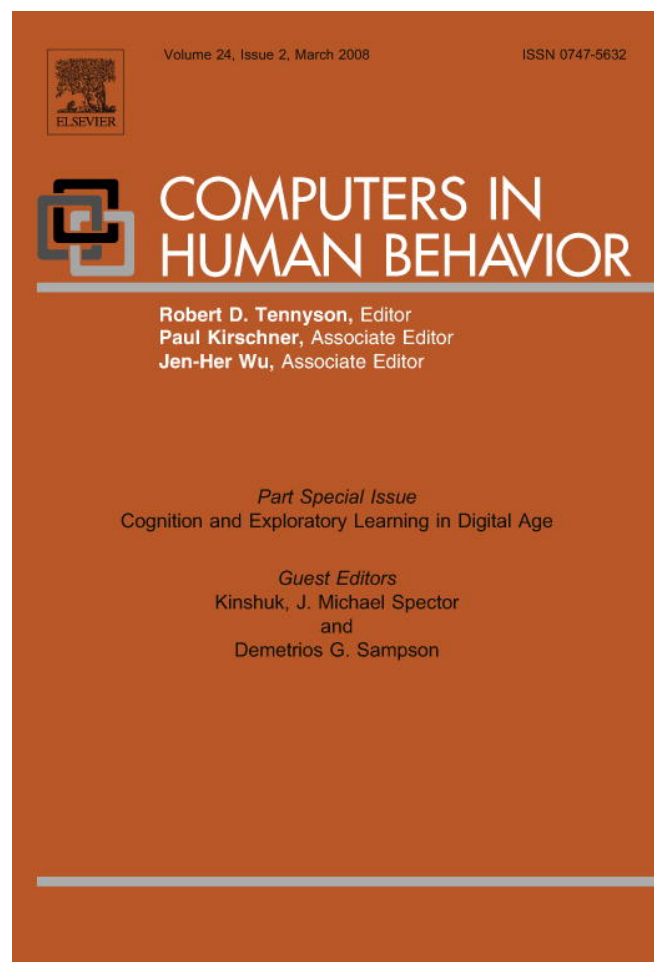


Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article was published in an Elsevier journal. The attached copy is furnished to the author for non-commercial research and education use, including for instruction at the author's institution, sharing with colleagues and providing to institution administration.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Coercing shared knowledge in collaborative learning environments

Paul A. Kirschner^{a,d,*}, Pieter J. Beers^{a,b}, Henny P.A. Boshuizen^a,
Wim H. Gijselaers^c

^a *Open Universiteit Nederland, The Netherlands*

^b *Delft University of Technology, The Netherlands*

^c *Maastricht University, The Netherlands*

^d *Utrecht University, P.O. Box 80140, 3508 TC, Utrecht, The Netherlands*

Available online 23 March 2007

Abstract

Multidisciplinary teams are used in industry, government, and education for solving complex problems because they allow different perspectives to be brought to bear on a problem and thus enrich the problem space. This, in turn, is expected to allow for rich problem analyses and solutions. However, multidisciplinaryity is not always advantageous. Good team solutions require team members to possess a good degree of common ground. To address this, researchers and educators often chose techniques such as collaboration scripts or scenarios to structure collaboration or how ICT-tools are used. They do this by making use of formalisms or constraints to structure conversation and discourse among collaborators with the aim of guiding the exchange of knowledge and information or both. Such techniques and tools have attained good results on cognitive aspects of group learning by focusing on task aspects. However, they have not explicitly addressed the problem of how teams with expertise diversity reach common ground. This article presents the results of a series of experiments that have shown that a tool that is capable of scripting the negotiation of both meaning and standpoint can have very positive effects on achieving common ground.

© 2007 Elsevier Ltd. All rights reserved.

* Corresponding author. Address: Utrecht University, P.O. Box 80140, 3508 TC, Utrecht, The Netherlands.
E-mail address: p.a.kirschner@fss.uu.nl (P.A. Kirschner).

1. Introduction

Ernst Rothkopf (1970, p. 325), paraphrasing a well-known saying, stated that “you can lead a horse to water but the only water that gets into his stomach is what he drinks”. What Rothkopf meant is that learning depends less on what teachers or instructional designers plan or want to have happen in learning situations than on what the learners themselves actually do. Central to this idea is that, just as in the case of the horse at a watering hole, what actually occurs is “a matter of choice on the part of the student. In relevant circumstances, students choose whether they will pay attention in lectures, read assignments, or review what has previously been read; rarely are these activities the only ones available” (Rhodes, 1993, p. 6).

The same goes for collaborative learning environments. Generally speaking, experience and research has shown that learners do not work or learn well in the collaborative learning environments that are designed and developed ‘for them’. Kirschner, Strijbos, and Kreijns (2003) accentuated this with respect to collaborative learning, speaking of a probabilistic as opposed to a causal learning model or instructional model. This is often the case for face-to-face environments, but is even more so for computer-supported collaborative learning (CSCL). To remedy this, teachers and instructional designers often make use of stopgap measures to get the students to work and learn together such as:

- Requiring a certain number of emails or contributions to discussion forums from each student in order for the student to be allowed to take an exam, complete and/or pass the course, or to acquire extra points towards a final grade.
- Active teacher or teaching-assistant involvement via moderation of the hard-to-achieve collaboration causing them to become very obvious and dominant in the environment. McFadzean and McKenzie (2001), for example, go as far as stating that tutor-facilitators should take of the job of steadily helping “learners through the process of clarifying and agreeing goals, exploring capabilities, and ensure feelings are widely expressed” (p. 479).

Why is this necessary? Why is it that these same students who do not, or will not, collaborate in CSCL-environments spend a great deal of their waking hours when not at school (primary or secondary education) or attending classes (higher education) emailing, MSN[®] or AOL[®] messaging, or Skyping[®] each other? Something is definitely wrong here either with the task, the social conditions, and/or the learning environment.

With respect to *tasks*, it is too often the case that the learning tasks are not suited to collaboration (Kirschner, 2005). They are often too closed (i.e., there is little room in the problem space to discuss), too easy (i.e., it can more efficiently be carried out by one person than by a team), or too controlled (i.e., there is little room for learner initiative). Pedagogically, thus, for collaborative learning to take place, the learning tasks need to be open, complex, and allow for student initiative. They need to be perceived as problems that are situated in real-life contexts (Brown, Collins, & Duguid, 1989) where the environment is rich in information and where there are no right answers, but many adequate or plausible ones. The tasks should be authentic, where meaning is negotiated through interactions with others, and where multiple perspectives exist (Von Glasersfeld, 1988).

A second problem is that the *social conditions* for collaboration are too often missing either because there is too much concentration on ‘on-task’ behavior in the environment (thus neglecting the fact that collaboration is very social) and/or the computer application used lacks tools that present opportunities to create a sound social space through awareness. A sound social space allows and/or stimulates the social conditions necessary for good collaboration such as acceptance, support, warmth, trust, and liking. Consequently, the necessary preconditions for good teamwork, based upon mutual interdependence and promotive interaction, do not exist. The result is a learning environment that is not sociable; that lacks the necessary affective work relationships, strong group cohesiveness, feelings of trust, respect, belonging and satisfaction, and strong sense of community needed for true collaboration (for an in-depth discussion see Kreijns & Kirschner, 2004; Kreijns, 2004).

Finally, and related to the previous problem, is that environments themselves are not suited to the specific collaborative working and learning required. By this we mean that they either have an *ontology* that does not match the problem or task at hand and/or they lack the tools that either *stimulate* or *coerce* the needed activities. With respect to ontology, there is often a mismatch between the ontology of the task and the ontology of the environment. We see this, for example, when an environment meant for problems/tasks in one domain such as the natural sciences – whose ontology concentrates on hypothesis generation and testing (e.g., Belvedere; Suthers, 2001) – is used for solving problems in domains with a completely different ontology such as the design domain (e.g., instructional or industrial design) or the argumentative domain (e.g., philosophy, law) and vice versa (Kirschner & van Bruggen, 2004; Van Bruggen, 2003).

Although there has been much research carried out on the first two problems, there has been very little research carried out on how these problems can be solved with the use of the specific technological affordances that CSCL-environments have available, and particularly with specific tools for coercing necessary collaborative activities.

2. Coercion

Merriam-Webster Online[®] defines “coerce” as meaning “to compel to an act or choice.” In these times of *politically correct speak*¹ – especially with respect to education and learning - coercion could be seen as something very negative. However, if we reflect on what teachers, parents, employers or whoever actually do, it is apparent that others are constantly ‘coerced’. We require or constrain certain acts or choices so as to shape or compel certain actions.

Many researchers have used information and communication technology (ICT) tools in their attempts to stimulate and sustain working in teams. These tools use *formalisms*, which is defined as sets of rules that, for example, structure conversation and discourse among collaborators so as to guide the exchange of knowledge and information. They are sets of moves or actions that users are allowed to make within a specific environment or situation. Specific formalisms have been tailored to influence specific aspects of problem

¹ The modern day equivalent of Newspeak from 1984 by George Orwell. Newspeak was the official language of Oceania, and had been devised to meet the ideological needs of Ingsoc, or English Socialism. The purpose of Newspeak was not only to provide a medium of expression for the world-view and mental habits proper to the devotees of Ingsoc, but to make all other modes of thought impossible.

solving, and a few ICT-tools have been developed to coerce (Dillenbourg, 2002) people to follow the rules of such formalisms.

Coercion, as used in this article, refers to the degree to which participants are required to follow a specific formalism. Another way of saying this is that it is the degree to which the participant is constrained in her or his behavior (e.g., actions or choice of actions) with using the formalism. Together, coercion and formalism constitute a *collaboration script*. The higher the script's coerciveness, the more participants are required to adhere to its formalism. Scripting requires "subjects... to make a particular type of speech act in a specific context" (Baker & Lund, 1997, p. 176). For Dillenbourg 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). A script that uses little or no coercion leaves participants many degrees of freedom and thus its usage attains a high degree of idiosyncrasy (i.e., learners use it as they see fit or understand it). A script with a high level of coercion severely constrains the number of options participants have, thus guiding them along the lines of the formalism and/or forcing them to use it as intended. The script as a whole acts as a *performance constraint* in that it makes actions unavailable that are not relevant to the instructed activities (Van Merriënboer & Sweller, 2005). In the research reported in this article, three types/levels of scripting were used.

3. Supporting negotiation for complex problem solving

This article focuses on coercing the negotiation of *common ground*, which can be seen as a shared cognitive frame of reference (Bromme, 2000). Many researchers agree upon the importance of common ground, for various reasons. For instance, learning occurs best when grounding forces co-learners to elaborate explanations (Webb, 1991) and negotiate decisions or meaning (Baker, Hansen, Joiner, & Traum, 1999). Boshuizen and Tabachneck-Schijf (1998) hold that both degree of common ground and co-operation (i.e., keeping track of what is said and making sure that the addressees should be able to understand the meaning) are of prime importance for collaboration of team members. They state that the "shared/non-shared dimension" – where shared representations are used and/or agreed upon by agents such that they are similar enough to be considered variants of the same representation – is crucial to understanding how teams deal with developing and sharing multiple representations when solving complex problems. Common ground, once achieved, can act as a shared interface between multiple representations among multiple actors (Boshuizen and Tabachneck-Schijf).

Bromme's (2000) theoretical account of the negotiation of common ground fits nicely in the work of Boshuizen and Tabachneck-Schijf (1998). Bromme theorizes about the causes of misunderstanding in the case of multiple representations, noting that people communicate on the basis of imperfect assumptions about the others' perspectives (Bromme, Rambow, & Nückles, 2001). Thus, one's perspective of the other affects one's own externalization of knowledge, and one's understanding of others' contributions, sometimes causing misunderstanding. Bromme introduces the concept of negotiation of common ground, in which participants iteratively make their own private understanding of the other explicit and providing feedback so as to reach common ground; a common cognitive frame of reference.

The negotiation of common ground is intrinsic to solving complex problems in multidisciplinary teams, because common ground is needed to afford the sharing of knowledge

and the subsequent construction of a shared problem representation. To do so, one must require problem solvers to explicitly verify their understanding of each other's contributions to a conversation and to explicitly articulate their own positions on those contributions.

4. Support framework

Beers, Boshuizen, Kirschner, and Gijsselaers (2005) hypothesized that unshared knowledge in one participant's head becomes newly constructed team knowledge via three intermediate forms and four processes (see Fig. 1).

In this 'cascade', negotiation of both meaning and position are fundamentally important for achieving common ground (Alpay, Giboin, & Dieng, 1998; Baker et al., 1999). Negotiation of meaning concerns making a private understanding of a contribution public, verifying whether and to what extent one's own understanding differs from what others intend, receiving feedback on this, re-verifying, etcetera. Negotiation of position concerns making a private opinion about a contribution public, checking whether one's own position is clear to others, and vice versa.

To support this, five support-principles for effective communication in teams that mimic the process of negotiation (see Table 1) were developed by Beers, Boshuizen, Kirschner, and Gijsselaers (2005). These support-principles require participants in the process of solving complex problems to follow certain rules.

Beers, Boshuizen, Kirschner, and Gijsselaers (2005) developed a set of rules that serves as a model for conversation (i.e., a formalism) for both contiguous and distributed groups to negotiate meaning and position to ultimately increase common ground. According to these rules, new conversation topics are introduced via 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 teammates, and can post *rejections* to messages that are unintelligible or objectively incorrect in their eyes. That formalism, aimed at making individual perspectives explicit to other team members, was developed, tested, and implemented in an ICT negotiation tool (NTool; see Fig. 2) which has been shown to be both valid, reliable and useful for

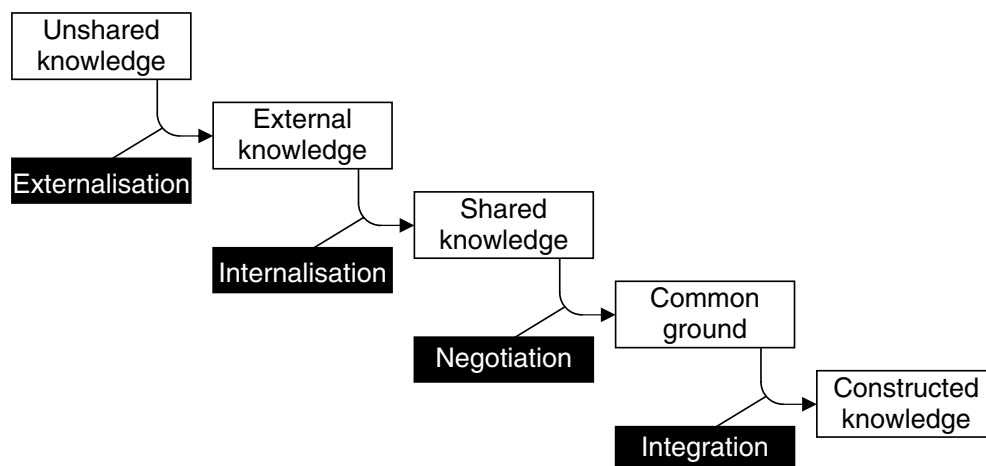


Fig. 1. From unshared knowledge to constructed knowledge.

Table 1
The NTool support principle

No.	Rule
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

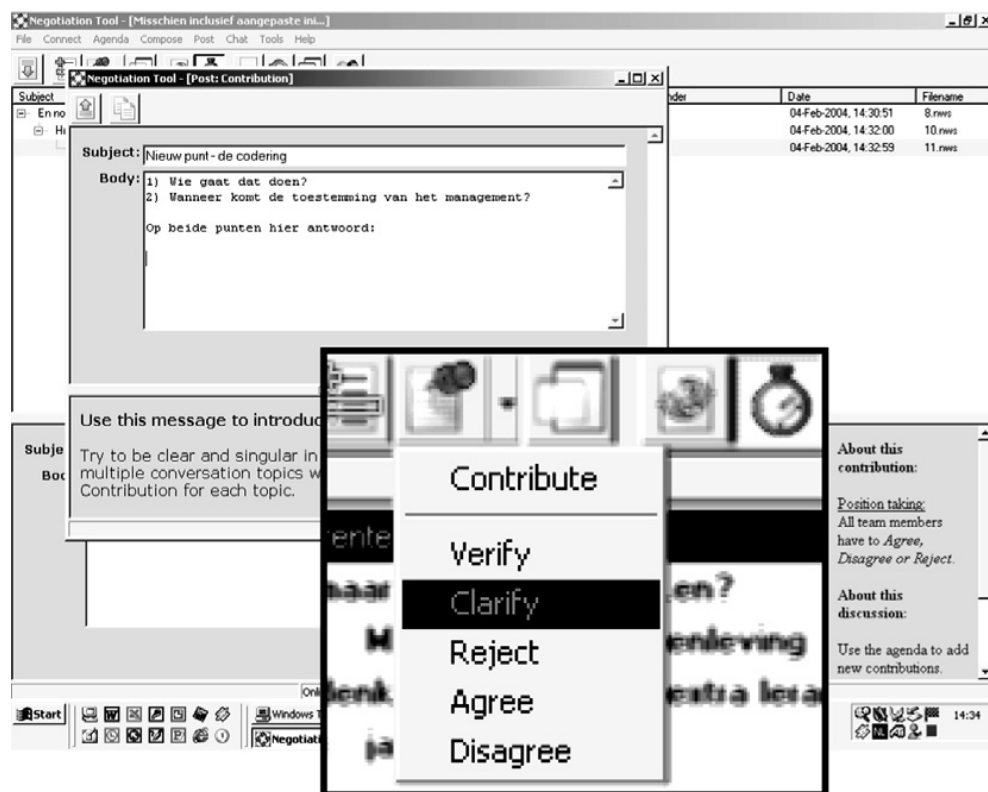


Fig. 2. NTool.

this purpose (Beers, Boshuizen, Kirschner, & Gijselaers, 2005; Beers, Boshuizen, Kirschner, Gijselaers, & Westendorp, 2007).

The rest of this article will deal with how NTool made the individual perspectives of the different team members explicit to the other team members and achieved shared knowledge in CSCL by coercing its users into negotiating common ground.

5. Method

The authors carried out four related empirical studies to determine whether it was possible to design and develop a tool that could facilitate the negotiation meaning and position while solving a complex problem in collaborative-learning environments. The research centered on answering the following research questions:

- Does coercion affect (negotiation of) meaning and position and thus facilitate teams in achieving shared knowledge and common ground?

and if so,

- How does the degree of coercion affect negotiation and common ground? and
- How does the degree of coercion disrupt discussions?

In the following sections, the general set-up of the studies will be discussed. First, the operationalization of the coercion is discussed. This is followed by a general discussion of the variables used and data collected in the studies. Finally, the most important results are discussed. For a full, in depth, discussion of the studies as well as of the methodological problems encountered and their solutions, the reader is referred to Beers, Boshuizen, Kirschner, and Gijssels (2005), Beers, Boshuizen, and Kirschner (2005); Beers, Boshuizen, Kirschner, and Gijssels (2007), and Beers, Boshuizen, Kirschner, Gijssels, and Westendorp (2007).

5.1. Scripting and coercion

Three levels or types of scripting were used. The first type was called *idiosyncratic*, a situation in which coercion is fairly low. Participants were supplied with the formalism and/or negotiation tool with specific message types (i.e., contribution, verification, clarification, agree, disagree, reject, accept), but without restrictions as to posting. In other words, they could use the different message types in the tool as they saw fit and in the order that they chose. They could first post 10 contributions and then discuss them one by one, completely discuss one contribution before posting a second, or even choose not to discuss or negotiate either meaning or position.

The second type was called *scripted* and used a medium level of coercion. Participants were again supplied with negotiation tool with the same specific message types, but were restricted in their use of it. This was done by ‘tweaking’ the tool to accept certain moves or acts or combinations of moves or acts, but to not accept others. Participants were required to negotiate meaning of all contributions first and then negotiate their positions. In other words, everyone needed to verify and choose to either accept or reject EACH of the contributions, but they did this by first negotiating the meaning of all of the contributions and then by negotiating their position on them.

The final type was called *stringent* and used a high level of coercion. Participants used the same tool, but it was now ‘tweaked’ such that they were required to verify each contribution – one contribution at a time – through a series of clarification–verification ‘moves’ before accepting or rejecting the contribution (i.e., negotiating the meaning of a specific contribution) and then agreeing or disagreeing with it (i.e., negotiating position on the contribution).

5.2. Variables and data

Common variables were used and comparable data were collected in all four studies. Two of the studies (1 and 3) were in face-to-face situations, while the other two were carried out in a CSCL-environment. Regardless of the setting, the same or similar variables were involved and similar data were collected.

5.2.1. Negotiation

The analysis goal was to identify aspects of communication that deal with negotiating common ground. To that end the same theoretical framework used for defining the formalism was used to characterize and measure negotiation. This strategy of using the same theory for both the design of CSCL and the analysis of its effects is described by Chi (1997). The main advantage of this approach is the relationship between theory and formalism. As stated earlier, the formalism used here is a set of rules and objects that together model conversation. 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 is based upon the combination of a linguistic approach to negotiation (Clark & Schaefer, 1989) and cognitivist one (Bromme, 2000). The linguistic approach describes what negotiation of common ground looks like in conversation, while the cognitivist approach focuses on how new knowledge is processed, and the role that prior knowledge plays in this process. The combination links the content of a collaborative learning process to the way people communicate with each other. The step from a theoretical framework of negotiation to a formalism for the support of negotiation resulted in the following earlier presented steps (see Table 1) as well as codes for cognitive conversation content directly related to solving the problem:

- *Contribution*: Introduction of a new topic of conversation.
- *Verification*: Direct or indirect request for information about the intended meaning of a contribution.
- *Clarification*: Reaction to a verification or a perceived lack of understanding in which the intended meaning of a contribution is elucidated.
- *Acceptance*: Reaction to a contribution in which the contribution is judged intelligible and/or correct.
- *Rejection*: Reaction to a contribution in which the contribution is judged unintelligible and/or incorrect.
- *Agreement*: Reaction to a contribution in which the sender voices his/her agreement with the contribution.
- *Disagreement*: Reaction to a contribution in which the sender voices his/her disagreement with the contribution.

As a result of the coding that actually took place (Beers, 2005), two extra categories were added that were not a part of the theoretical basis of the research, namely elaboration and regulation.

- *Elaboration*: During initial coding, it was found that people built upon each other's statements (i.e., cognitive content) without explicitly negotiating common ground. This resulted in the category *Elaboration*, defined as a statement that expands upon a contribution either by adding information to it or by summarizing it.
- *Task regulation*: When the coders noted that much conversation dealt with regulating what the participants were doing. Two regulative categories were added, namely *Monitoring* the problem solving-process (e.g., "I think we have not yet arrived at a good problem definition"), and *Regulation* of the conversation (e.g., "Could you make a note of that on the board please").

5.2.2. Common ground

Common ground was operationalised as the extent to which the content of each individual representation was present in other individual representations. To characterize the content of the individual representations, the discussion content itself was characterized (see Fig. 3). The discussion episodes identified earlier served as a basis for characterizing the discussion content.

Each episode was first numbered and summarized. Then, the content of all individual representations, both initial (pre-test) and subsequent to collaboration (post-test), and the group representation was characterized. The summaries were used to identify the content within the individual representations. For every individual representation, the topics that were and were not represented were assessed. In Fig. 3, for example, episode 7 is present in Jane's initial individual representation, in the group discussion, and in all post-tests. By repeating this procedure for all 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, the mean number of pre-tests and post-tests that a contribution was present was computed for each group. This mean number of post-tests per contribution was used as a measure of common ground.

5.2.3. Cognitive load

Any task that needs to be carried out, either individually or in a team brings with a certain amount of cognitive load (Sweller, 1988). Some of this load is intrinsic to the task itself and is determined by the complexity of the task itself (i.e., the number of elements that need to be processed by person or persons carrying out the task and the interactivity between the elements). Added to this intrinsic cognitive load are the load caused by processing activities not directly beneficial to carrying out the task at hand (extraneous load) and the load caused by processing activities that are beneficial (germane load). Since the team members have to both carry out the task AND work with NTool (including the learning of how to use it), it might be the case that the combination of these two activities

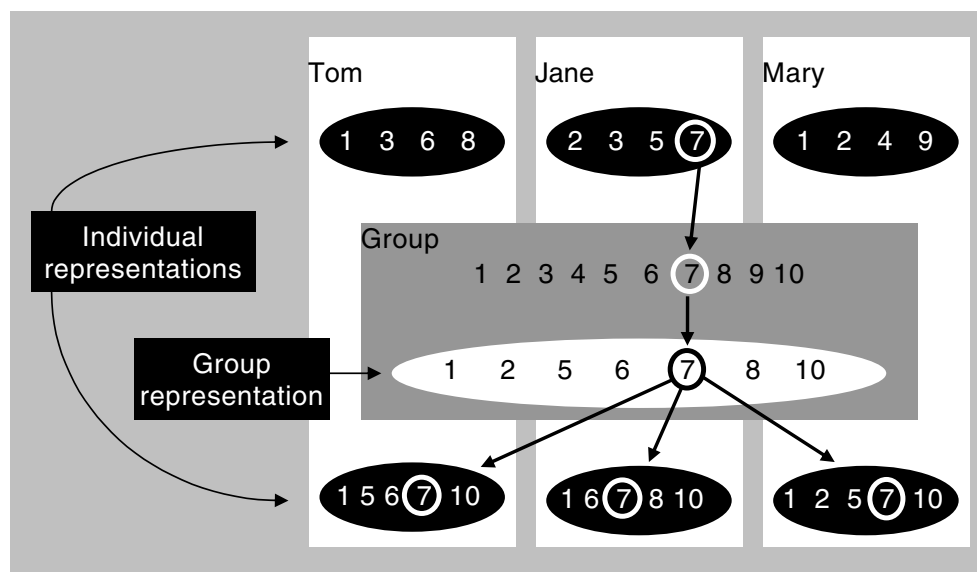


Fig. 3. Analysis of common ground; numbers indicate episodes.

could tax working memory to such an extent that the respondents are hampered in their primary activity (i.e., cognitive overload). To see whether this was the case, cCognitive load was measured in two studies through self-report of invested mental effort on a scale ranging from 1 (very, very low mental effort) to 9 (very, very high mental effort) (Paas, Tuovinen, Tabbers, & Van Gerven, 2003). Mental effort refers to the cognitive capacity actually allocated to solve the problem and can be considered to reflect the actual cognitive load (Sweller, Van Merriënboer, & Paas, 1998).

5.3. The four studies

All studies were variations on the same methodological theme, that is, they all involved three-person groups assigned the task of collaboratively solving a complex problem. Also, both prior to and after group assignment, participants had to solve the same problem individually. To promote construction of individual perspectives, as well as allow the researchers to determine participants' individual representations, participants first wrote individually down their solutions (pre-test). After this, they carried out the task collaboratively after which they were again asked to individually carry out the task (post-test). In the studies with cognitive load-measurements, load was measured three times, namely after the pre-test, after the team collaboration, and after the post-test. In the NTool practical studies, this process was carried out twice.

Two face-to-face studies were carried out to gain insight in 'raw' effects of the formalism without coercion; in other words, to test whether the formalism affected collaboration. Two subsequent studies used an ICT-implementation of the formalism that enabled the three levels of coercion. There were some differences between studies in terms of task and participants, but all studies are the same in that they involve three-person multidisciplinary teams solving the solution of a complex problem of interest to all disciplines involved (see Table 2).

The pen-and-paper studies used a whiteboard and flip-over as working aids. Half of the groups were instructed to use the formalism when working with a whiteboard and flip-over

Table 2
Studies, measurements and experimental conditions

	Conditions	Cognitive load	Task	Disciplines	N ^a
Face-to-face pilot 1	Idiosyncratic and Formalism groups		Investment decision	Marketing, Macro-economics, and Organisation Science	6
Face-to-face pilot 2	Idiosyncratic and Formalism groups	X	Solve high-school drop-out	Psychology, Economics, and Cultural Sciences	4
NTool laboratory study	Low, Medium, and High Coercion		Solve high-school drop-out	Psychology, Economics, and Cultural Sciences	17
NTool practical study	Low and High Coercion	X	1: Design a Floating House 2: "Amersfoort-by-the-Sea"	High-tech Metal-Electrics, Infrastructure, and Architecture	22

^a N is the number of groups, not individual participants.

to structure their collaboration (support principle). The other groups could use these materials any way they wanted to (no support principle). Participants in these studies were senior university students.

In the *NTool studies* (Beers, 2005; Beers, Boshuizen, Kirschner, & Gijsselaers, 2005; Beers, Boshuizen, & Kirschner, 2005) three versions of NTool with low, medium and high coercion (the idiosyncratic, scripted, and stringent condition, respectively) were developed, implemented, and tested. Before starting on the task, participants received instructions about how to use Ntool, about the formalism, and about the specific performance constraints (i.e., the scripting) during problem solving. Each team was supplied with three computers running NTool, one for each participant. The first NTool study was carried out in a laboratory setting with senior university students. The second NTool study took place in a ‘real’ educational setting, with second-year senior secondary vocational education students.

6. Results

6.1. Quality of negotiation

Table 3 shows all statistically significant results on negotiation.

In the first face-to-face study, formalism groups worked longer, discussed more contributions, and negotiated the contributions more thoroughly (see the amount of negotiation of meaning per contribution), than Idiosyncratic groups. The number of verifications and clarifications was higher in the formalism groups. The difference in number of clarifications was marginally significant ($p = 0.072$). Furthermore, in the formalism groups the mean number of participants per conversation-episode was significantly higher than in the Idiosyncratic groups, $\chi^2(2, N = 150) = 8.77, p < 0.05$. No other differences were statistically significant, although all of the observed differences were sizeable and in the expected direction.

Table 3
Main negotiation results

	Metric	Statistical test result
Study 1	Most clarifications in formalism groups	$U(N = 6) = 0.500, p = 0.072$
	Most participation per episode in formalism groups	$\chi^2(2, N = 150) = 8.77, p < 0.05$
Study 2	No statistically significant differences	
Study 3	Coercion correlated with Verification	$r_s(N = 17) = 0.63, p < 0.01$
	Coercion correlated with Clarification	$r_s(N = 17) = 0.54, p < 0.05$
	Most contributions in Idiosyncratic groups	$U(N = 17) = 4.00, p < 0.005$
	Most acceptance messages in Scripted groups	$U(N = 17) < 0.001, p < 0.005$
	Most regulation messages in Scripted groups	$U(N = 17) = 4.00, p < 0.01$
	Coercion correlated with negotiation per contribution	$r_s(N = 17) = 0.83, p < 0.0005$
Study 4	Most clarifications in Stringent groups	$F(1, 11) = 6.17, p < 0.05$
	Most regulation in Stringent groups	$F(1, 11) = 4.56, p = 0.06$
	Most ‘Other’ communication in Stringent groups	$F(1, 12) = 6.05, p < 0.05$
	Most negotiation per contribution in Stringent groups	$F(1, 11) = 8.17, p < 0.05$

In Study 2, no statistically significant differences were found with regard to negotiation. However, contrary to the first study, this time contributions were most heavily negotiated in the Idiosyncratic groups, formalism groups made more contributions. Furthermore, most negotiation took place in the Idiosyncratic groups, in terms of both verifications and clarifications. Finally, the Formalism groups regulated most. It appears that in the second face-to-face pilot differences were sizeable, but in the unexpected direction. On the whole, the data from the two face-to-face studies lend little credence to whatever hypothesis about negotiation and the formalism.

In the first NTool-study, NTool was shown to be increasingly effective with increasing coercion. Statistical testing (Kruskal–Wallis) revealed significant differences between the conditions for the number of contributions $\chi^2(2, N = 17) = 8.85, p < 0.05$, verifications, $\chi^2(2, N = 17) = 7.08, p < 0.05$, clarifications, $\chi^2(2, N = 17) = 7.33, p < 0.05$, acceptance messages, $\chi^2(2, N = 17) = 10.58, p < 0.01$ and regulation messages, $\chi^2(2, N = 17) = 8.03, p < 0.05$. Spearman correlations revealed significant correlations between coercion and verification, $r_s(N = 17) = 0.63, p < 0.01$, and coercion and clarification, $r_s(N = 17) = 0.54, p < 0.05$. Post hoc contrasting of Idiosyncratic groups with scripted and Stringent groups revealed a significantly higher number of contributions in the Idiosyncratic groups, $U(N = 17) = 4.00, p < 0.005$. Finally, contrasting of Scripted groups with Idiosyncratic and Stringent groups revealed a significantly higher number of acceptance $U(N = 17) < 0.001, p < 0.005$ and regulation messages $U(N = 17) = 4.00, p < 0.01$ 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. Finally, Kruskal–Wallis tests showed that the amount of negotiation of meaning per contribution differed significantly between the different versions of NTool, $\chi^2(2, N = 17) = 11.17, p < 0.005$. Coercion was significantly correlated with negotiation per contribution, $r_s(N = 17) = 0.83, p < 0.0005$. These results indicate that contributions were most heavily negotiated in the Stringent groups and least heavily in the Idiosyncratic groups.

In the second NTool-study, ANOVA revealed a significant main effect of coercion on verification, $F(1, 11) = 5.61, p < 0.05$; Stringent teams made more verifications than Idiosyncratic. Furthermore, a non-significant, but notable effect was found for coercion on regulation, $F(1, 11) = 4.56, p = 0.06$, 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, $F(1, 12) = 6.05, p < 0.05$; ‘other’ communication occurred most in the Stringent groups. Finally, there was a significant main effect of coercion on negotiation of meaning per contribution, $F(1, 11) = 8.17, p < 0.05$; negotiation per contribution was highest in the Stringent groups.

6.2. Common ground

Neither face-to-face study yielded statistically significant results. In the first, the number of discussion topics mentioned in the post-tests (overlap after collaboration) by all members was the same for both conditions, which means that no difference in common ground was found. In the second, Idiosyncratic groups achieved the most common ground after collaboration as shown by the content overlap in post-tests.

In the first NTool-study, no statistically significant differences were found for common ground with regard to pre-tests, $\chi^2(2, N = 17) = 1.78, p = 0.41$. The distribution of contri-

Table 4
Main common ground results

	Metric	Statistical test result
Study 1	No statistically significant differences	
Study 2	No statistically significant differences	
Study 3	Coercion correlated with common ground	$r_s(N = 17) = 0.57, p < 0.05$
Study 4	Most common ground in Idiosyncratic groups	$F(1, 12) = 8.46, p < 0.05$

Table 5
Main cognitive load results

	Metric	Statistical test result
Study 2	Most cognitive load in Formalism groups	$U(N = 12) = 7.00, p < 0.05^a$
Study 4	Most cognitive load in Stringent groups	$F(1, 52) = 3.67, p = 0.06$

^a One-tailed test.

butions across post-tests was significantly different between conditions, $\chi^2(2, N = 17) = 6.14, p < 0.05$. Subsequent Spearman correlation testing showed that the distribution of contributions across post-tests was significantly correlated with coercion, $r_s(N = 17) = 0.57, p < 0.05$ (see Table 4); the higher the coercion, the higher the number of post-tests a contribution would end up in.

Finally, in the second NTool-study, there was a significant effect of coercion on common ground, but not in the expected direction., $F(1, 12) = 8.46, p < 0.05$; common ground was highest in the Idiosyncratic groups.

6.3. Cognitive load

In the second face-to-face study, participants in the Scripted groups reported significantly more cognitive load after the post-tests, $U(N = 12) = 7.00, p < 0.05$, one-tailed. This suggests that Scripted collaboration required more processing in the post-test than Idiosyncratic collaboration. Likewise, in the second NTool-study a marginally significant main effect of coercion on cognitive load was found, $F(1, 52) = 3.67, p = 0.06$, suggesting that cognitive load was highest in the Stringent groups. In sum, it appears that overall cognitive load was higher in the Stringent groups than in the Idiosyncratic groups (see Table 5).

7. Conclusions

What can be concluded from these studies? The primary research question underlying the four studies was: How can collaborative construction of knowledge in multidisciplinary teams be supported so as to help the teams to better solve complex problems? The hypothesis underlying the research was that encouraging people to make their individual perspectives tangible to fellow team members would facilitate the negotiation of common ground resulting in more common ground. The formalism used in the research aimed at achieving this by requiring each team member to explicitly verify her or his own understanding of the contributions of other team members. Finally, it was assumed that the

more coercively the formalism was applied, the more effective it would be in terms of increasing grounding activity in communication and the amount of common ground a group would achieve.

Studies 1 and 2 yielded contradictory results, though not statistically significant. Study 3 – with an electronic NTool – showed that coercion is positively related to both the negotiation of and the amount of common ground. However, it also appeared that the specific way that the medium coercion condition (i.e., Scripted) was designed had unexpected disruptive effects on communication, as the users of this version used more regulation activities than users of the high and low coercion versions. Finally, Study 4 in a practical educational setting using the electronic NTool, revealed that high coercion with NTool resulted in the most negotiation activities. However, contrary to expectations, and also contrary to previous findings, common ground was highest in the non-coerced groups. The high coercion teams appear to have been more taxed than the non-coerced teams, as shown by increased ‘other’ communication and increased regulation.

In other words, coercion increases negotiation, but differentially affects common ground. At the university level, coercion did not appear to disrupt collaboration while this may have been the case at the secondary vocational school level. Coercion taxes the cognitive resources of the students (high CL), but this extra cognitive load is primarily germane (beneficial for learning) – something that education strives towards. Finally, the use of the formalism may have a longer lasting effect than an effect on the specific situation. That those students using the formalism chose to wait for a clarification shows that they might be getting used to presenting and requesting clarification of contributions when working together.

7.1. Negotiation and common ground

The experiments showed that the formalism affected negotiation activity, and that negotiation increased with coercion, confirming the stated hypothesis. Results of common ground differed between studies; in the NTool laboratory-study (Study 3) common ground *increased* with coercion, while common ground in the practical educational study (Study 4) *decreased* with coercion. The latter result is especially interesting because negotiation in that same study increased with coercion.

The difference between the NTool laboratory-study and the NTool 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 high degrees of both domain knowledge and workgroup experience. The participants in the practical educational study were considerably younger, less knowledgeable, and less experienced with working in groups.

Also, 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, resulting in increased 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 were affected by NTool, but not the underlying cognitive processes. This, however, 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 reorganize their own knowledge before they could integrate the others' knowledge. It is known that new knowledge can trigger reorganization 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 affect the processing of others' knowledge during collaboration, it may have been the case that the teams that negotiated most also needed to reorganize their prior knowledge more. As there was no additional time to do this during discussion, this reorganization must 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.

In sum, the studies showed that NTool and its accompanying formalism affect negotiation; the more coercively the formalism is applied, the more negotiation of common ground takes place. Furthermore, there is evidence that NTool can increase common ground, although this effect may be bound to specific user conditions, such as motivation and sophistication of prior knowledge.

7.2. An NTool effect mechanism

The main hypothesis included an assumption about a mechanism for NTool to affect communication, namely that explicit verification of understanding would make explicit the differences in perspective between group members. This, in turn would offer a starting point for reaching common ground. However, it was not clear on how NTool would achieve this. Though not reported here, exploratory interviews were carried out to study the actual mechanism through which the formalism affected grounding processes and to compare it with the hypothesized mechanism. This was done for both face-to-face studies (Studies 1 and 2) and the NTool laboratory study (Study 3).

The interview data suggested that the scripting did this by keeping participants from immediately voicing their opinions, allowing them to devote time to considering their fellows team members' contributions instead. This data supported the hypothesis that encouraging people to make their individual perspectives known and thus tangible – instead of implicit or tacit – to their fellow team members facilitates the negotiation of common ground. 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 allowed 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 immediately voice their opinion in reaction to a new contribution. Using NTool restrained them from doing so, which may have led to additional consideration and verification of contributions. Subsequent clarification may have changed their opinions from 'disagreeing' to 'accepting' and sometimes 'agreeing' upon a contribution. In sum, it seems that first verifying and clarifying contributions instead of immediately taking a

position increases the chances that contributions are accepted, and become part of common ground.

7.3. *The value of cognitive load measurements*

Cognitive load measurements were employed in two studies to determine whether coercion had adverse effects on communication that are not apparent when only measures of intended outcomes are used. These measurements made it possible to gain insight into whether high coercion taxed the users' cognitive capacities too much for an effect on grounding to take place. This would be the case, for example, when stringent use of the formalism resulted in significantly higher levels of cognitive load during collaboration. A comparison of the Stringent and Idiosyncratic versions in the secondary vocational education study (Study 4) showed that this indeed might have been the case for that specific population. No cognitive load measurements were taken in the laboratory NTool study (Study 3), but the data suggested that the Scripted version of NTool might have caused cognitive overloading for the population of senior year university students (based on increased regulation activity in that version), whereas this did not seem the case for the Idiosyncratic and Stringent versions.

Cognitive load measurements can also help detect differences in processing *after* the stringent formalism had been used which could suggest a lasting effect not attributable to extraneous effects of the coercion itself. Such effects were shown in both the face-to-face cognitive load study (Study 2) and the secondary vocational education study (Study 4). Although in the latter study results of cognitive load measurements seemingly contradicted 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 techniques and tools, while they are highly unobtrusive and easy to implement.

Acknowledgements

This research was supported and made possible by the Netherlands Organization for Scientific Research (NWO) within the stimulation program Society and the Electronic Superhighway under Grant No. 014-43-704.

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.
- Baker, M. J., & Lund, K. (1997). Promoting reflective interactions in a computer-supported collaborative learning environment. *Journal of Computer Assisted Learning*, 13, 175–193.
- 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.
- Beers, P. J. (2005). *Negotiating common ground: tools for multidisciplinary teams*. Unpublished PhD thesis, Open University of the Netherlands, Heerlen, The Netherlands.
- Beers, P. J., Boshuizen, H. P. A., Kirschner, P. A., & Gijssels, W. H. (2005). Computer support for knowledge construction in collaborative learning environments. *Computers in Human Behavior*, 21, 623–643.

- Beers, P. J., Boshuizen, H. P. A., & Kirschner, P. A. (2005). 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.
- Beers, P. J., Boshuizen, H. P. A., Kirschner, P. A., & Gijsselaers, W. (2007). Common ground, complex problems and decision making. *Group Decision and Negotiation*, 15, 529–556.
- Beers, P.J., Boshuizen, H.P.A., Kirschner, P.A., Gijsselaers, W. Westendorp, J. (2007). Measuring effects of ICT-tools. *Educational Technology Research and Development*, in press.
- 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., 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.
- Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18(1), 32–41.
- 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., & Schaefer, E. F. (1989). Contributing to discourse. *Cognitive Science*, 13, 259–294.
- 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.
- Kirschner, P. A. (2005). Learning in innovative learning environments. *Computers in Human Behavior*, 21, 547–554.
- Kirschner, P. A., Strijbos, J. W., & Kreijns, K. (2003). Designing integrated collaborative e-learning. In W. Jochems, J. van Merriënboer, & R. Koper (Eds.), *Integrated e-learning: Implications for pedagogy, technology & organization*. London: Taylor & Francis Books Ltd.
- Kirschner, P. A., & van Bruggen (2004). Learning and understanding in virtual teams. *Cyberpsychology and Behavior*, 7, 135–140.
- Kreijns, K., & Kirschner, P. A. (2004). Determining sociability, social space and social presence in (a)synchronous collaborating teams. *Cyberpsychology and Behavior*, 7, 155–172.
- Kreijns, K. (2004). *Sociable CSCL environments: social affordances, sociability, and social presence*. Unpublished PhD thesis, Open University of the Netherlands, Heerlen, the Netherlands.
- McFadzean, E., & McKenzie, J. (2001). Facilitating virtual learning groups: a practical approach. *The Journal of Management Development*, 2, 470–494.
- 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, 63–71.
- Rhodes, D. M. (1993). Revision of Johnson, H. C., Rhodes, D. M., & Rumery, R. E. (1975). The assessment of teaching in higher education: a critical retrospect and a proposal. *Higher Education*, 4, 173–199. Available at http://west.bradley.edu/TroupISU/Dent/drhodes_web/secure/docs/primary/jorhruat.pdf.
- Rothkopf, E. Z. (1970). The concept of mathemagenic activities. *Review of Educational Research*, 40, 325–336.
- Suthers, 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. (1988). Cognitive load during problem solving: effects on learning. *Cognitive Science*, 12, 257–285.
- Sweller, J., Van Merriënboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10, 251–295.
- 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 Merriënboer, J. J. G., & Sweller, J. (2005). Cognitive load theory and complex learning: recent developments and future directions. *Educational Psychology Review*, 17, 147–177.

- Von Glasersfeld, E. (1988). *Cognition, construction of knowledge and teaching*. (Eric Document Reproduction Service No. ED 294 754).
- Webb, N. (1991). Task-related verbal interaction and mathematics learning in small groups. *Journal of Research in Mathematics Education*, 22, 366–389.