a contribution. In sum, it seems that verifying and clarifying contributions instead of immediately taking a position increases the chances that contributions are accepted, and become part of common ground.

The Value of Cognitive Load Measurements

In the last two studies we employed measurements of cognitive load in reaction to some unexpected results in the NTool laboratory study. It appeared that ICT-tools may have adverse effects on communication that are not apparent when only measures of intended outcomes are used. In comparison, cognitive load measurements offered an additional view on the effects of NTool.

In the first place, we used cognitive load measurements to gain insight in whether NTool taxed the users' cognitive capacities too much for an effect on grounding to take place. This would be the case, for example, when NTool versions that were more difficult to use resulted in significantly higher levels of cognitive load during collaboration. A comparison of the Stringent and Idiosyncratic versions in the secondary vocational education study showed that this indeed might have been the case for that specific participant population. No cognitive load measurements were taken in the laboratory NTool study, but the data suggested that the Scripted version of NTool might have caused cognitive over-loading for the population of senior year university students (based on increased regulation activity with that version), whereas this did not seem the case for the Idiosyncratic and Stringent versions.

In the second place, we used cognitive load measurements to detect differences in processing *after* NTool had been used which would suggest a lasting effect that could not be attributed to extraneous effects of NTool itself. Such effects were shown in both the face-to-face cognitive load study and the secondary vocational education study. Although in the latter study results of cognitive load measurements were in seeming contradiction with the common ground measurements, they still offered additional explaining power. Overall it may be concluded that cognitive load measurements can provide additional insights in the (possibly adverse) effects of ICT-tools, while they are highly unobtrusive and easy to implement.

Directions for Further Research

Our results suggest a number of directions for further research. First, there are some unanswered questions about the dynamic relation between negotiation of common ground and common ground itself. This thesis is built upon the implicit assumption of a positive, (curvi)linear relationship between negotiation processes and common ground. However, some of our results contradict this assumption.

A second direction for future research lies in different views on the nature of common ground and how to measure it. This thesis has used quite conservative measures of common ground, focussing on the most explicit and tangible aspects of common ground, while neglecting more ephemeral aspects. The question can be raised as to whether such different conceptualisations in actuality refer to different constructs, and what consequences they have for common ground.

A third direction for further research is the continued development of NTool. NTool has proven itself in the laboratory, and it has been shown that NTool can be adapted to the classroom. More research is needed into *how* NTool should be adapted to different environments.

Finally, none of the studies in this thesis has tested NTool or its formalism in practical professional contexts, where different interests are at stake, which might affect grounding processes in different and unexpected ways. Further research on the NTool formalism (not necessarily with actual use of NTool itself) may give insight in the non-cognitive, non-task processes that affect support of grounding processes.

The Relation Between Negotiation and Common Ground

There seems to be an implicit assumption that negotiation of common ground is positively related to common ground itself. Our laboratory study confirms this view, but the results obtained from the secondary vocational training study institute cast some doubt. A different relation between negotiation and common ground, where negotiation initially causes a decrease in common ground, and only after that starts to affect common ground positively, provides an explanation for the differences between the laboratory study and the practical educational study. Figure 7.1b depicts such a relation, compared to a (curvi) linear relation (Figure 7.1a).

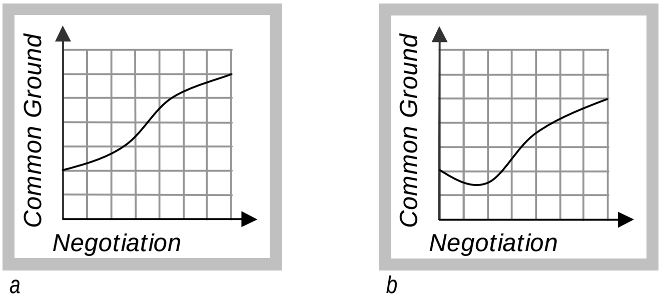

Figure 7.1a and b: The relation between negotiation and common ground.

How can such a relationship be explained? If we take the perspective of one team member, we can see that the others' contributions stem from different knowledge bases than his/her own, particularly in the case of a multidisciplinary team. This also means that their contributions are not immediately available to him for meaningful representation, as his/her own representations are not compatible with them (Boshuizen & Tabachneck-Schijf, 1998). Boshuizen and (Tabachnek-)Schijf argue that to be able to integrate two representations, one needs to restructure one's own representations so as to create an 'interface' that allows for the meaningful translation between representations. The need for reorganising one's own knowledge before integrating knowledge from a different perspective could then initially *increase* the extent to which team members are conscious of the representational differences before they start addressing their differences through negotiation. This would lead to the hypothesis that in collaboration, especially in the case of multidisciplinary teams, common ground initially decreases before it increases as team members become aware of their representational differences. In other words, the amount of time between the onset of collaboration and measuring common ground could be instrumental in determining whether and increase in negotiation is accompanied by an increase in common ground.

A Reflection on the Nature of Common Ground

The specific measures of common ground used in this thesis may have been influenced by research done within the field of CSCL. CSCL is basically an educational field of research, aimed at the educational goal of retention of what has been learned or dealt with. This is reflected by our measures of common ground taken *after* actual communication had taken place. Measures were done with individual, independent participants instead of actual teams, and only explicitly mentioned topics were considered for the common ground measurement. This way of measuring common ground is neglectful of more temporary and/or implicit aspects of common ground. Indeed, all aspects of common ground that depend on contextual characteristics of the collaboration are excluded. In other words, the measures reported in this thesis provide a very conservative estimate of common ground, with an emphasis on explicit, transferable knowledge.

Alternatively, common ground might also be conceived as being very ephemeral, depending heavily on specific situations, and being highly contextualised. Associated measures could, for example, consist of parroting each others concepts, using each other's language, and referring to shared objects within a group’s immediate surroundings. Such a conception of common ground would resemble an interface between people, the short-lived capacity that people feel they know how to communicate with each other, and are confident in understanding the other and making oneself understood (Boshuizen & Tabachneck-Schijf, 1998). Such measures would capture the more implicit and short-lived aspects of common ground as well, although it could result in an overestimation of the actual common ground present within a group.

The above views of common ground, being either ephemeral and bound by the collaboration context, or lasting, being retained and available for conscious reflection after communication, may seem to oppose each other, but they can also be seen as the far ends of a continuum that ranges from implicit, highly contextualised and non-transferable to explicit, decontextualised and transferable aspects of common ground. The question of choosing a particular common ground measure then depends on the more important question of “common ground for what?” For example, as previously noted, from an educational perspective retention of knowledge is an important issue, which means that for assessment purposes it is feasible to take a view of decontextualised, explicit, and transferable common ground. The question then must be raised whether problem-solving groups require the same extent of transferability and explicitness.

Future research could focus on different measures of common ground to test whether common ground indeed can range from highly ephemeral to highly solid, or whether further theoretical distinctions need to be made. For instance, researchers could correlate different measures of common ground statistically. A highly ephemeral measurement of common ground might be counting the number of instances that team members use the same words during short segments of conversation, and to correlate a weighted mean over these segments with a more static measure such as has been used in this thesis. Furthermore, it could be explored through interviews what believes about common ground people have. In this respect, an expression like “being on the same wavelength” is a clear example of a concept of common ground.

Continued Development of NTool

The studies reported here have taken the NTool formalism from the drawing table to the computer. However, from start to finish we did not do any major modifications of the formalism rules besides adapting it to allow for different degrees of coercion. Some question remains as to whether the NTool formalism can be implemented in other ways, resulting in other collaboration scripts than the present idiosyncratic, stringent and scripted ones. For example, the current kinds of collaboration scripts did not allow participants to work on shared content; they could only post individual messages and they could not edit each other's messages, let alone write one together. In other words, in the present variants of NTool the actual writings are still individually made and owned, while it could be argued that requiring users to write together could increase the extent to which the users feel a need to negotiate common ground.

Another option for further development might lie in making the type of support more dynamic over time. Little can be learned about changing negotiation needs and habits as time proceeds from the studies in this thesis, while a problem solving endeavour self-evidently involves specific phases such as problem definition, problem analysis, and solution development. It may be the case that as a team collaboratively performs a task, its negotiation needs change over time, according to the (problem-solving) phase it is in. With regard to NTool the associated challenge then is how to accommodate these task phases in the NTool design.

Finally, the current design of NTool was entirely text-based. Other types of external representation than text only, such as diagrams, could add powerful structuring tools to the current NTool design, which would result on possible combinations of problem- or argumentation-structuring support and support for common ground. Obviously, such designs of decision-support would require more intensive training and planning than the current NTool versions to accommodate possible cognitive overload caused by the instructions. Nonetheless it would be interesting to test such a design in a 2 X 2 factorial design, using experimental conditions for no support at all, support for grounding only, support for argumentation- or problem-structuring only, and the two support principles combined.

NTool and Professional Practice

None of the studies in this thesis tested NTool or its formalism in practical professional contexts where different interests are at stake that might affect the grounding processes in unexpected ways. Such task-contexts are richer than the tasks from this study in many ways, and involve other task-processes as well. Also, further research on the NTool formalism (not necessarily with actual use of NTool itself) may give insight in the non-cognitive, non-task processes that affect support of grounding processes.

First, again the combination of negotiation support and other types of group support needs to be mentioned. However, one may expect that professionals already use some method of problem structuring or argumentation structuring. This means that the research context becomes very messy, and will require more qualitative methods of enquiry such as explorative interview techniques and participant observation. It would be especially interesting to take an NTool-like type of decision-support and to study whether, and if so, how its use is affected by communication as it appears in the professional practice.

More specifically, it may be the case that NTool can trigger a number of effects that did not occur in the studies reported here, because none of the studies manipulated the need for strategic behaviour, and the studied settings did not require it. The professional practice may be expected to behave differently in this respect. For instance, one could envision that users try to evade the NTool rules on the one hand, and also that users call their peers to play by the NTool rules to prevent such behaviour. In other words, it may be expected that a certain amount of tool appropriation can occur, with NTool uses that may vary greatly from the way NTool was intended to be used.

And finally, a study in a professional context would provide insight in the usefulness of NTool. In the professional context, useful not only means that NTool achieves what it intends to, but also that its use is regarded valuable.

References

Alpay, L., Giboin, A., & Dieng, R. (1998). Accidentology: An example of problem solving by multiple agents with multiple representations. In M. W. Van Someren, P. Reimann, H. P. A. Boshuizen & T. De Jong (Eds.), *Learning with multiple representations* (pp. 152-174). Oxford, UK: Elsevier.

Avouris, N. M., Dimitracopoulou, A., & Komis, V. (2003). On analysis of collaborative problem solving: an object-oriented approach. *Computers in Human Behavior, 19*(2), 147-167.

Baker, M. J., Hansen, T., Joiner, R., & Traum, D. (1999). The role of grounding in collaborative learning tasks. In P. Dillenbourg (Ed.), *Collaborative learning: Cognitive and computational approaches* (pp. 31-63). Amsterdam, The Netherlands: Pergamon / Elsevier Science.

Baker, M. J., & Lund, K. (1997). Promoting reflective interactions in a computer-supported collaborative learning environment. *Journal of Computer Assisted Learning, 13*, 175-193.

Barron, B. (2003). When smart groups fail. *The Journal of the Learning Sciences, 12*(3), 307-359.

Barros, B., & Verdejo, M. F. (1999, July). *An approach to analyse collaboration when shared structured workspaces are used for carrying out group learning processes.* Paper presented at the International Conference on Artificial Intelligence in Education, Le Mans, France.

Barrows, H. S., & Tamblyn, R. M. (1980). *Problem-Based Learning. An approach to medical education*. New York, USA: Springer.

Bechky, B. A. (2003). Sharing meaning across occupational communities: The transformation of understanding on a production floor. *Organization Science, 14*(3), 312-330.

Beers, P. J., Boshuizen, H. P. A., Kirschner, P. A., & Gijselaers, W. H. (2005). Computer support for knowledge construction in collaborative learning environments. *Computers in Human Behavior, 21*(4), 623-643.

Beers, P. J., Van Asselt, M. B. A., Vermunt, J. D., & Kirschner, P. A. (2003). Policy, information and learning. *The Journal of Workplace Learning, 15*(2), 70-79.

Boland, R. J., & Tenkasi, R. V. (1995). Perspective making and perspective taking in communities of knowing. *Organization Science, 6*(4).

Boshuizen, H. P. A., & Tabachneck-Schijf, J. M. (1998). Problem solving with multiple representations by multiple and single agents: An analysis of the issues involved. In M. W. Van Someren, P. Reimann, H. P. A. Boshuizen & T. De Jong (Eds.), *Learning with multiple representations* (pp. 137-151). Oxford, UK: Elsevier.

Bromme, R. (2000). Beyond one's own perspective: The psychology of cognitive interdisciplinarity. In P. Weingart & N. Stehr (Eds.), *Practicing interdisciplinarity* (pp. 115-133). Toronto, Canada: University of Toronto Press.

Bromme, R., & Nückles, M. (1998). Perspective-taking between medical doctors and nurses: A study on multiple representations of different experts with common tasks. In M. W. Van Someren, P. Reimann, H. P. A. Boshuizen & T. De Jong (Eds.), *Learning with multiple representations* (pp. 175-196). Oxford, UK: Elsevier.

Bromme, R., Rambow, R., & Nückles, M. (2001). Expertise and estimating what other people know: The influence of professional experience and type of knowledge. *Journal of Experimental Psychology: Applied, 7*(4), 317-330.

Buckingham Shum, S., & Hammond, N. (1994). Argumentation-based design rationale: What use at what cost? *International Journal of Human-Computer Studies, 40*(4), 603-652.

Buckingham Shum, S. J., MacLean, A., Bellotti, V. M. E., & Hammond, N. V. (1997). Graphical argumentation and design cognition. *Human-Computer Interaction, 12*, 267-300.

Capozzoli, M., McSweeney, L., & Sinha, D. (1999). Beyond kappa: A review of interrater agreement measures. *The Canadian Journal of Statistics, 27*(1), 3-23.

Chang, H.-T., & Lee, A. T. (2001). The relationship between psychological safety, organisation context support and team learning behaviour in Taiwan. *Global Journal of Engineering Education, 5*(2), 185-192.

Checkland, P. (1981). *Systems thinking, systems practice*. Chichester, UK: John Wiley.

Chi, M. T. H. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *The Journal of the Learning Sciences, 6*(3), 271-315.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 127-149). Washington DC, USA: American Psychological Association.

Clark, H. H., & Schaefer, E. F. (1989). Contributing to discourse. *Cognitive Science, 13*, 259-294.

Conklin, E. J., & Weil, W. (1997). *Wicked Problems: Naming the pain in organizations* (White Paper): Group Decision Support Systems, Inc.

Conklin, J., & Begeman, M. L. (1987, November). *gIBIS: A hypertext tool for team design deliberation.* Paper presented at the Hypertext'87, Chapel Hill, US.

Courtney, J. F. (2001). Decision making and knowledge management in inquiring organizations: toward a new decision-making paradigm for DSS. *Decision Support Systems, 31*, 17-38.

Derry, S. J., Adams DuRussel, L., & O'Donnell, A. M. (1998). Individual and distributed cognitions in interdisciplinary teamwork: A developing case study and emerging theory. *Educational Psychology Review, 10*(1), 25-56.

DeTombe, D. J. (2002). Complex societal problems in operational research. *European Journal of Operational Research, 140*, 232-240.

Dillenbourg, P. (2002). Over-scripting CSCL: The risks of blending collaborative learning with instructional design. In P. A. Kirschner (Ed.), *Three worlds of CSCL: Can we support CSCL?* (pp. 61-91). Heerlen, The Netherlands: Open Universiteit Nederland.

Dillenbourg, P., Baker, M., Blaye, A., & O'Malley, C. (1995). The evolution of research on collaborative learning. In P. Reimann & H. Spada (Eds.), *Learning in humans and machine: Towards an interdisciplinary learning science* (pp. 189-211). Oxford, UK: Elsevier.

Douglas, M., & Wildavsky, A. B. (1982). *Risk and culture: An essay on the selection of technical and environmental dangers*. Berkeley, USA: University of California Press.

Eden, C., & Ackermann, F. (2001). Group decision and negotiation in strategy making. *Group Decision and Negotiation, 10*, 119-140.

Edmondson, A. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly, 44*, 350-383.

Evans, M. W., & Marciniak, J. (1987). *Software quality assurance and management*. New York, USA: John Wiley &Sons.

Fischer, F., Bruhn, J., Gräsel, C., & Mandl, H. (2002). Fostering collaborative knowledge construction with visualization tools. *Learning and Instruction, 12*, 213-232.

Fischer, G., Nakakoji, K., & Ostwald, J. (1995, August 23-25). *Supporting the Evolution of Design Artifacts with Representation of Context and Intent.* Paper presented at the Designing Interactive Systems 1995 Conference, Ann Arbor, USA.

Forrester, J. W. (1971). Counterintuitive behavior of social systems. *Technology Review, 73*(3), 52-68.

Garrison, D. R., Anderson, T., & Archer, W. (2000). Critical inquiry in a text-based environment: computer conferencing in higher education. *The Internet and Higher Education, 2*(2-3), 87-105.

Hasan, H., & Gould, E. (2001). Support for the sense-making activity of managers. *Decision Support Systems, 31*, 71-86.

Hillman, D. C. A. (1999). A new method for analyzing patterns of interaction. *The American Journal of Distance Education, 13*(2), 37-47.

Hisschemöller, M., & Hoppe, R. (1995-96). Coping with intractable controversies: the case for problem structuring in policy design and analysis. *Knowledge and Policy: The International Journal of Knowledge Transfer and Utilization, 8*(4), 40-60.

Johnson, D. W., & Johnson, R. T. (1994). Structuring academic controversy. In S. Sharan (Ed.), *Handbook of cooperative learning methods* (pp. 66-81). Westport, US: Praeger.

Jonassen, D. H. (2000). *Computers as mindtools for schools: Engaging critical thinking*. Columbus, US: Prentice-Hall.

Kanselaar, G., Erkens, G., Andriessen, J., Prangsma, M., Veerman, A., & Jaspers, J. (2003). Designing argumentation tools for collaborative learning. In P. A. Kirschner, S. J. Buckingham Shum & C. S. Carr (Eds.), *Visualizing argumentation: Software tools for collaborative and educational sense-making* (pp. 51-73). London, UK: Springer.

Kirschner, P., Strijbos, J.-W., Kreijns, K., & Beers, P. J. (2004). Designing electronic collaborative learning environments. *Educational Technology Research and Development, 52*(3), 47-66.

Kirschner, P., Van Bruggen, J., & Duffy, T. (2003, June). *COVASE: Collaborative visualization for constructivist learning.* Paper presented at the CSCL 2003, Bergen, Norway.

Kreijns, K. (2004). *Sociable CSCL environments: Social affordances, sociability, and social presence.* Unpublished PhD-thesis, Open University of the Netherlands, Heerlen, the Netherlands.

Krippendorf, K. (1980). *Content analysis: An introduction to its methodology*. Beverly Hills, US: Age.

Kunz, W., & Rittel, H. W. J. (1970/1979). *Issues as elements of information systems* (Working Paper No. 131). Heidelberg, Germany - Berkeley, USA.

Landis, J., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159-174.

Leontjev, A. N. (1981). The problem of activity in psychology. In J. V. Wertsch (Ed.), *The concept of activity in soviet psychology* (pp. 37-71). Armonk, USA: Sharp.

Lesgold, A., Rubinson, H., Feltovich, P., Glaser, R., Klopfer, D., & Wang, Y. (1988). Expertise in a complex skill: diagnosing X-ray pictures. In M. T. H. Chi, R. Glaser & M. J. Farr (Eds.), *The nature of expertise*. Hillsdale, NJ: Erlbaum.

Lipponen, L., Rahikainen, M., Lallimo, J., & Hakkarainen, K. (2003). Patterns of participation and discourse in elementary students' computer-supported collaborative learning. *Learning and Instruction, 13*, 487-509.

Lomi, A., Larsen, E. R., & Ginsberg, A. (1997). Adaptive learning in organizations: A system-dynamics-based exploration. *Journal of Management, 23*(4), 561-582.

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

Mulder, I. (1999). *Understanding technology mediated interaction processes: A theoretical context* (GigaCSCW/D1.4.1 No. TI/RS/99042). Enschede, The Netherlands: Telematica Instituut.

Mulder, I., Swaak, J., & Kessels, J. (2002). Assessing group learning and shared understanding in technology-mediated interaction. *Educational Technology & Society, 5*(1), 35-47.

Nieveen, N. (1999). Prototyping to reach product quality. In J. van den Akker, R. M. Branch, K. Gustafson, N. Nieveen & T. Plomp (Eds.), *Design approaches and tools in education and training* (pp. 125-135). Dordrecht, the Netherlands: Kluwer Academic Publishers.

Noldus, L. P. J. J., Trienes, R. J. H., Hendriksen, A. H. M., Jansen, H., & Jansen, R. G. (2000). The Observer Video-Pro: New software for the collection, management, and presentation of time-structured data from videotapes and digital media files. *Behavior Research Methods, Instruments & Computers, 32*, 197-206.

Norman, D. A. (1993). *Things that make us smart: Defending human attributes in the age of the machine*.Cambridge, US: Perseus Publishing.

O'Donnell, A. M., & Dansereau, D. F. (1992). Scripted cooperation in student dyads: A method for analyzing and enhancing academic learning and performance. In R. Hertz-Lazarowitz & N. Miller (Eds.), *Interaction in cooperative groups: The theoretical anatomy of group learning* (pp. 120-141). New York, USA: Cambridge University Press.

Ostwald, J. (1996). *Knowledge construction in software development: The evolving artifact approach.* Unpublished PhD Thesis, University of Colorado, Boulder, USA.

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist, 38*(1), 63-71.

Rittel, H. W. H., & Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy Sciences, 4*, 155-169.

Rotmans, J. (1998). Methods for IA: The challenges and opportunities ahead. *Environmental Modeling and Assessment, 3*(3, Special Issue: Challenges and opportunities for Integrated Assessment), 155-179.

Rotmans, J., Kemp, R., & Van Asselt, M. (2001). More evolution than revolution: Transition management in public policy. *Foresight/ the journal of future studies, strategic thinking and policy, 03*(01), 1-17.

Rourke, L., & Anderson, T. (2004). Validity in qualitative content analysis. *Educational Technology Research and Development, 52*(1), 5-18.

Rourke, L., Anderson, T., Garrison, D. R., & Archer, W. (2001). Methodological issues in the content analysis of computer conference transcripts. *International Journal of Artificial Intelligence in Education, 12*, 8-22.

Rourke, L., Anderson, T., Garrison, R., & Archer, W. (1999). Assessing social presence in asynchronous text-based computer conferencing. *Journal of Distance Education, 14*(2), 50-71.

Salomon, G., & Perkins, D. N. (1998). Individual and social aspects of learning. *Review of Research in Education, 23*, 1-24.

Scaife, M., & Rogers, Y. (1996). External cognition: how do graphical representations work? *International Journal of Human-Computer Studies, 45*, 185-213.

Selvin, A. (2002). Fostering collective intelligence: Helping groups use visualized argumentation. In P. A. Kirschner, S. Buckingham Shum & C. Carr (Eds.), *Visualizing argumentation: Software tools for collaborative and educational sesne-making*. London, UK: Springer Verlag.

Senge, P. M. (1990). *The fifth discipline: Art and practice of the learning organization*. New York, USA: Doubleday.

Soller, A. L. (2002). *Computational analysis of knowledge sharing in collaborative distance learning.* Unpublished PhD-thesis, University of Pittsburgh, Pittsburgh, USA.

Sterman, J. D. (1994). Learning in and about complex systems. *System Dynamics Review, 10*(2-3), 291-330.

Strauss, A. L. (1987). *Qualitative analysis for social scientists*. Cambridge, UK: Cambridge University Press.

Strijbos, J.-W. (2004). *The effect of roles on computer-supported collaborative learning.* Unpublished PhD-thesis, Open University of the Netherlands, Heerlen, the Netherlands.

Strijbos, J. W., Martens, R. L., Prins, F. J., & Jochems, W. M. G. (in press). Content analysis: what are they talking about? *Computers & Education*.

Sullivan Palincsar, A. (1998). Social constructivist perspectives on teaching and learning. *Annual Review of Psychology, 49*, 345-375.

Suthers, D. (1999, January). *Representational support for collaborative inquiry.* Paper presented at the 32nd Hawaii International Conference on System Sciences, Maui, Hawaii, US.

Suthers, D., Girardeau, L., & Hundhausen, C. (2003, June). *Deictic roles of external representations in face-to-face and online collaboration.* Paper presented at the CSCL, Bergen, Norway.

Suthers, D. D. (2001). Towards a systematic study of representational guidance for collaborative learning discourse. *Journal of Universal Computer Science, 7*(3). Retrieved 23 January 2004, from http://www.jucs.org/jucs_7_3/towards_a_systematic_study

Sweller, J., Van Merriënboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*(3), 251-295.

Thomas, A. P., Bull, P., & Roger, D. (1982). Conversational exchange analysis. *Journal of Language and Social Psychology, 1*(2), 141-155.

Toulmin, S. E. (1958). *The uses of argument*.Cambridge,UK: Cambridge University Press.

Van Asselt, M. B. A. (2000). *Perspectives on uncertainty and risk*. Dordrecht: Kluwer.

Van Bruggen, J. M. (2003). *Explorations in graphical argumentation; The use of external representations in collaborative problem solving.* Unpublished PhD-thesis, Open University of the Netherlands, Heerlen, The Netherlands.

Van Bruggen, J. M., Kirschner, P. A., & Jochems, W. (2002). External representation of argumentation in CSCL and the management of cognitive load. *Learning and Instruction, 12*, 121-138.

Van den Bossche, P., Gijselaers, W., Segers, M., & Woltjer, G. (2005). *Sharing expertise in management: A study on team learning and its effect on shared mental models*. Paper presented at the Academy of Management, Hawaii.

Van den Bossche, P., Gijselaers, W., & Segers, M. (2004, April). *Effects of social factors on shared cognition and teamwork in collaborative learning environments.* Paper presented at the AERA, San Diego, USA.

Van Merriënboer, J. J. G., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review, 17*(2), 147-177.

Van Someren, M. W., Reimann, P., Boshuizen, H. P. A., & De Jong, T. (Eds.). (1998). *Learning with multiple representations*.Oxford, UK: Elsevier.

Vennix, J. A. M. (1996). *Group model building: Facilitating team learning using system dynamics*. Chichester, UK: John Wiley & Sons.

Verschuren, P. J. M. (1997). Ontwerpgericht onderzoek: Een methodologische bezinning [Design-based research: A methodological reflection]. *Bedrijfskunde, 69*(3), 82-89.

Walsh, J. P. (1995). Managerial and organizational cognition: Notes from a trip down memory lane. *Organization Science, 6*(3), 280-321.

Woltjer, G. (2003). *Decisions and macroeconomics: Development and implementation of a simulation game* (Research memorandum No. rm03002). Maastricht, The Netherlands: METEOR, Maastricht research school of Economics of Technology and Organizations; Maastricht University.

Summary

Society is increasingly confronted with complexity. Complex societal problems require novel ways of problem conceptualisation, using knowledge from different scientific disciplines and societal perspectives. Solving such problems is therefore typically done in multidisciplinary or multi-stakeholder teams whose team members have a variety of backgrounds and can bring multiple perspectives to bear on the problem. However, multiple perspectives, although a requirement, are no guarantee for successful problem solving. In order for a multidisciplinary team to arrive at novel problem solutions, its team members have to achieve some measure of common ground, that is, a shared cognitive frame of reference. This thesis aims to add to the growing body of knowledge and techniques of supporting complex problem solving. The research question of this thesis is: How can we support the collaborative construction of knowledge in multidisciplinary teams so as to help the teams to better solve complex problems?

Complex problem solving is here conceived of as a collaborative process that starts with unshared disciplinary knowledge from individuals, and ends with the construction of new knowledge, which takes the form of a problem solution. The negotiation of common ground is an important aspect of this process. According to theories about the negotiation of common ground from the domains of linguistics and the cognitive sciences, people communicate on the basis of assumptions they hold about their fellow team members' (disciplinary) perspective. The negotiation of common ground, thus, is closely related to the existence of multiple perspectives in teams. Our efforts have therefore concentrated on supporting the negotiation of common ground.

Theory on the negotiation of common ground served as a starting point for the development of NegotiationTool (NTool), an ICT-tool to support complex problem solving in multidisciplinary teams. The heart of NTool is a *formalism*, a set of rules that serves as a model for conversation, aimed at making individual perspectives explicit to other team members. The formalism aimed to do so by encouraging the *verification* and *clarification* of each others' contributions to an (electronic) discussion so as to facilitate making differences in understanding explicit, and by requesting participants to explicitly *accept* or *reject* others' contributions, so as to indicate whether they understood them, or judged them intelligible, and to explicitly *agree* or *disagree* with others' contributions, so as to ensure that all positions in the team are explicit. Note that this also enables the case of *agreeing to disagree*, when a contribution is perfectly intelligible and cannot be proven wrong from the contributor's perspective, but is still disagreeable from another's perspective.

We hypothesised that (1) encouraging people to make their individual perspectives tangible to their fellow team members would facilitate the

negotiation of common ground, and also (2) would result in more common ground. This thesis is an account of the iterative testing and development of NTool.

All studies in this thesis roughly followed the same format. Multidisciplinary teams collaboratively solved a complex problem. Analysis of negotiation was done through study of the communication process using techniques from content analysis to create a process overview of negotiating common ground. In the case of the face-to-face studies, the communication was videotaped and analysed with video-coding software. In the case of the computer-mediated discussions the communication was cached on a server and accessible with NTool.

Chapter 2 reports the first study in this project, a face-to-face pilot with a pen-and-paper implementation of the formalism. Participants received a set of coloured markers, a whiteboard and a flip-over. Participants in the formalism condition were instructed to (1) write down new topics on the whiteboard, (2) to write down verifications and clarifications on the flip-over, (3) and to share their positions on the whiteboard again. Other participants could use the materials idiosyncratically. The formalism was shown to affect the negotiation of common ground in the expected way, that is, instructing participants to explicitly verify and clarify their contributions appeared to increase negotiation of common ground. Also, it appeared from interview data that participants with the instruction were more committed to the negotiation of common ground.

Chapter 3 reports the first study with NTool. This was a laboratory study that explored the effects of different levels of coercion on the effectiveness of NTool. Coercion is the use of scripting to constrain participants' actions so as to make them follow the formalism rules more closely. Three conditions were studied, with three corresponding NTool versions. All versions used the same message types for making new *contributions* to the discussion, *verification* and *clarification* of those contributions, and *deciding upon* the contributions (*accepting* and *rejecting*, *agreeing* or *disagreeing*). The Idiosyncratic version (low coercion) used no coercion constrain the use of certain message types at certain times. The Scripted version (medium coercion) resulted in two discussion phases by coercing participants to first clarify and verify all contributions. Only after that had been done were participants given the options to decide upon the contributions. The Stringent version (high coercion) only allowed discussion of one contribution at a time, requiring participants to verify and clarify and then decide upon one contribution before being able to start discussing the next one.

The study showed that coercion was positively related to both the negotiation of common ground and to the amount of common ground. However, it also appeared that the specific way in which the medium coercion condition was designed had some unexpected disruptive effects on communication, as the users of this version used a lot more regulation activities than the users of the high and low coercion versions of NTool.

In Chapter 4, again a face-to-face study is reported. This study built on the unexpected results of the laboratory study. The main aim of this study was to isolate effects of the instruction for using the formalism from the effects of coercion. Aside from the outcome measures used in the previous studies (negotiation of common ground and common ground itself) this study used cognitive load measurements and stimulated recall interviews to gain insight in the overall effects of the formalism. The cognitive load measures showed that the the instruction alone probably had germane effects on common ground and the interview data suggested that the instruction did so by keeping participants from immediately voicing their opinions, who devoted some time to consider their fellows' contributions instead. The interview data supported the hypothesis that encouraging people to make their individual perspectives tangible to their fellow team members facilitates the negotiation of common ground.

Chapter 5 reports the second study with NTool, with high and low coercion versions. This study took place in a secondary vocational education institution. The main goal of the study was to test the robustness of the results from the laboratory study in a more ecologically valid setting. All previously used measures (negotiation, common ground, and cognitive load) were used. Furthermore, questionnaires about psychological safety (a shared belief held by members of a team that the team is safe for interpersonal risk taking) were included to control for possible confounding effects. Again, high coercion resulted in the most negotiation of common ground, but, conversely, also in the least common ground itself. The Stringent version of NTool may have been more taxing than the Idiosyncratic version, as shown by increased 'other' communication and increased regulation. This result is different from the laboratory study, and may be explained by the participants' educational level (senior year university students versus second year vocational education students). Finally, cognitive load was higher in the Stringent groups than in the Idiosyncratic groups, and psychological safety was shown to positively affect negotiation. This means that a sense of being safe within a team affects the extend to which people explicitly verify their understanding.

Overall, it can be concluded that NTool does affect negotiation of common ground and also common ground itself. The actual use of NTool has to be carefully tailored to the intended audience. Furthermore, the studies and the results raised some methodological questions and some directions for further research.

The methodological questions (Chapter 6) pertained to issues with validity and reliability of the used methodologies and the comparability of results across the various studies in this thesis. It was concluded that content analysis of communication is especially difficult in the case of latent content (i.e., content that is not apparent from surface aspects of communication and therefore open to subjective interpretation). Analysis of latent content not only requires clear definitions of what is to be analysed, but also specific rules to deal with ambiguous cases so as to ensure reliability. Sometimes such rules can endanger the validity of a coding scheme. Furthermore, it was shown that the comparison of analyses across studies that employ the same analysis scheme for content analysis is not as straightforward as it might seem. To keep on the safe side, we only compared studies with regard to statistically significant differences in this thesis, and not the analyses themselves.

Chapter 7 offers some for future research. First, it would be interesting to study whether the relation between common ground and negotiation of common ground is a positive one. This thesis is built upon the implicit assumption of a positive, (curvi)linear relationship between negotiation processes and common ground. However, some of our results contradict this assumption. An alternative hypothesis is that negotiation might initially lower common ground, and only result in more common ground after being sustained for some time. A second direction lies in the nature of common ground and how to measure it. This thesis has used quite conservative measures of common ground, focussing on the most explicit and tangible aspects of common ground, while neglecting more ephemeral aspects. The question can be raised as to whether such different conceptualisations in actuality refer to different constructs, and what consequences they have for common ground. Thirdly, NTool has proven itself in the laboratory, and it has been shown that NTool can be adapted to the classroom. It would be interesting to study how NTool might be altered to make it adaptable to a greater variety of environments. Finally, none of the studies in this thesis has tested NTool or its formalism in practical professional contexts, where different interests are at stake, which might affect grounding processes in different and unexpected ways. Further research on the NTool formalism (not necessarily with actual use of NTool itself) may give insight in the non-cognitive, non-task processes that affect support of grounding processes.

Samenvatting

Complexiteit speelt een steeds grotere rol in de maatschappij. Complexe vraagstukken vragen om een vernieuwende probleemaanpak waarbij kennis uit verschillende wetenschappelijke disciplines en maatschappelijke perspectieven gebruikt wordt. Daarom worden complexe vraagstukken vaak opgelost in multidisciplinaire teams of teams met verschillende belanghebbenden, immers, in zulke teams kunnen de teamleden vanuit hun uiteenlopende achtergronden verschillende perspectieven op het probleem werpen. Maar multidisciplinariteit alleen is geen garantie voor goede oplossingen. Wil een multidisciplinair team tot vernieuwende oplossingen komen dan zullen de teamleden tot *common ground* moeten komen, dat is een gemeenschappelijk cognitief referentiekader. Dit proefschrift heeft als doel een bijdrage te leveren aan de kennis van en technieken voor de ondersteuning van het oplossen van complexe vraagstukken. De onderzoeksvaag is: Hoe kunnen we gezamenlijke kennisconstructie in multidisciplinaire teams ondersteunen, opdat die teams beter complexe problemen kunnen oplossen?

Het oplossen van complexe problemen is hier geconceptualiseerd als een collaboratief proces dat begint met ongedeelde, individuele kennis uit verschillende disciplines en dat eindigt met de constructie van nieuwe kennis in de vorm van oplossingen. De negotiatie van common ground (*grounding*) is een belangrijk onderdeel van dit proces. Theorieën over grounding uit de linguïstiek en de cognitieve wetenschappen geven aan dat mensen communiceren op basis van specifieke aannames over het (disciplinaire) perspectief van hun teamgenoten. Daarom is grounding relevant voor de verschillende perspectieven in multidisciplinaire teams. Dit onderzoek is gericht op de ondersteuning van de negotiatie van common ground.

Theorieën over grounding staan aan de basis van de ontwikkeling van NegotiatieTool (NTool), een ICT-tool voor de ondersteuning van multidisciplinaire teams die zich bezighouden met complexe problemen. De kern van NTool is een zogenaamd *formalisme*. Een formalisme is een set van regels die dient als communicatiemodel. Het formalisme van NTool is erop gericht om individuele perspectieven expliciet te maken voor de andere teamleden, door (1) ze ertoe te bewegen actief hun kennis van andermans bijdragen aan (elektronische) discussies te *verifiëren* en te *verhelderen*, opdat verschillen in begrip expliciet worden, en (2) ze te vragen expliciet de bijdragen van anderen te *accepteren* danwel *af te wijzen*, om helder te maken of iedereen het heeft begrepen, en (3) ze expliciet hun *opinie te laten geven*, opdat de verschillende posities in het team expliciet zijn. Dit betekent trouwens niet dat iedereen het uiteindelijk met elkaar eens hoeft te zijn, men kan immers ook iemands positie begrijpen, diegene niet kunnen overtuigen, en het toch met iemand oneens zijn. In het Engels gebruiken we hiervoor de term *agree to disagree*.

Onze hypothesen waren dat (1) het expliciet maken van individuele perspectieven tot meer grounding zou leiden, en dat (2) dat ook tot meer common ground zou leiden. Dit proefschrift verhaalt de stapsgewijze ontwikkeling van NTool.

De studies in dit proefschrift waren grofweg hetzelfde opgezet. Multidiscplinaire teams kregen telkens de opdracht om gezamenlijk een probleem op te lossen. Voor het analyseren van grounding pasten we inhoudsanalyse toe op het communicatieproces.

Hoofdstuk 2 bespreekt de eerste pilot-studie van een pen-en-papier versie van het formalisme. De deelnemers kregen een stel gekleurde stiften, een schoolbord en een flip-over. Deelnemers in de formalisme-conditie kregen de instructie om (1) nieuwe onderwerpen op het schoolbord te schrijven, (2) die over en weer te verifiëren en verhelderen op de flip-over, (3) en hun mening over elk onderwerp weer op het schoolbord te zetten. De andere deelnemers mochten de stiften, het schoolbord en de flip-over naar eigen inzicht gebruiken (de idiosyncratische conditie). Het formalisme bleek grounding te beïnvloeden zoals verwacht, wat wil zeggen dat het expliciet verifiëren en verhelderen van bijdragen leek te leiden tot meer grounding. Daarbij bleek uit interviews dat de deelnemers in de formalisme-conditie meer belang hechtten aan de negotiatie van common ground. Deze resultaten gaven ons voldoende vertrouwen om een ICT-implementatie van het formalisme te maken.

Hoofdstuk 3 bespreekt de eerste studie met NTool. In dit laboratoriumonderzoek werd het effect op grounding onderzocht van de mate van stringentie (*coercion*) waarmee het formalisme werd opgelegd. De mate van stringentie werd bepaald door het aantal vrijheidsgraden dat de deelnemers hadden bij het gebruik van NTool. We testten drie versies van NTool. Alle drie maakten gebruik van een vaste set berichttypes, namelijk *contributie* voor een nieuw gespreksonderwerp, speciale berichttypes voor *verificatie* en *verheldering*, en berichttypes voor het *beslissen* over de gespreksonderwerpen (*acceptatie* en *afwijzing*, *eens* en *oneens*). De Idiosyncratische versie (lage coercion) had geen beperkingen voor het gebruik van de berichttypen. De geScripte versie (gemiddelde coercion) verdeelde de discussie in twee fasen door deelnemers in eerste instantie alleen het gebruik van *contributies*, *verificaties* en *verhelderingen* toe te staan, en pas nadat iedereen alle contributies had geverifieerd de andere berichttypen. De Stringente versie (hoge coercion) beperkte de discussie tot één onderwerp tegelijkertijd, waarbij deelnemers eerst een contributie moesten verifiëren en verhelderen, en er daarna over beslissen, voordat ze verder mochten met de volgende contributie.

Dit onderzoek toonde een positieve relatie tussen coercion en grounding aan, en ook tussen coercion en common ground zelf. Het bleek echter ook dat de geScripte versie tot onverwacht veel regulatie leidde. Waarschijnlijk gaf de specifieke manier waarop coercion in deze versie geïmplementeerd was de deelnemers problemen bij het gebruik van NTool.

Hoofdstuk 4 is een verslag van een tweede studie met de pen-en-papier versie van het formalisme, naar aanleiding van de onverwachte resultaten van de laboratoriumstudie. In dit onderzoek werden de effecten van de formalisme-instructie gescheiden van effecten van coercion. Daarbij voegden we metingen van cognitieve belasting, en interviews aan de hand van videobeelden van de samenwerking, toe aan de metingen uit de vorige studies (grounding en common ground), om een meer algemeen beeld te krijgen van de effecten van het formalisme. De cognitieve belasting-metingen gaven aan dat de instructie voor het formalisme op zichzelf (zonder coercion) een positief effect hadden op common ground, en uit de interviews bleek dat dit kwam doordat de instructie deelnemers ervan weerhield onmiddellijk hun mening te geven; in plaats daarvan besteedden ze hun aandacht aan het overwegen van de contributies van hun teamgenoten. De interviewdata ondersteunden de hypothese dat het expliciet maken van individuele perspectieven tot meer grounding zou leiden.

Hoofdstuk 5 gaat over de tweede studie met NTool, nu met alleen de Idiosyncratische en de Stringente versie. Het onderzoek vond plaats op een MBO-instelling, met als doel het testen van de robuustheid van de resultaten uit de laboratoriumstudie in een omgeving met een hogere ecologische validiteit. We gebruikten alle eerdere maten (grounding, common ground, cognitieve belasting), en ook een vragenlijst over psychologische veiligheid (een gemeenschappelijk vertrouwen onder teamleden dat in hun team ruimte is voor het nemen van interpersoonlijke risico’s) om te controleren voor confounding. Weer gaf de Stringente versie de meeste grounding, maar deze keer ook de minste common ground. Misschien was de Stringente versie van NTool meer belastend dan de Idiosyncratische, blijkens de verhoogde ‘overige’ communicatie en de verhoogde regulatie. Dit is een verschil met het resultaat van de laboratoriumstudie, dat misschien verklaard kan worden door het opleidingsniveau van de deelnemers (vierdejaars universiteitsstudenten versus tweedejaars middelbaar beroepsonderwijs). Verder bleek de cognitieve belasting hoger bij deelnemers uit de Stringente groepen, en psychologische veiligheid bleek een positieve invloed te hebben op grounding. Dit betekent dat een gevoel van veiligheid in het team invloed heeft op de mate waarin teamleden expliciet hun begrip verifiëren.

Dit leidt tot de conclusie dat NTool zowel common ground als de negotiatie ervan beïnvloedt. Bij de inzet van NTool in de praktijk moet zorgvuldig rekening gehouden worden met de beoogde gebruiker. Daarnaast gaven de verschillende onderzoeken aanleiding tot een aantal methodologische vragen en een aantal mogelijkheden voor verder onderzoek.

De methodologische vragen (hoofdstuk 6) hebben betrekking op validiteit en betrouwbaarheid van de gebruikte methoden en de vergelijkbaarheid van de analyses tussen de verschillende studies. Ten eerste concludeerden we dat inhoudsanalyses extra lastig zijn in het geval van latente inhoud. Latente inhoud is inhoud die niet zonder meer af te lezen is aan oppervlaktekenmerken van communicatie, en daarom vatbaar voor meerdere interpretaties. Analyse van latente content vereist niet alleen heldere definities van de te analyseren concepten, maar ook specifieke regels voor het omgaan met ambigue gevallen om tot een betrouwbare analyse te komen. Soms kunnen zulke regels ter bevordering van betrouwbaarheid de validiteit van een analyseschema in gevaar brengen. Verder bleek de vergelijking van verschillende studies op basis van de analyses zelf minder eenvoudig dan we verwachtten. Daarom hebben we de verschillende onderzoeksresultaten alleen vergeleken op basis van statistisch significante toetsuitkomsten.

Hoofdstuk 7 geeft een aantal suggesties voor verder onderzoek. Ten eerste zou het interessant zijn om de relatie tussen grounding en common ground te bestuderen. Dit proefschrift is impliciet uitgegaan van een positieve (curvi)lineaire relatie tussen grounding en common ground. Sommige van onze resultaten spreken deze aanname tegen. Een alternatieve hypothese zou zijn dat common ground bij negotiatie in eerste instantie lager wordt, en pas naarmate de negotiatie vordert leidt tot een toename van common ground. De tweede suggestie voor onderzoek betreft de aard van common ground en hoe common ground, gegeven die aard, te meten. In dit proefschrift is gebruik gemaakt van vrij conservatieve maten voor common ground, gericht op de expliciete en tastbare aspecten ervan, waarbij de meer efemere aspecten niet aan bod kwamen. Dit leidt tot de vraag of deze verschillende conceptualisaties (expliciet en tastbaar versus efemeer) eigenlijk verschillende constructen betreffen, en wat de consequenties daarvan zijn voor ons begrip van common ground. Ten derde heeft NTool zich bewezen in het laboratorium, en er is ook aangetoond dat het aangepast kan worden voor gebruik in de klas. Het zou interessant zijn te bestuderen hoe NTool verder kan worden ontwikkeld voor toepassing in uiteenlopender omgevingen. En ten vierde, geen van onze studies vond plaats in de praktische professionele context, waar verschillen in belang spelen die mogelijkerwijs het grounding-proces op nog onvermoede wijze beïnvloeden. Verder onderzoek aan het NTool-formalisme (niet noodzakelijkerwijs met NTool zelf) zou inzicht kunnen bieden in de niet-cognitieve en niet-taak-processen die grounding beïnvloeden.

Curriculum Vitae

PJ Beers (Pieter Jelle) was born 7 October 1977 in Deventer and grew up in Laren (Gld), the Netherlands. He attended primary and secondary school in Lochem and graduated in 1996. He studied Environmental Health Sciences at Maastricht University and obtained his Masters' degree in 2001 following a one-year internship with the International Centre for Integrative Studies, Maastricht University, on a thesis titled "Policy makers and learning from ICT-tools." He started as a PhD-student with OTEC, Open University of the Netherlands, on 15 August, 2001, on a project named "Knowledge sharing and knowledge building in expert teams with ICT", on which this thesis is based.

Publications

Journal Articles

Beers, P. J., Boshuizen, H. P. A., Kirschner, P. A., & Gijselaers, W. (submitted). The Analysis of Negotiation of Common Ground in CSCL.

Beers, P. J., Boshuizen, H. P. A., Kirschner, P. A., & Gijselaers, W. (submitted). Support of Grounding in a Classroom.

Beers, P. J., Boshuizen, H. P. A., Kirschner, P. A., Gijselaers, W. & Westendorp, J. (submitted). Measuring Effects of ICT-Tools.

Beers, P. J., Boshuizen, H. P. A., Kirschner, P. A., & Gijselaers, W. (submitted). Common gound, complex problems and decision making.

Beers, P. J., Boshuizen, H. P. A., Kirschner, P. A., & Gijselaers, W. H. (2005). Computer support for knowledge construction in collaborative learning environments. *Computers in Human Behavior, 21*(4), 623-643.

Kirschner, P., Strijbos, J.-W., Kreijns, K., & Beers, P. J. (2004). Designing electronic collaborative learning environments. *Educational Technology Research and Development, 52*(3), 47-66.

Beers, P. J., Van Asselt, M. B. A., Vermunt, J. D., & Kirschner, P. A. (2003). Policy, information and learning. *The Journal of Workplace Learning, 15*(2), 70-79.

Conference Papers

Beers, P. J., Boshuizen, H. P. A., & Kirschner, P. A. (2005, August). The analysis of negotiation of common ground in CSCL. In J. W. Strijbos & F. Fischer (Chairs), *Measurement challenges in collaborative learning research.* Symposium conducted at the 11th EARLI Conference, Nicosia, Cyprus.

Beers, P. J., Boshuizen, H. P. A., & Kirschner, P. A. (2005, August). NTool: Facilitating the negotiation of common ground in multidisciplinary teams. In P. J. Beers (Chair), *Multiple research perspectives on shared understanding.* Symposium conducted at the 11th EARLI Conference, Nicosia, Cyprus.

Kirschner, P. A., Beers, P. J., Boshuizen, H. P. A., & Gijselaers, W. H. (2005, August). Coercing shared knowledge in collaborative learning environments. In M. Bannert & J. Zumbach (Chairs), *Analyzing (self-)monitoring in computer assisted learning.* Symposium conducted at the 11th EARLI Conference, Nicosia, Cyprus.

Beers, P. J., Kirschner, P. A., Boshuizen, H. P. A., & Gijselaers, W. H. (2005, May - June). *Coercing knowledge construction in collaborative learning environments.* Paper presented at the Computer Supported Collaborative Learning 2005, Taipei, Taiwan.

Beers, P. J., Boshuizen, E., & Kirschner, P. (2005, May - June). CSCL-analyse en negotiatie van common ground [CSCL-analysis and negotiation of common ground]. In J. W. Strijbos (Chair), *Methodologische uitdagingen voor*

CSCL onderzoek [Methodological challenges for CSCL research]. Symposium conducted at the ORD 2005, Gent, Belgium.

Schurer, A., Beers, P. J., & Boshuizen, H. P. A. (2005, May - June). *De ondersteuning van common ground in multidisciplinaire teams [Support of common ground in multidisciplinary teams].* Poster presented at the ORD 2005, Gent, Belgium.

Beers, P. J., Boshuizen, H. P. A., & Kirschner, P. A. (2004, July). *Computer support for knowledge construction in collaborative learning environments.* Paper presented at the Special Interest Meeting of EARLI SIG 6 and 7, Tübingen, Germany.

Beers, P. J., Boshuizen, H., & Kirschner, P. (2004, June). NegotiatieTool: Een ICT-tool voor multidisciplinaire teams. In R. Martens (Chair), *Tools voor het uitwisselen van kennis: welke kennis kunnen we er over uitwisselen?* Symposium conducted at the 31st OnderwijsResearch Dagen (ORD), Utrecht, The Netherlands.

Beers, P. J., Boshuizen, H., & Kirschner, P. (2004, June). Samenwerkend leren, Common Ground, en het NegotiatieTool. In H. Boshuizen (Chair), *Effectiever samenwerkend leren - Een kruisbestuiving van tools en theorieën.* Symposium conducted at the 31st OnderwijsResearch Dagen (ORD), Utrecht, The Netherlands.

Westendorp, J., Beers, P. J., & Boshuizen, H. (2004, June). *De cognitieve belasting door grounding in teams.* Paper presented at the 31st OnderwijsResearch Dagen (ORD), Utrecht, The Netherlands.

Beers, P. J., Boshuizen, H. P. A., & Kirschner, P. (2004, April). Computer support for knowledge construction in collaborative learning environments. In P. A. Kirschner (Chair), *Learning in innovative learning environments.* Symposium conducted at the AERA, San Diego, California, USA.

Kirschner, P. A., Strijbos, J. W., Kreijns, K., & Beers, P. J. (2003, October). *Designing electronic collaborative learning environments.* Paper presented at the AECT 2003, Anaheim, USA.

Beers, P. J., Boshuizen, H. P. A., & Kirschner, P. A. (2003, August). Agreeing to disagree: Perspective, negotation and common ground in teams. In P. A. Kirschner (Chair), *The social psychological dimension of social interaction and the effects of cultural backgrounds in CSCL.* Symposium conducted at the 10th EARLI Conference, Padova, Italy.

Beers, P. J., Boshuizen, H. P. A., & Kirschner, P. A. (2003, August). Negotiating shared understanding in collaborative problem solving. In S. Järvela & F. Fischer (Chairs), *Awareness and shared understanding of distributed collaboration: Analysis and facilitation.* Symposium conducted at the 10th EARLI Conference, Padova, Italy.

Beers, P. J., Boshuizen, H. P. A., & Kirschner, P. A. (2003, May). Ondersteuning van negotiatie in teams. In S. Akkerman (Chair), *Het bereiken van een gemeenschappelijke kennisbasis in teams: Een zoektocht naar een gemeenschappelijk kader.* Symposium conducted at the 30th OnderwijsResearch Dagen (ORD), Kerkrade, The Netherlands.

Beers, P. J., Boshuizen, H. P. A., Kirschner, P. A., & Van den Bossche, P. (2002, August / September). *Decision-support and complexity in decision making.* Paper presented at the 5th Junior Researchers of EARLI (JURE) Conference, Amsterdam, The Netherlands, and at the 6th ICO Summerschool, Rethymnon, Greece.

Van den Bossche, P., Gijselaers, W., Segers, M., Beers, P. J., Kirschner, P., & Boshuizen, E. (2002, June). *Decision-making in complex environments, Multiple representations and multiple agents.* Paper presented at the EARLI Learning and professional development SIG meeting, Turku, Finland.

Beers, P. J., Van Merriënboer, J., Kirschner, P., Van den Bossche, P., Gijselaers, W., & Boshuizen, E. (2002, May). *Tools for supporting policy makers: An integrated theoretical framework.* Poster presented at the 29th OnderwijsResearch Dagen (ORD), Antwerp, Belgium.

Beers, P. J., Kirschner, P. A., Van den Bossche, P., & Gijselaers, W. (2002, April). *Decision-support and complexity in decision making.* Paper presented at the 3rd European Conference on Organisational Knowledge, Learning and Capabilities (OKLC), Athens, Greece.

Other Publications

Kirschner, P.A., Kreijns, K., Beers, P.J., & Strijbos, J.-W. (2005). Designing electronic tools for collaborative learning environments. *Educational Technology, 45*(5), 48-52.

Beers, P. J., Boshuizen, E., & Kirschner, P. (2004, November). *Kennisconstructie in multidisciplinaire teams: experimenteren met ICT.* Paper presented at the SURF Onderwijsdagen, Utrecht, The Netherlands.

Beers, P. J. (2001). *Policy makers and learning from ICT-tools.* Unpublished Master's Thesis, Maastricht University, Maastricht, The Netherlands.

Van Asselt, M. B. A., Mellors, J., Rijkens-Klomp, N., Greeuw, S. C. H., Molendijk, K. G. P., Beers, P. J., et al. (2001). *Building blocks for participation in Integrated Assessment: A review of participatory methods* (ICIS Working Paper No. I01-E003). Maastricht, The Netherlands: International Centre for Integrated Studies (ICIS).

Amelung, B., Beers, P. J., Rotmans, J., Van Asselt, M. B. A., & Hogervorst, J. (1999). *Global tourism, Towards an integrated approach* (ICIS Working Paper No. I99-E007). Maastricht, The Netherlands: International Centre for Integrated Studies (ICIS).

ICO Dissertation Series

In the ICO Dissertation Series dissertations are published of graduate students from faculties and institutes on educational research within the following universities: University of Twente, University of Groningen, Maastricht University, University of Amsterdam, Utrecht University, Open University, Leiden University, Wageningen University, Technical University of Eindhoven, and Free University (and formerly University of Nijmegen and University of Tilburg). Over one hundred dissertations have been published in this series. The most recent ones are listed below.

107. Veermans, K.H. (09-01-2003). *Intelligent Support for Discovery Learning: Using Opportunistic Learner Modeling and Heuristics to Support Simulation Based Discovery Learning*. Enschede: University of Twente.
108. Tjepkema, S. (10-01-2003). *The Learning Infrastructure of Self-Managing Work Teams*. Enschede: University of Twente.
109. Snellings, P.J.F. (22-01-2003). *Fluency in Second Language Writing: The Effects of Enhanced Speed of Lexical Retrieval*. Amsterdam: University of Amsterdam.
110. Klatter, E.B. (23-01-2003). *Development of Learning Conceptions during the Transition from Primary to Secondary Education*. Nijmegen: University of Nijmegen.
111. Jellema, F.A. (21-02-2003). *Measuring Training Effects: the Potential of 360-degree Feedback*. Enschede: University of Twente.
112. Broekkamp, H.H. (25-02-2003). *Task Demands and Test Expectations: Theory and Empirical Research on Students' Preparation for a Teacher-made Test*. Amsterdam: University of Amsterdam.
113. Odenthal, L.E. (21-03-2003). *Op zoek naar balans: Een onderzoek naar een methode ter ondersteuning van curriculumvernieuwing door docenten*. Enschede: University of Twente.
114. Kuijpers, M.A.C.T. (21-03-2003). *Loopbaanontwikkeling: onderzoek naar 'Competenties'*. Enschede: University of Twente.
115. Jepma, IJ. (01-07-2003). *De schoolloopbaan van risico-leerlingen in het primair onderwijs*. Amsterdam: University of Amsterdam.
116. Sotaridona, L.S. (05-09-2003). *Statistical Methods for the Detection of Answer Copying on Achievement Tests*. Enschede: University of Twente.
117. Kester, L. (05-09-2003). *Timing of Information Presentation and the Acquisition of Complex Skills*. Heerlen: Open University.
118. Waterreus, J.M. (05-09-2003). *Lessons in Teacher Pay: Studies on Incentives and the Labor Market for Teachers*. Amsterdam: University of Amsterdam.
119. Toolsema, B. (23-10-2003). *Werken met competenties*. Enschede: University of Twente.
120. Taks, M.M.M.A. (20-11-2003). *Zelfsturing in leerpraktijken: Een curriculumonderzoek naar nieuwe rollen van studenten en docenten in de lerarenopleiding*. Enschede: University of Twente.
121. Driessen, C.M.M. (21-11-2003). *Analyzing Textbook Tasks and the Professional Development of Foreign Language Teachers*. Utrecht: Utrecht University.

122. Hubers, S.T.T. (24-11-2003). *Individuele leertheorieën en het leren onderzoeken in de tweede fase*. Eindhoven: Technical University of Eindhoven.
123. Sun, H. (04-12-2003). *National Contexts and Effective School Improvement: An Exploratory Study in Eight European Countries*. Groningen: University of Groningen.
124. Bruinsma, M. (09-12-2003). *Effectiveness of Higher Education: Factors that Determine Outcomes of University Education*. Groningen: University of Groningen.
125. Veneman, H. (01-07-2004). *Het gewicht van De Rugzak: Evaluatie van het beleid voor leerlinggebonden financiering*. Groningen: University of Groningen.
126. Annevelink, E. (27-08-2004). *Class Size: Linking Teaching and Learning*. Enschede: University of Twente.
127. Emmerik, M.L. van (22-09-2004). *Beyond the Simulator: Instruction for High Performance Tasks*. Enschede: University of Twente.
128. Vries, B. de (15-10-2004). *Opportunities for Reflection: E-mail and the Web in the Primary Classroom*. Enschede: University of Twente.
129. Veenhoven, J. (05-11-2004). *Begeleiden en beoordelen van leerlingonderzoek: Een interventiestudie naar het leren ontwerpen van onderzoek in de tweede fase van het voortgezet onderwijs*. Utrecht: Utrecht University.
130. Strijbos, J.W. (12-11-2004). *The Effect of Roles on Computer-Supported Collaborative Learning*. Heerlen: Open University.
131. Hamstra, D.G. (22-11-2004). *Gewoon en Anders: Integratie van leerlingen met beperkingen in het regulier onderwijs in Almere*. Groningen: University of Groningen.
132. Lubbers, M.J. (09-12-2004). *The Social Fabric of the Classroom: Peer Relations in Secondary Education*. Groningen: University of Groningen.
133. Nijman, D.J.J.M. (10-12-2004). *Supporting Transfer of Training: Effects of the Supervisor*. Enschede: University of Twente.

Dewiyanti, S. (25-02-2005). *Learning Together: A positive experience. The effect of reflection on group processes in an asynchronous computer-supported collaborative learning environment.* Heerlen: Open University.

Stoof, A. (04-03-2005). *Tools for the identification and description of competencies*. Heerlen: Open University.

Groot, R.W.A. de (10-03-2005). *Onderwijsdecentralisatie en lokaal beleid*. Amsterdam: Universiteit van Amsterdam.

Salden, R.J.C.M. (22-04-2005). *Dynamic Task Selection in Aviation Training.* Heerlen: Open University.

Gijlers, A.H. (23-09-2005). *Confrontation and co-construction: Exploring and supporting collaborative scientific discovery learning with computer simulations.* Enschede: University of Twente.

Stevenson, M.M.C. (27-09-2005). *Reading and writing in a foreign language: A comparison of conceptual and linguistic processes in Dutch and English.* Amsterdam: University of Amsterdam.

Beers, P.J. (25-11-2005). *Negotiating Common Ground: Tools for Multidisciplinary Teams*. Heerlen: Open Universiteit Nederland.