Computer-Supported Collaborative Learning
DOI 10.1007/s11412-011-9110-3

Comparing the effects of representational tools in collaborative and individual inquiry learning

Bas Kolloffel · Tessa H. S. Eysink · Ton de Jong

Received: 6 June 2010 / Accepted: 9 January 2011

Abstract Constructing a representation in which students express their domain understanding can help them improve their knowledge. Many different representational formats can be used to express one's domain understanding (e.g., concept maps, textual summaries, mathematical equations). The format can direct students' attention to specific aspects of the subject matter. For example, creating a concept map can emphasize domain concepts and forming equations can stress arithmetical aspects. The focus of the current study was to examine the role of tools for constructing domain representations in collaborative inquiry learning. The study was driven by three questions. First, what are the effects of collaborative inquiry learning with representational tools on learning outcomes? Second, does format have differential effects on domain understanding? And third, does format have differential effects on students' inclination to construct a representation? A pre-test post-test design was applied with 61 dyads in a (face-to-face) collaborative learning setting and 95 students in an individual setting. The participants worked on a learning task in a simulation-based learning environment equipped with a representational tool. The format of the tool was either conceptual or arithmetical or textual. Our results show that collaborative learners outperform individuals, in particular with regard to intuitive knowledge and situational knowledge. In the case of individuals a positive relation was observed between constructing a representation and learning outcomes, in particular situational knowledge. In general, the effects of format could not be linked directly to learning outcomes, but marked differences were found regarding students' inclination to use or not use specific formats.

Keywords Inquiry learning · Simulations · External representations · Representational tools · Mathematics

Introduction

In collaborative inquiry learning, students are viewed as active agents in the process of knowledge acquisition. Collaborative inquiry learning unites two approaches: inquiry

B. Kolloffel (✉) · T. H. S. Eysink · T. de Jong
University of Twente, Enschede, The Netherlands
e-mail: b.j.kolloffel@gw.utwente.nl

learning and collaborative learning (Bell et al. 2010; Saab et al. 2007). In inquiry learning students learn through exploration and scientific reasoning. In an empirical comparison study, inquiry learning has been found to be among the most effective and efficient methods of active learning (Eysink et al. 2009). In collaborative learning two or more students construct knowledge together while they work towards the solution of a problem or assignment. Research has shown that collaboration between students can enhance learning (Lou et al. 2001; Slavin 1995; van der Linden et al. 2000). The combination of the two might lead to very powerful learning environments.

In (collaborative) inquiry learning, students investigate a domain by making observations, posing questions, collecting empirical data, organizing and interpreting the data in light of the posed questions, and drawing conclusions. This not only requires them to plan and execute inquiry processes, but also to select, process, analyze, interpret, organize, and integrate information into meaningful and coherent knowledge structures (Mayer 2002, 2004). Many things can and will go wrong in these processes unless students are provided with guidance and scaffolding during their inquiry process (de Jong 2005, 2006; de Jong and van Joolingen 1998; Quintana et al. 2004; Reiser 2004; Sharma and Hannafin 2007). Computer technology can support students and facilitate the inquiry learning process in many ways, for example by offering computer simulations for exploring, experimenting, and collecting empirical data (de Jong 2006; de Jong and van Joolingen 1998; Park et al. 2009; Rieber et al. 2004; Trundle and Bell 2010); tools for building and running dynamic models (Löhner et al. 2005; Sins et al. 2009; van Joolingen et al. 2005); tools for storing, Q1 editing, organizing, visualizing, and sharing data (Nesbit and Adesope 2006; Novak 1990; Suthers 2006; Suthers et al. 2008; Toth et al. 2002); and last but not least, tools for communication and exchanging information with others (e.g., chat tools, e-mail, online forums, message boards, threaded discussions) (Lund et al. 2007; Suthers et al. 2003). Collaboration can also fulfil a scaffolding function in inquiry learning. For example, during inquiry learning, students have to make many decisions (e.g., which hypothesis to test, what variables to change). In a collaborative setting, the presence of a partner stimulates students to make their plans and reasoning about these decisions explicit (Gijlers and de Jong 2009). Through externalization students express and explain ideas, ask for clarifications or arguments and generate new ideas or hypotheses. The process of expressing ideas through externalization and explanation stimulates students to rethink their own ideas and might even make them aware of possible deficits in their reasoning (Cox 1999; Kaput 1995; van Boxtel et al. 2000).

In the case of collaborative learning it is logical to think of speech or typed chat messages as primary media to externalize and explain ideas. Chat is a fast way of exchanging messages; talking in particular in a face-to-face setting is even faster, more elaborate, and richer in the sense that it provides both verbal and non-verbal information (e.g., gesturing, nodding, pointing, facial expressions, and intonation of speech) (Janssen et al. 2007; Strømsø et al. 2007; van der Meijden and Veenman 2005; van Drie et al. 2005). On the other hand, the speed of these media might sometimes be a disadvantage as well. Speech and chat are often fragmented, incoherent, jumping from one subject to another, and since they are volatile (speech more than chat) they do not lend themselves very much for reflection and consideration afterwards.

Another, more lasting way of externalizing and expressing ideas is by means of creating artefacts or models representing a domain or topic. This can for example be done in the form of writing a summary (Foos 1995; Hidi and Anderson 1986), creating a drawing (Van Meter et al. 2006; Van Meter and Garner 2005), building a runnable computer model (Löhner et al. 2003; Manlove et al. 2006), or constructing a concept map (Nesbit and Q2

Adesope 2006; Novak 1990, 2002). Furthermore, it should be noted that these activities are not reserved for collaborative learning settings only, but can just as well be applied in individual learning. Artefacts like these reflect the students' current overview and understanding of the domain, crystallize it as it were. It is open to be viewed, viewed again, discussed, elaborated, manipulated, and reorganized. But there might be an aspect that is even more important for learning. Externalizations show more than simply what students know and understand. Equally if not more important are the elements and aspects of the domain that are *not* represented, incorrectly represented or only partly represented. Externalization elicits self-explanation effects, and because the process of externalization requires students to go back and forth between their mental representations and the external representations they are constructing it can make them aware of unnoticed gaps and/or ambiguities in their mental representations (Cavalli-Sforza et al. 1994; Cox 1999; Kaput 1995). This in turn is important information that can be used to extend, refine and disambiguate their domain knowledge.

Representational tools: Tools for constructing externalizations

Computer technology can be used for creating and sharing externalizations. These tools are often referred to as *representational tools* (Suthers and Hundhausen 2003). Perhaps the most common example of a representational tool is the concept mapping tool (Novak 1990, 2002), but many other forms are available as well. Suthers and Hundhausen (2003) argue that in collaborative learning constructing external representations may form the pivot around which students share and discuss knowledge. Gijlers and de Jong (submitted) found that students who used a shared concept mapping tool in a collaborative simulation-based inquiry learning task showed significantly enhanced levels of intuitive knowledge compared to collaborating dyads that did not use a shared concept mapping tool. Intuitive knowledge is considered a quality of conceptual knowledge that taps on understanding how changes of one variable affect other variables (Swaak and de Jong 1996). Gijlers and de Jong (submitted) observed that in the concept mapping condition the intuitive knowledge scores were significantly and positively related to the percentage of chat messages related to conclusion and interpretation.

Effects of format on learning and communication

Representational tools can be used to store, display, manipulate, organize and share information, but also to support, scaffold, and even direct inquiry, communication, and knowledge construction processes. The representational format of a tool, also referred to as "notation" or "notational system" (e.g., Kaput 1995; Suthers 2008; Suthers et al. 2008; Wilensky 1995), can play a key role in learning. Kaput (1995) remarks: "different notation systems support dramatically different forms of reasoning, although the differences are strongly influenced by interactions between the knowledge structures associated with the notations and the prior knowledge to the reasoning" (p. 148). The properties of formats influence which information is attended to and how people tend to seek, organize and interpret information (e.g., Ainsworth and Loizou 2003; Cheng 1999; Larkin and Simon 1987; Zhang 1997). For example, constructing a concept map draws students' attention to Q3 key concepts in the subject matter and to the relations between those concepts (Nesbit and Adesope 2006) which can help students to enhance and refine their conceptual knowledge and understanding.

Suthers and Hundhausen (2003) compared three different formats of representational tools (concept maps, evidence matrix, and text) that were integrated in an electronic learning environment in which students explored a sequence of information pages about complex science and public health problems. It was found that pairs using an evidence matrix representation discussed and represented issues of evidence more than pairs using other representations. Second, pairs using visually structured representations (concept map, evidence matrix) revisited previously discussed ideas more often than pairs using text. Third, it was observed that the evidence matrix not only prompted novices to consider relevant relationships, but made them spend considerable time and resources on irrelevant issues as well.

van Drie et al. (2005) also compared three different formats of representational tools. They compared argumentative diagrams, lists, and matrices in a historical writing task in a computer-supported collaborative learning (CSCL) environment. It was found that matrices consisting of a table format that could be filled in by the students, supported domain-specific reasoning and listing arguments, whereas argumentative diagrams, organizing and linking arguments in a two-dimensional graphical way, made students focus more on the balance between pro and con arguments.

A study by Ertl et al. (2008) illustrates how pre-structuring a representational tool prompted the students' attention particularly to specific information that was relevant to the task. They used a task about Attribution Theory. Students were required to identify and name *causes*, to classify values of *consensus* and *consistency*, and to describe the *attribution* in students with school problems. Twenty-seven triads were provided with a representational tool, twenty-six triads did not have a representational tool. The tool consisted of a content scheme, that is, a table in which causes, consensus, consistency, and attribution could be filled in by the students. It was found that triads provided with a scheme, scored higher with respect to determining consensus, consistency, and attribution. This study suggests that the effects of a representational tool can depend to a large extent on the mapping between the tool on the one hand and the goals and aims of the learning task on the other hand. In this domain, causes, consensus, consistency, and attribution were the main aspects.

Do representational tools always work?

Formats can have different affordances, not only in the sense that they focus the attention of students on different aspects of the subject matter, but also with regard to how accessible or easy to use they are for students. Format can therefore play a critical role in the likelihood that students engage in constructing a representation and use a representational tool as intended. Kolloffel et al. (2010) studied the effects of representational tools used by individual students in a learning environment about combinatorics and probability theory. Three different formats of representational tools were tested: a concept mapping tool, a tool for creating arithmetical representations (e.g., formulas, equations), and a textual representational tool, which resembled simple word processing software. Each of the tools was integrated in a simulation-based inquiry learning environment. It was found that students who used a representational tool showed significantly higher post-test scores, and they also showed enhanced levels of situational knowledge, which is a prerequisite for going beyond the superficial details of problems. Furthermore, when students were provided with a conceptual or textual representational tool they were much more likely to construct representations than when provided with a representational tool with an arithmetical format.

In a similar sense is offering tools and scaffolds not a guarantee that the learning outcomes improve. Clarebout and Elen (2006; see also: Clarebout et al. 2009) pointed out that tools are often used inadequately or not at all by students. Inadequate use of tools is for example using a tool “to gather but not organize or synthesize problem-related information” (Jiang et al. 2009). They argue that the likelihood that students will use a tool depends on a complex interplay of factors, including (but not limited to) prior knowledge (high or low, both can stimulate or inhibit tool-use), motivation and goal orientation, self-regulation strategies, and domain-related interest (Jiang et al. 2009).

Talking or chatting about the subject matter in collaborative inquiry learning can be seen as a way of externalizing and expressing knowledge. Yet, tools can be useful to direct the attention of students toward specific aspects of the domain that might be overlooked otherwise.

Research questions

The focus of the current study was to examine the role of representational tools in collaborative inquiry learning. The study was driven by the following questions. First, what are the effects of collaborative inquiry learning with representational tools on learning outcomes? Second, does format of the tool have differential effects on domain understanding? And third, does the format of the tool have differential effects on students' inclination to use a representational tool?

In the current study the format participants could use to construct a representation was experimentally manipulated. Three representational tools were developed, each designed in such a way that it constrained the format that could be used to construct a representation. One tool allowed only conceptual input, another one allowed only arithmetical input, and a third one could only be used to create texts (these tools will be described in more detail in the Method section).

In order to gain a more full appreciation of the collaborative aspect in this study, the results were contrasted to a twin study reported earlier (see Kolloffel et al. 2010), that took place in an individual inquiry learning setting. In the collaborative inquiry learning setting the students communicated face-to-face with each other. Following existing literature on the comparison learning outcomes in individual and collaborative learning settings (e.g., Lou et al. 2001; Slavin 1995; van der Linden et al. 2000), it was hypothesized that learning outcomes in the collaborative learning setting would be higher than for those in the individual learning setting.

The format used to construct a representation was assumed to have differential effects on knowledge construction and domain understanding. Creating a *conceptual* representation like a concept map was hypothesized to point the students' attention at the identification of concepts and their relationships (Nesbit and Adesope 2006). A concept map is relatively easy to construct, especially if there are not too many concepts and relations (van Drie et al. 2005). Because this format is easy to understand and use, it was assumed that participants would be inclined to use it. The focus of students on the domain concepts was hypothesized to result in enhanced levels of knowledge about the conceptual aspects of the domain, rather than procedural or situational aspects.

Constructing representations in an *arithmetical* format was assumed to direct the students' attention mainly towards procedural domain aspects (e.g., the ability to calculate the probability of an event). Therefore, it was hypothesized that constructing an arithmetical representation would foster the acquisition of procedural knowledge rather than knowledge about conceptual and situational aspects. Regarding the likelihood that students would construct a representation, it was hypothesized that compared to other formats students would have difficulty constructing arithmetical representations (cf. Tarr and Lannin 2005), however, discussing the arithmetical aspects of the domain with a peer in a collaborative learning setting could have a beneficial effect on students' inclination to use the arithmetical tool.

The third format for constructing a domain representation was a *textual* format. This format particularly allows students to express their knowledge in their own words. The current domain could easily be described in terms of everyday life contexts and situations. Constructing textual representations was assumed to direct the student's attention to situational and conceptual aspects, although the textual format was not expected to emphasize domain concepts as strongly as the concept maps were supposed to do. It was expected that students would not experience much difficulty with using the textual format. Overall, this is one of the most commonly used formats inside and outside educational settings. Therefore, it was assumed that many participants would be inclined to use this representational tool.

Method

Participants

In the *collaborative learning study*, 128 secondary education students entered the experiment. In total, the data of 61 pairs could be analyzed. The average age of these 56 boys and 66 girls was 14.62 years (SD=.57). In the twin study, the *individual learning study*, 95 secondary education students, 50 boys and 45 girls, participated (Kolloffel et al. 2010). The average age of the students was 14.62 years (SD=.63). All data were collected in two subsequent years in the same school with the same teachers and the same method. The experiments employed a between-subjects design with the format of the provided representational tool (conceptual, arithmetical, or textual) as the independent variable. Students were randomly assigned to conditions. Of the 61 pairs in the collaborative setting, 22 pairs were in the Conceptual condition, 19 pairs in the Arithmetical condition, and 20 pairs in the Textual condition. Of the 95 students in the individual learning setting, 33 were in the Conceptual condition, 30 in the Arithmetical condition, and 32 in the Textual condition. The domain of combinatorics and probability theory was part of the regular curriculum and both experiments took place some weeks before this subject would be treated in the classroom. The students attended the experiment during regular school time; therefore, participation was obligatory. They received a grade based on their post-test performance.

Domain

The instruction was about the domain of combinatorics and probability theory, which involves determining how many different combinations can be made with a set of elements and the probability that one or more combinations will be observed in a random experiment. Some of the key concepts in this domain are *replacement* (are elements allowed to occur more than once in a combination?) and *order* (is the specific order of elements in a combination relevant information?). On basis of these two concepts, four so-called problem categories can be distinguished (replacement and order relevant; no replacement and order relevant; and so on). An example of a problem which comes under the category "replacement and order relevant" is the following: what is the probability that a thief will guess the 4-digit PIN-code of your credit card correctly in one go?. It is possible that a digit is observed more than once in a code (replacement). Second, it is necessary but not sufficient to know which four digits comprise the code because one also needs to know the specific order in which the digits appear in the code (order relevant).

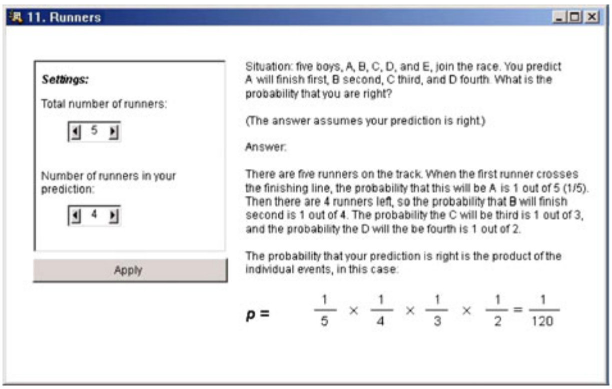
Learning environment

The instruction about the domain was implemented in a simulation-based inquiry learning environment, called Probe-XMT, which was created with SIMQUEST authoring software (de Jong et al. 1998; Swaak and de Jong 2001). .Computer simulations can be used by students to inquire into a domain. The simulation displays a state or situation of the domain and some of the elements or variables that play a role in that domain can be changed by the user. Each time the user makes a change, the simulation shows the effects of the change on the state or situation. The idea behind this instruction is that by systematically changing variables and observing the consequences of those changes, the students can explore and learn to master the key concepts and principles of the domain (de Jong 2005, 2006; de Jong and van Joolingen 1998). An example of a computer simulation in Probe-XMT is displayed in Fig. 1.

The simulation in Fig. 1 is about predicting the outcome of a footrace. Relevant variables here are for example the total number of runners and the range of the prediction (e.g., predicting only the winner, or the top 3, or the top 10, and so on). In the box on the left-hand side of the simulation, students could enter the values of those variables. On the right-hand side of the simulation the resulting effects of the values on number of possible combinations and the probability that a certain prediction would be true could be observed. In this case, this consisted of a text and an equation that changed whenever the values of the variables were changed. In an earlier study, the combination of text and equations was found to have computational benefits and benefits in terms of learning outcomes compared to other formats, e.g., tree diagrams (Kolloffel et al. 2009).

Probe-XMT consisted of five sections (not displayed in Fig. 1). Four of these sections were devoted to each of the four problem categories. The fifth section aimed at connecting and integrating these four problem categories. Each section used a different cover story, that is, an everyday life example of a situation in which combinatorics and probability played a role, exemplifying the problem category treated in that section. The example of the footrace (see above) was used as cover story for problem category "no replacement; order relevant".

Fig. 1 Screen dump Probe-XMT simulation



The example of the thief and the credit card was used as a cover story in the “replacement; order relevant” section. In the fifth (integration) section, the cover story applied to all problem categories. In each section the students’ inquiry activities were guided by a series of questions (both open-ended and multiple-choice items) and assignments, all based on the cover story of that particular section. Information about user actions in the learning environment, including time-on-task, path through the learning environment, and simulation use, were registered in log files.

Representational tools

For this study an electronic on-screen representational tool was added to the learning environment Probe-XMT. This tool could be used to construct an overview or summary of the domain’s main concepts, principles, variables, and their mutual relationships. Depending on the experimental condition to which a participant was assigned, the format of this tool was either conceptual, or arithmetical, or textual. (This will be explained in more detail later). In each condition the tool was available at all times in the learning environment and therefore the participants could use it any time they wanted during their learning process. Operating the tool was easy and straightforward. Participants received a demonstration of how to use the tool beforehand and there was plenty of time to practice using the tool before the actual experiment started. Furthermore, during the experiment help and assistance with using the tools was available at all times.

As mentioned before, the experimental manipulation focused on the format of the representational tool. There was a tool with a conceptual format, a tool with an arithmetical format, and a tool with a textual format. The *conceptual representational tool* (see Fig. 2) could be used to create a concept map of the domain. Students could draw circles representing domain concepts and variables. Keywords could be entered in the circles. The circles could be connected to each other by arrows indicating relations between concepts and variables. The nature of these relations could be specified by attaching labels to the arrows.

In the *arithmetical representational tool* (see Fig. 3), students could use variable names, numerical data, and mathematical operators (division signs, equation signs, multiplication signs, and so on) in order to express their knowledge.

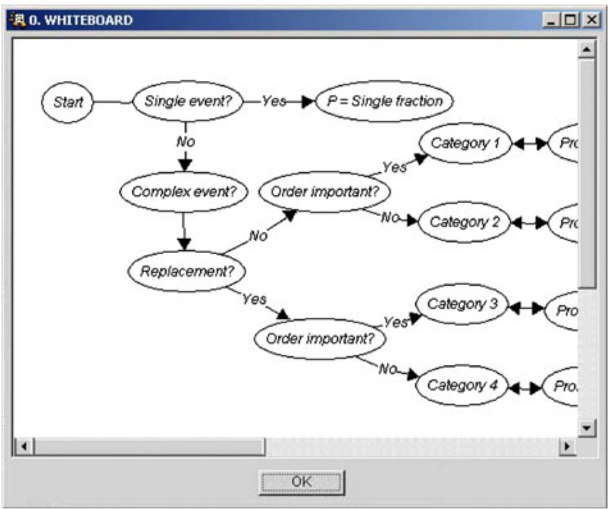


Fig. 2 Conceptual representational tool (Concept map created by participants)

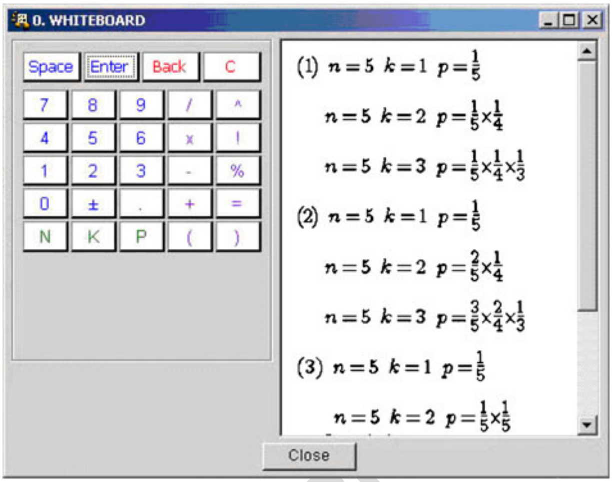


Fig. 3 Arithmetical representational tool (Input on the right side created by participants)

Finally, the *textual representational tool* (see Fig. 4) resembled simple word processing software, allowing textual and numerical input.

In theory, the participants could have used paper and pencil to “bypass” the representational tool. Experimenters were present in the classroom at all times and this behavior was not observed. Participants were focused on the computer screen, meanwhile talking with each other about the subject matter, assignments, navigation, and so on. No artefacts were created outside the electronic learning environment. In the current study, the effects of representational tools were tested outside the lab, in real classroom settings. The tools were intended as means to *support* students while learning, not as means to *assess* learning. Assessment is mostly obligatory in classroom settings, whereas making use of support is not. For reasons of ecological validity, the use of the representational tool was therefore not obligatory, although students were strongly advised to use the tool and they were informed that using the tool would help them to better prepare themselves for the post-test.

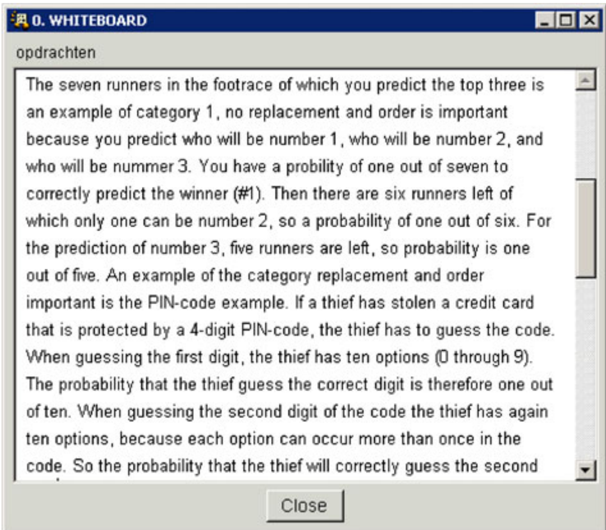


Fig. 4 Textual representational tool (Text created by participants)

Knowledge measures

Two knowledge tests were used in this experiment: a pre-test and a post-test. The tests contained 12 and 26 items respectively. The sensitivity and reliability of the test items have been established in recent years in a number of studies performed across Germany and The Netherlands (see e.g., Berthold and Renkl 2009; Eysink et al. 2009; Kolloffel et al. 2009; Gerjets et al. 2009; Wouters et al. 2007). The pre-test was aimed at measuring (possible differences in) the prior knowledge of the students. The post-test was specifically designed to measure the effects of external representations on domain knowledge. Well-structured and organized mathematical knowledge is thought to include conceptual, intuitive, procedural, and situational understanding (e.g., Fuchs et al. 2004; Garfield and Ahlgren 1988; Hiebert and Lefevre 1986; Rittle-Johnson and Koedinger 2005; Rittle-Johnson et al. 2001; Sweller 1989). The post-test consisted of different types of items, each aimed at measuring one of these types of knowledge.

Conceptual knowledge is the implicit or explicit understanding of principles underlying and governing a domain and of the interrelations between pieces of knowledge (Rittle-Johnson et al. 2001) developed by establishing relationships between pieces of information or between existing knowledge and new information. The post-test contained 12 multiple choice items aimed at measuring conceptual knowledge. Four of these items were intended to measure regular conceptual knowledge (see Fig. 5 for an example).

Eight items were intended to measure *intuitive conceptual knowledge* (see Fig. 6 for an example). Intuitive conceptual knowledge reflects the extent to which conceptual understanding has become automated. The idea behind intuitive conceptual knowledge is that as students' conceptual understanding becomes deeper and more automated, this will increase the speed with which they can assess concepts and their relations in problem situations and also enable them to accurately predict how these concepts and relations will respond to changes. Items measuring conceptual knowledge and intuitive conceptual knowledge differed in three respects (Eysink et al. 2009): first, the situation described in the problem statement regarding the intuitive items was the same for each item and was presented prior to the items instead of being presented with each separate item; second, the intuitive items offered two alternatives instead of four; finally, students were asked to answer the intuitive items as quickly as possible, as intuitive knowledge is characterized by a quick perception of the meaningful situation (Swaak and de Jong 1996).

Procedural knowledge is "the ability to execute action sequences to solve problems" (Rittle-Johnson et al. 2001, p.346). The post-test contained 10 open-ended items aimed at measuring procedural knowledge (see Fig. 7 for an example).

You have a deck of cards from which you select 4 cards. You predict that you will select an ace, a king, a queen and a jack in this specific order. Does it matter whether you put back the selected cards before each new selection or not?

a. Yes, your chances increase when you put back the selected cards
b. Yes, your chances decrease when you put back the selected cards
c. No, your chances remain the same whether you put back the selected cards or not
d. This depends on whether the deck of cards is complete or not

Fig. 5 Post-test item measuring conceptual knowledge

(Answer the following question(s) as quickly as possible)

There are a number of marbles in a bowl. Each marble has a different color. You will pick at random (e.g., blindfolded) a number of marbles from the bowl, but before you do you predict which colors you will pick.

The chance your prediction proves to be correct is higher in case of:

a. No replacement; order not important
b. Replacement; order important

Fig. 6 Post-test item measuring intuitive conceptual knowledge

Situational knowledge (de Jong and Ferguson-Hessler 1996) enables students to relate a problem to everyday, real-life situations, and to analyze, identify, and classify a problem, to recognize the concepts that underlie the problem, and to decide which operations need to be performed to solve the problem. Four multiple-choice items were included in the post-test to measure this type of knowledge (see Fig. 8 for an example).

The correct answers to the items presented in Figs. 5, 6, 7, and 8, are respectively: answer B; answer A; (1/10)*(1/10)=1/100; and answer A.

Procedure

The experiments were performed in three sessions all separated by a one-week interval, and took place in a real school setting. The procedures in both the individual and collaborative setting were identical.

In session one, students received some background information about the purpose of the study, the domain of interest, learning goals, and so on. This was followed by the pre-test. In both the individual and the collaborative setting, students completed the pre-test individually. It was announced that the post-test would contain more items of greater difficulty than the pre-test, but that the pre-test items nonetheless would give an indication of what kind of items to expect on the post-test. At the end of the pre-test, a printed introductory text was handed out to the students in which the domain was introduced. The duration of the first session was limited to 50 min. During the last 15 min of the session, the students received an explanation of how their representational tool could be operated and they could practice with the tool.

A week later, in session two, the students worked with the learning environment and had to construct a domain representation using a representational tool. The duration of this session was set at 70 min. Students in the individual learning setting worked alone. In the collaborative learning setting students were allowed to choose their partner themselves. Communication between students was on a face-to-face basis: the collaborating students were sitting next to each other, using the same computer terminal. They worked together on the assignments, simulations, and the representational tool in the learning environment.

In a pop music magazine you see an ad in the rubric FOR SALE in which a ticket for a spectacular concert of your favorite pop group is offered. Unfortunately the last 2 digits of the telephone number, where you can obtain information about the ticket, are not readable anymore. You really like to have the ticket and decide to choose the 2 digits randomly. What is the probability that you dial the correct digits on your first trial?

Fig. 7 Post-test item measuring procedural knowledge

Fig. 8 Post-test item measuring situational knowledge

You throw a dice 3 times and you predict that you will throw 6-4-2 in that order. What is the characterization of this problem?

a. order important; replacement
b. order important; no replacement
c. order not important; replacement
d. order not important; no replacement

Despite the possibility of following a non-linear path through the learning environment, students were advised to keep to the order of sections and assignments because they built upon each other.

The third session was set at 50 min. First, students were allowed to use the learning environment for 10 min in order to refresh their memories with regard to the domain. Then all students had to close their domain representations and learning environments, and had to complete the post-test. In both the individual and the collaborative setting, students completed the post-test individually.

Data preparation

A scoring rubric (see Appendix) was used to assess whether the domain representations constructed by the students reflected the concepts of replacement and order, presented calculations, referred to the concept of probability, indicated the effect of size of (sub)sets on probability, and the effects of replacement and order on probability. The scoring rubric was designed in such a way that all types of representations could be scored on the basis of exactly the same criteria. The maximum number of points that could be assigned on the basis of the rubric was eight points.

Results

Prior knowledge

Two measures of prior knowledge were obtained, a pre-test score and math grade. The reliability, Cronbach's α, of the pre-test was .40 in the individual setting and .48 in the collaborative setting. The pre-test reliabilities were rather low, but sufficient for the purpose of verifying that students did not have too much prior knowledge and that there were no differences between settings and/or conditions. Second, students were asked for their latest school report grade in mathematics. This grade, which can range from 1 (very, very poor) to 10 (outstanding) was interpreted as an indication of the student's general mathematics achievement level. It should be noted that this measure was reported by the students themselves and since no data from the school regarding math grades was available to the experimenters, the accuracy and reliability of the reported math grades should be considered with care. In Table 1 math grade and pre-test measures are presented.

Three-way ANOVAs with setting (individual or collaborative), format (Conceptual, Arithmetical, Textual), and tool-use (Tool-use or No-tool-use) as factors were performed to test for a priori differences with respect to math grade (general mathematics achievement level) and pre-test score (prior knowledge). A difference regarding *math grade* was observed with respect to setting, $F(1,205)=5.37$, $p<.05$, and tool-use, $F(1,205)=6.97$, $p<.01$. Furthermore, an interaction between setting and tool-use, $F\ (1,205)=7.24$, $p<.01$,

Table 1 Math grade and pre-test measures

	Representational format							
	Conceptual (indiv. n=33) collab. n=44)		Arithmetical (indiv. n=30) collab. n=38)		Textual (indiv. n=32) collab. n=40)		Total (indiv. n=95) collab. n=122)	
	M	SD	M	SD	M	SD	M	SD
MATHGRADE (max. 10)								
Individual (Total)	6.46	1.61	5.89	1.55	6.25	1.54	6.21	1.57
Tool-use	7.12	1.52	6.62	1.63	6.73	1.37	6.89	1.46
No-tool-use	5.75	1.41	5.71	1.51	5.82	1.58	5.75	1.48
Collaborative (Total)	6.56	1.32	6.98	1.20	6.80	1.00	6.77	1.19
Tool-use	6.75	1.33	6.75	1.49	6.82	1.10	6.75	1.09
No-tool-use	6.40	1.32	7.04	1.14	6.79	0.93	6.84	0.94
PRE-TEST (max. 12)								
Individual (Total)	5.70	1.36	5.43	1.85	5.25	1.59	5.46	1.60
Tool-use	5.94	1.39	5.67	1.75	5.40	1.60	5.68	1.51
No-tool-use	5.44	1.32	5.38	1.91	5.12	1.62	5.32	1.65
Collaborative (Total)	5.70	1.94	5.71	1.52	6.35	1.88	5.92	1.81
Tool-use	5.90	1.89	6.00	1.51	5.67	1.65	5.83	1.70
No-tool-use	5.54	2.00	5.63	1.54	6.91	1.90	5.97	1.88

was observed. On average, the math grades of students in the collaborative learning setting were somewhat higher compared to the individual students. Furthermore, in the individual learning setting it was observed that students who used a representational tool had higher math grades compared to individuals who did not use a tool. The math grades of individuals who used a tool were equal to those of students in the collaborative setting. If applicable, math grade was entered as a covariate in subsequent analyses. With regard to *pre-test scores*, no significant differences were found for setting (F (1,205)=3.12, p=.08), format (F (2,205)=0.06, p=.95), or tool-use (F (1,205)=0.13, p=.72). No interactions were observed either.

Learning task

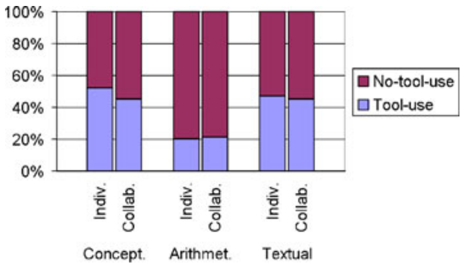
Use of representational tools

One of the research questions was about the students' inclination to use a representational tool and whether or not the format of the tool affected this inclination. The percentages of students in each condition who used a representational tool to construct a domain representation are displayed in Fig. 9.

When provided with a *conceptual tool*, 52% of the individual students and 45% of the pairs of students used it. A Chi-Square analysis showed that these percentages do not differ significantly, $X^2(1, N=55)=0.19$, *n.s.*

Of students provided with an *arithmetical tool*, 20% of the individuals and 21% of the pairs used it, which is no significant difference, $X^2(1, N=49)=0.01$, *n.s.*

Fig. 9 Percentage of students in each condition who did or did not construct a representation



When provided with a *textual tool*, 47% of the individuals and 45% of the pairs of students used it, which again is no significant difference, $X^2(1, N=52)=0.02$, *n.s.*

As can be observed in Fig. 9, the patterns of tool use are quite similar for the individual and the collaborative setting. The overall picture is that about 50% of the students provided with a conceptual or textual tool used the tool. Of students provided with the arithmetical tool, about 20% actually used the tool. A Chi-Square analysis showed that these differences between conditions are significant, $X^2(2, N=156)=10.58$, $p<.01$. Compared to students in the Arithmetical condition, students in the Conceptual condition used their tool more often ($X^2(1, N=104)=9.30$, $p<.01$) and so did students with a textual tool ($X^2(1, N=101)=7.49$, $p<.05$). No difference was observed between the Conceptual and the Textual condition ($X^2(1, N=107)=0.09$, *n.s.*).

The hypothesis that students would be inclined more to use a conceptual or a textual tool rather than an arithmetical tool, was therefore confirmed by the data. However, the data also show that the hypothesis that collaboration could have a stimulating effect on using the arithmetical tool was not confirmed.

Quality of constructed representations

In Table 2 the average quality scores of the constructed representations are displayed. In the case of representations constructed by pairs, the representations are considered a group product and therefore the quality scores are assigned to pairs and not to individuals. All representations were scored by two raters who worked independently. The inter-rater agreement was .89 (Cohen's Kappa) for the individual setting and .92 for the collaborative setting.

A two-way ANOVA with setting (individual vs. collaborative learning) and format as factors showed that with regard to quality scores there was no main effect of setting ($F(1,55)=3.69$, $p=.06$), no main effect of format ($F(2,55)=1.57$, $p=.22$), and no interaction effect ($F(2,55)=0.71$, $p=.50$).

Table 2 Quality scores of constructed representations

	Representational format											
	Conceptual (indiv. n=17; coll. 10 pairs)				Arithmetical (indiv. n=6; coll. 4 pairs)				Textual (indiv. n=15; coll. 9 pairs)			
	M	SD	Min	Max	M	SD	Min	Max	M	SD	Min	Max
Individual	2.38	0.99	1	5	2.67	1.97	1	6	2.67	0.98	1	4
Collaborative	2.70	1.42	1	5	4.00	1.41	2	5	3.00	0.87	2	4

Time-on-task

The log files provided data about the amount of time students spent on the learning task (see Table 3). Time-on-task is conceived here as the time that elapsed between the moment the participants started their learning environment and the moment they closed it. In the learning environment the participants worked through the five sections (see section Learning environment), read the cover stories, read and worked on the assignments and simulations, and used a representational tool that was integrated into their learning environment. The tool was at the participants' disposal throughout the time they spent in the learning environment.

The data presented in Table 3 were analyzed by means of a three-way ANOVA with setting (individual vs. collaborative learning), format, and tool-use as factors. Note that in the case of collaborative learning the process measures of the dyads were analyzed, not the measures of the individual students of the dyad. With regard to time-on-task it was found that there was a main effect of setting (F (1,143)=5.09, p<.05). The average time-on-task of students in the collaborative learning setting was lower than that of students in the individual learning stetting. If applicable, time-on-task was entered as a covariate in subsequent analyses. No main effects were observed for format (F(2,143)=0.10, p=.91), or tool-use (F(1,143)=2.97, p=.09). No interaction effects were observed.

Learning outcomes

Both in the collaborative and the individual setting students completed the post-test individually. The reliability, Cronbach's α, of the post-test was .80 in the individual setting and .78 in the collaborative setting. All post-test measures were analyzed and compared by means of ANOVAs with setting (individual or collaborative), format (Conceptual, Arithmetical, or Textual), and tool-use as factors.

Two research questions were related to learning outcomes. The first was: what are the effects of collaborative inquiry learning with representational tools on learning outcomes? The second, concerned the differential effects of different formats on learning outcomes. These questions and hypotheses will be addressed in the following paragraphs.

Table 3 Time-on-task (min.)

Setting	Representational format							
	Conceptual (indiv. n=33; coll. 22 pairs)		Arithmetical (indiv. n=30; coll. 19 pairs)		Textual (indiv. n=32; coll. 19 pairs)		Total (indiv. n=95; coll. 60 pairs)	
	M	SD	M	SD	M	SD	M	SD
Individual (Total)	69.64	13.95	66.95	17.61	66.64	18.32	67.78	16.57
Tool-use	70.84	14.17	70.62	15.98	70.50	16.85	70.67	15.13
No-tool-use	68.38	14.05	66.04	18.19	63.23	19.38	65.86	17.33
Collaborative	62.36	4.98	60.90	8.41	62.85	4.31	62.05	6.06
Tool-use	64.88	3.98	65.68	2.16	63.13	5.08	64.34	4.19
No-tool-use	60.25	4.88	59.63	9.04	62.59	3.75	60.63	6.64

Post-test overall scores

The post-test overall scores are displayed in Table 4. Post-test overall scores and math grade covaried and the same was true for post-test overall scores and time-on-task, so math grade and time-on-task were entered as covariates.

It was found that students in the collaborative learning setting obtained significantly higher post-test overall scores (F (1,201)=17.33, p<.001) compared to individual learners. The hypothesized beneficial effect of collaborative learning, observed in many other studies as well, was therefore also present in the current data. Furthermore, an interaction was observed between setting and tool-use (F (1,201)=6.23, p<.05) (see Fig. 10). The interaction indicates that students in the individual learning setting who used a representational tool obtained higher scores than individuals who did not, and the scores of those tool-using individuals equaled the post-test overall scores of students in the collaborative setting. The hypothesis that using a representational tool leads to better learning outcomes compared to not using such a tool, was therefore only partly confirmed. This effect was only observed in the individual learning setting. Inspection of the post-test scores of collaborating students not using a tool (see Table 4) suggests that dyads in the textual tool condition obtained lower scores (M=18.17) compared to dyads in the conceptual and arithmetical tool condition (means respectively 19.27 and 19.66). An additional analysis showed that this difference is not significant.

Conceptual and intuitive knowledge

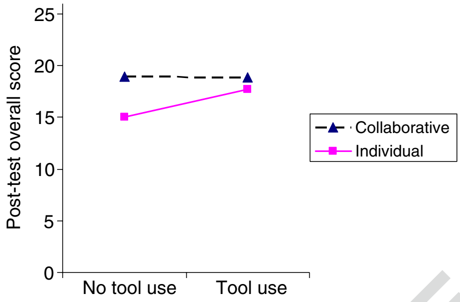
The average scores on *conceptual knowledge* items are displayed in Table 5.

No main effect of format of the representational tool was observed. The hypothesized beneficial effect of constructing a concept map on conceptual understanding was not confirmed by the data. The ANOVA showed an interaction between setting and tool-use (F (1,205)=6.37, p<.05) (see Fig. 11). This interaction indicates that students in the individual setting using a representational tool obtained slightly higher scores on conceptual knowledge compared to collaborating students (whether or not using a representational tool) and individuals not using a tool.

Table 4 Post-test overall scores (corrected for math grade and time-on-task; max. 26 points)

	Representational format							
	Conceptual (indiv. n=33) collab. n=44)		Arithmetical (indiv. n=30) collab. n=38)		Textual (indiv. n=32) collab. n=38)		Total (indiv. n=95) collab. n=120)	
	M	SE	M	SE	M	SE	M	SE
Individual (Total)	16.01	0.63	17.42	0.82	15.70	0.63	16.38	0.41
Tool-use	17.16	0.88	18.73	1.45	16.21	0.93	17.36	0.65
No-tool-use	14.86	0.90	16.11	0.74	15.20	0.87	15.39	0.50
Collaborative (Total)	19.43	0.54	18.40	0.71	18.18	0.58	18.67	0.36
Tool-use	19.60	0.79	17.13	1.26	18.19	0.84	18.31	0.57
No-tool-use	19.27	0.73	19.66	0.66	18.17	0.80	19.03	0.43

Fig. 10 Interaction between setting and tool use regarding the post-test overall score



Another aspect of conceptual knowledge was intuitive knowledge (see Table 6). Both math grade and time-on-task covaried with intuitive knowledge, therefore they were entered as covariates.

The ANCOVA showed that students in the collaborative learning setting obtained higher scores with respect to *intuitive* knowledge (F (1,201)=70.46, p<.001). Also, an interaction was observed between setting and format (F (2,201)=3.22, p<.05) (see Fig. 12). The interaction indicates that individual students with an arithmetical representational tool obtained higher intuitive knowledge scores than other individuals.

Procedural knowledge

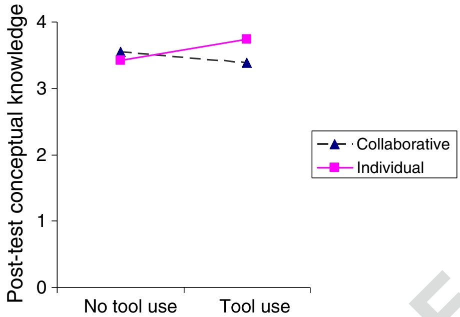
The mean scores on procedural knowledge items are displayed in Table 7. Here, both math grade and time-on-task covaried with procedural knowledge, so they were entered as covariates.

The ANCOVA indicated no significant differences between setting, format, and tool-use. Only one, rather complicated interaction was observed between setting, format, and tool-use (F (2,201)=4.22, p<.05). This interaction is possibly caused by a relatively high score of individuals in the arithmetical condition who used the tool and the relatively low scores of collaborating students in the same condition who used the tool. No main effect of format was found, so the hypothesized beneficial effect of constructing an arithmetical

Table 5 Conceptual knowledge score (max. 4 points)

	Representational format						Total (indiv. n=95) collab. n=122)	
	Conceptual (indiv. n=33) collab. n=44)		Arithmetical (indiv. n=30) collab. n=38)		Textual (indiv. n=32) collab. n=40)			
	M	SD	M	SD	M	SD	M	SD
Individual (Total)	3.52	0.62	3.53	0.57	3.59	0.56	3.55	0.58
Tool-use	3.65	0.61	3.50	0.55	3.93	0.26	3.74	0.50
No-tool-use	3.37	0.62	3.54	0.59	3.29	0.59	3.42	0.60
Collaborative (Total)	3.64	0.49	3.55	0.65	3.28	0.85	3.49	0.68
Tool-use	3.60	0.50	3.25	0.89	3.22	0.88	3.39	0.75
No-tool-use	3.67	0.48	3.63	0.56	3.32	0.84	3.55	0.64

Fig. 11 Interaction between setting and tool use regarding conceptual knowledge



representation on the acquisition of procedural knowledge was not confirmed by the data. In general, equal levels of procedural knowledge can be obtained with other formats.

Situational knowledge

The post-test scores on situational knowledge are displayed in Table 8. Neither math grade nor time-on-task covaried here, so they were left out of the ANOVA.

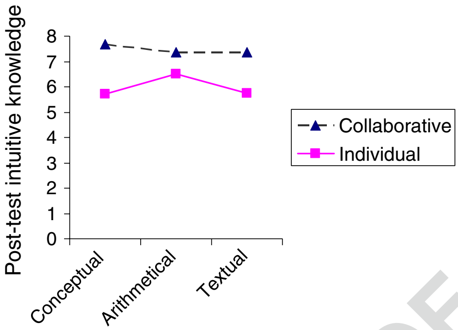
The analysis indicated a significant difference between settings (F (1,205)=8.00, p<.01), formats (F (2,205)=4.22, p<.05), and tool-use (F(1,205)=5.02, p<.05). Students in the collaborative setting obtained higher situational knowledge scores compared to individuals; students in the arithmetical condition outperformed students in the textual condition; and students who used a tool to externalize, obtained significantly higher scores than students who did not use a tool. Furthermore, an interaction between setting and tool-use was observed (F (1,205)=10.12, p<.01) (see Fig. 13). The interaction shows that students in the individual learning setting who used a representational tool obtained higher situational knowledge scores than individuals who did not and their scores were equal to those of tool-users in the collaborative setting.

The main effect observed for format of the representational tool was significant, yet it disconfirmed the hypothesis that creating a textual representation would enhance situational knowledge more than constructing an arithmetical or conceptual representation.

Table 6 Intuitive conceptual knowledge score (corrected for math grade and time-on-task; max. 8 points)

	Representational format							
	Conceptual (indiv. n=33) collab. n=44)		Arithmetical (indiv. n=30) collab. n=38)		Textual (indiv. n=32) collab. n=38)		Total (indiv. n=95) collab. n=120)	
	M	SE	M	SE	M	SE	M	SE
Individual (Total)	5.62	0.21	6.46	0.28	5.70	0.22	5.92	0.14
Tool-use	5.93	0.30	6.72	0.49	5.74	0.32	6.13	0.22
No-tool-use	5.31	0.31	6.19	0.25	5.66	0.30	5.72	0.17
Collaborative (Total)	7.72	0.18	7.39	0.24	7.38	0.20	7.50	0.12
Tool-use	7.81	0.27	7.45	0.43	7.38	0.29	7.54	0.19
No-tool-use	7.63	0.25	7.33	0.23	7.39	0.27	7.45	0.15

Fig. 12 Interaction between setting and format regarding intuitive conceptual knowledge



Discussion and conclusion

In (collaborative) inquiry learning, students plan and execute inquiry processes and select, process, analyze, interpret, organize, and integrate information into meaningful and coherent knowledge structures. Expressing and externalizing one's ideas and understandings, for example in the form of constructing a domain representation, have been found to foster these processes. One of the questions addressed in the current study was: does creating a domain representation affect learning outcomes in collaborative inquiry learning? Second, the nature of the domain representations can be quite different, depending on the representational format used (e.g., circles, arrows, and keywords in concept maps; words in written summaries; numbers, formulas, and equations in arithmetic). The next research question was: does the format used to create a domain representation differentially affect students' domain understanding by emphasizing or de-emphasizing aspects of the learning materials? And third, does the representational format have differential effects on students' inclination to construct a representation? These questions were explored in the domain of combinatorics and probability theory. Three different representational tools were developed, each designed to constrain the format students could use to construct a domain representation.

Table 7 Procedural knowledge (corrected for math grade and time-on-task; max. 10 points)

	Representational format							
	Conceptual (indiv. n=33) collab. n=44)		Arithmetical (indiv. n=30) collab. n=38)		Textual (indiv. n=32) collab. n=38)		Total (indiv. n=95) collab. n=120)	
	M	SE	M	SE	M	SE	M	SE
Individual (Total)	3.97	0.38	4.28	0.50	3.71	0.38	3.99	0.24
Tool-use	4.22	0.53	4.86	0.87	3.50	0.55	4.20	0.39
No-tool-use	3.72	0.54	3.69	0.44	3.92	0.52	3.77	0.30
Collaborative (Total)	4.57	0.32	3.87	0.43	4.50	0.35	4.32	0.21
Tool-use	4.65	0.48	2.70	0.75	4.93	0.50	4.09	0.34
No-tool-use	4.49	0.44	5.04	0.40	4.08	0.48	4.54	0.26

Table 8 Situational knowledge score (max. 4 points)

	Representational format							
	Conceptual (indiv. n=33) collab. n=44)		Arithmetical (indiv. n=30) collab. n=38)		Textual (indiv. n=32) collab. n=40)		Total (indiv. n=95) collab. n=122)	
	M	SD	M	SD	M	SD	M	SD
Individual (Total)	2.93	0.20	3.17	0.26	3.01	0.18	2.93	0.13
Tool-use	3.41	0.28	3.67	0.46	3.07	0.29	3.38	0.21
No-tool-use	2.46	0.29	2.68	0.24	2.30	0.28	2.47	0.16
Collaborative (Total)	3.50	0.17	3.71	0.23	3.01	0.18	3.41	0.11
Tool-use	3.55	0.25	3.75	0.40	2.73	0.26	3.34	0.18
No-tool-use	3.46	0.23	3.66	0.21	3.28	0.25	3.47	0.13

The first research question focused on the effects of collaborative inquiry learning with representational tools on learning outcomes. In order to test whether collaborative aspects influence inquiry learning with representational tools, the learning outcomes of students in a collaborative learning setting were compared to learning outcomes of students in an individual learning setting. Following existing literature on the comparison learning outcomes in individual and collaborative learning settings (e.g., Lou et al. 2001; Slavin 1995; van der Linden et al. 2000), it was hypothesized that learning outcomes in the collaborative learning setting would be higher than those in the individual learning setting. Our data were in line with findings reported in other studies: in the collaborative inquiry learning setting the overall learning results were significantly higher than in the individual setting, regardless of whether or not the dyads had used a representational tool to externalize their knowledge. In the individual inquiry learning setting, tool-use did make a difference. The post-test overall performance of individuals who externalized their knowledge was close to the performance of collaborating students, whereas the overall performance of individuals who had not engaged in externalization was significantly lower.

Collaborative learners outperformed individuals in particular on intuitive knowledge and situational knowledge. The observation that collaborative learners (regardless of whether or not they constructed a representation) outperformed individuals (even those who did construct a representation), implies that, in this study, intuitive knowledge is enhanced by collaborative learning and not by constructing representations per se. Intuitive knowledge is particularly fostered by interpretation and sense-making processes (Gijlers and de Jong

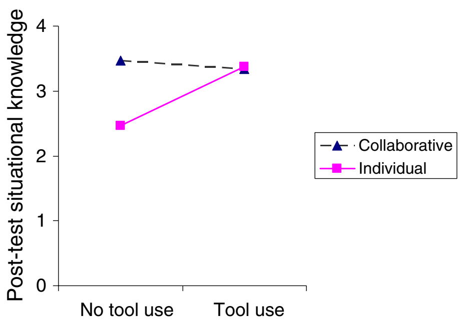


Fig. 13 Interaction between setting and tool-use regarding situational knowledge

submitted; Reid et al. 2003; Zhang et al. 2004), which suggests that collaboration stimulates these processes in a way that goes beyond the effects of externalizing knowledge by means of a representational tool alone. Q4

Situational knowledge, which is a prerequisite for going beyond the superficial details of problems in order to recognize the concepts and structures that underlie the problem (e.g., Fuchs et al. 2004), was also fostered by collaboration, although not exclusively: here collaboration, the format of representational tools, and tool-use all contributed to the acquisition of situational knowledge. Apparently all forms of externalization help to gain understanding of problem structures in this domain.

The second research question focused on the influence of representational format used to construct a representation on knowledge construction and domain understanding. Creating a *conceptual* representation like a concept map was hypothesized to enhance knowledge about the conceptual aspects of the domain, rather than procedural or situational aspects. Constructing representations in an *arithmetical* format was assumed to foster the acquisition of procedural knowledge and using a *textual* format was assumed to improve students' attention to situational knowledge. The results show that there is no evidence for this hypothesized mapping between representational format and the enhancement of a specific kind of understanding. For example, constructing a concept map does not enhance conceptual understanding. The mapping that was observed however, was in an unexpected direction: students who constructed an arithmetical representation showed enhanced levels of situational knowledge on the post-test compared to students who created a textual representation. Furthermore, an interaction effect indicated that individuals creating an arithmetical representation also showed enhanced levels of intuitive conceptual knowledge compared to other individuals.

Although the arithmetical format was the only representational format that could be directly linked to the enhancement of a specific type of knowledge (situational knowledge) and in the case of learning in an individual setting also to higher levels of intuitive conceptual knowledge, this representational format turned out to have some disadvantages as well. These came to light when answering the third research question: does the representational format have differential effects on students' inclination to construct a representation? In the case of concept maps it was assumed that participants would be inclined to use it. This representational format is relatively easy to understand and use, especially if there are not too many concepts and relations (van Drie et al. 2005). Regarding arithmetical formats it was hypothesized that students would have difficulty constructing them (cf. Tarr and Lannin 2005), however, discussing the arithmetical aspects of the domain with a peer in a collaborative learning setting was assumed to have a beneficial effect on students' inclination to use this representational format. The third format for constructing a domain representation considered here was a *textual* format. The current domain could easily be described in terms of everyday life contexts and situations. It was expected that students would not experience much difficulty with using the textual format. Overall, this is one of the most commonly used formats inside and outside educational settings. Therefore, it was assumed that many participants would be inclined to use this representational tool.

In both the collaborative setting and the individual setting the formats of the tools did not lead to differential effects on the quality of the constructed representations, these were similar across settings and formats. The results did show differences with regard to students' inclination to use a representational tool. Clarebout and Elen (2006, 2009a, b; see also: Jiang et al. 2009) observed that tools, which are integrated into learning environments are often used inadequately or not at all by students. The current study added to this insight that the format of representational tools affects the students' inclination to use a tool and engage in constructing a domain representation. About 20% of the students provided with the arithmetical representational tool used it. Representational tools with a conceptual or textual format were found to be used substantially more by students to engage in constructing a representation (around 50% use). This behavior turned out to be consistent in both settings. The usage percentages were remarkably similar in both the individual and the collaborative learning setting. Possibly, the arithmetical format is more difficult to use to construct a domain representation. Another possibility is that students failed to view mathematical symbols as reflections of principles and structures, but rather perceived them as indicators of which operations need to be performed (Atkinson et al. 2003; Cheng 1999; Greenes 1995; Nathan et al. 1992; Niemi 1996; Ohlsson and Rees 1991). This would mean Q5 that the textual and the conceptual format are more close to the code in which students can explain the domain to themselves, or maybe students consider those formats more suited to express their knowledge to the outside world. A complementary explanation could be that the use of arithmetical formats requires more advanced levels of domain understanding. To domain experts (e.g., teachers, university students of mathematics) the arithmetical representational format might be a convenient and efficient way of expressing and externalizing knowledge. Perhaps in the case of novices, still at the stage of trying to get some grip on the subject matter, it might not be an easy and straightforward representational format to express oneself and to externalize one's knowledge.

Some of the limitations of the current study will be discussed below along with some suggestions for future research. The quantitative approach used in the study showed how representational format affects students' inclination to use a representational tool. A qualitative research methodology (e.g., case-studies, interviews with participants) in a next study can possibly help to understand the motives of students to use or not use a certain representational format. A second point is the constraining of the format in the current study. In a next study, it could be useful to investigate the effects of allowing students to express and externalize their knowledge without being constrained to using a specific representational format. Another suggestion is to explore whether specific representational tools can be used in a complementary fashion, for example to support different stages or tasks during the learning process. For example, using concept maps in the early stages to help students identify key concepts, using textual representations to situate the identified concepts in contexts, and using an arithmetical format in the final stages of the learning task to stimulate students to express their knowledge in a more abstract way.

Another issue is the communication between students. The analyses did not include the actual communication between students. Maybe this would have shed some light on additional effects of representational tools and their formats on collaboration. In studies by Suthers and Hundhausen (2003) and van Drie et al. (2005) for example, it was found that the format of representational tools influenced communication and the activities performed by collaborating students.

Another question regarding collaborative inquiry learning concerns the medium through which students communicate with each other. In the current study, students worked in a face-to-face setting, sitting next to each other. Face-to-face communication is considered to be rich in the sense that it provides both verbal and non-verbal information (e.g., gesturing, nodding, pointing, facial expressions, and intonation of speech), but it also allows students to communicate faster and much more elaborate, which can be crucial in the case of interpretation and sense-making. There is no guarantee that the results of the current study would have been found in a setting in which students communicated through chat. Chat communication in collaborative settings is known to put some constraints on communication. For example, in chatting, students tend to be much more succinct, to focus more on technical and organizational issues instead of domain aspects, and to easily jump from topic to topic. This can have positive effects (e.g., brainstorming), but can also be detrimental when the situation requires students to focus on one topic (Strømsø et al. 2007; Kerr and Murthy 2004; Anjewierden et al. 2007). In this case, a shared representational tool may not only stimulate interpretation and conclusion activities, but also serves as an additional channel for communication and reasoning. This is in line with Van Drie et al. (2005) who remarked that (when students communicate via chat) a “representational tool does not only function as a cognitive tool that can elicit elaborative activities, but also as a tool *through which* students communicate” (p. 598). It would be interesting to explore the relation between mode of communication, externalization, and the effects on knowledge acquisition in a future study.

Acknowledgements The authors gratefully acknowledge the Netherlands Organisation for Scientific Research (NWO) for funding these studies (project number 411-02-162).

Appendix

Table 9 Q6

	Represented?	Conceptual tool	Arithmetical tool	Textual tool	PNT
A	The concept of “Replacement”	-Literally, or descriptive	Two formulas or calculations in which “replacement” varies	-Literally, or descriptive	1
		Examples:	Examples:	Examples:	
		-“Replacement”	-“(1/n) × (1/n) × (1/n) = P (1/n) × (1/(n-1)) × (1/(n-2)) = P”	-“Replacement”	
		-“Category 1: without replacement; order important”	-“1/5 × 1/4 × 1/3 1/5 × 1/5 × 1/5”	-“Category 1: without replacement; order important”	
		-...[Runners, BK]... then you have to do 1/7 × 1/6 × 1/5 because each time there is one runner fewer”	-“p=1/10 × 1/10 × 1/10 p=1/5 × 1/4 × 1/3”	-“...If there are 7 runners, then the chance is 1 out of 7 (1/7), if that runner passes the finish, then there are 6 runners left, then there is a chance of 1 out of 6 (1/6), and so on.	
B	The concept of “Order”	-Literally, or descriptive	Two formulas or calculations in which “order” varies	-Literally, or descriptive	1
		Examples:	Examples:	Examples:	
		-“Order”	-“(1/n) × (1/n) × (1/n) (k/n) × ((k-1)/n) × ((k-2)/n)”	-“Order”	
		-“Category 1: without replacement; order important”	-“1/5 × 1/4 × 1/3 3/5 × 2/4 × 1/3”	-“Category 1: without replacement; order important”	

t9.13 **Table 9** (continued)

		Represented?	Conceptual tool	Arithmetical tool	Textual tool	PNT
t9.12			-"...If there are 7 runners and you predict the top 3 without specifying the positions of specific runners in the top 3..."		-"...At a game of Bingo, order is not important"	
t9.13	C	Calculation	-Formal, literally, descriptive, or a concrete calculation	Formal (formula) or a concrete calculation	-Formal, literally, descriptive, or a concrete calculation	1
t9.14			Examples:	Examples:	Examples:	
t9.15			-p = acceptable outcomes/possible outcomes	-"(1/n) × (1/n) × (1/n)"	-p = acceptable outcomes/possible outcomes	
t9.16			- 1/5 × 1/4 × 1/3	-"1/5 × 1/4 × 1/3"	- 1/5 × 1/4 × 1/3	
t9.17			-... when you also bet on the order in which the marbles will be selected, your chance is: 1/5 and 1/4 is 1/20..."		-... when you also bet ont he order in which the marbles will be selected, your chance is: 1/5 and 1/4 is 1/20..."	
t9.18	D	Probability	-Literal reference to the term "probability"/p, or a description of the concept	-Literal reference to the term "p"	-Literal reference to the term "probability"/p, or a description of the concept	1
t9.19			-Expression of a concrete probability (e.g. a fraction), but then it need to be made clear in the context (e.g. by a calculation) where the probability comes from	-Expression of the outcome of a calculation	-Expression of a concrete probability (e.g. a fraction), but then it need to be made clear in the context (e.g. by a calculation) where the probability comes from	
t9.20			Examples:	Examples:	Examples:	
t9.21			-"In order to calculate 'p' the chances need to be multiplied."	-"p = (1/n) × (1/n) × (1/n)"	-"In order to calculate 'p' the chances need to be multiplied."	
t9.22			-p = 1/5 × 1/4 × 1/3	-"p = 1/5 × 1/4 × 1/3"	-p = 1/5 × 1/4 × 1/3	
t9.23			-"...In that case [student refers to a situation outlined earlier], the probability is 1/10"	-"1/5 × 1/4 × 1/3 = 1/60"	-"...In that case [student refers to a situation outlined earlier], the probability is 1/10"	
t9.24	E	Effect of n on probability	-Descriptive or on basis of calculations showing the effect (in the latter case, k needs to be constant)	A formula or a series of calculations showing the effect (in the latter case, k needs to be constant)	-Descriptive or on basis of calculations showing the effect (in the latter case, k needs to be constant)	1
t9.25			Examples:	Examples:	Examples:	
t9.26			-"fewer options = higher chance"	-"(1/n) × (1/n) × (1/n) = 1/n^3"	-"If the number of elements you can choose from increases, the chance will be smaller that you will select a specific element"	

t9.28 **Table 9** (continued)

		Represented?	Conceptual tool	Arithmetical tool	Textual tool	PNT
t9.27 t9.28			-"If fewer runners attend the race, the chance your prediction is correct will increase"	-"1/5 × 1/4 × 1/3=1/60 1/6 × 1/5 × 1/4=1/120"	-"If fewer runners attend the race, the chance your prediction is correct will increase"	
t9.29	F	Effect of *k* on probability	-Descriptive or on basis of calculations showing the effect (in the latter case, n needs to be constant)	A formula or a series of calculations showing the effect (in the latter case, k needs to be constant)	-Descriptive or on basis of calculations showing the effect (in the latter case, *n* needs to be constant)	1
t9.30			Examples:	Examples:	Examples:	
t9.31			-"with 1 choice → 1/possible outcomes; with more choices → number of choices/ possible outcomes"	-"$(1/n) \times (1/n)=1/n^2$ $(1/n) \times (1/n) \times (1/n)=1/n^3$"	-"When your prediction is less elaborate, the probability that your prediction will be correct increases"	
t9.32			-"If you only predict who will win the race and not the top 3, then the chance is greater that your prediction will be correct"	-"1/5 × 1/4=1/20 1/5 × 1/4 × 1/3=1/60"	-"If you only predict who will win the race and not the top 3, then the chance is greater that your prediction will be correct"	
t9.33	G	Effect of replacement on probability	-Descriptive or on basis of calculations showing the effect (in the latter case, *n* and *k* need to be constant)	A series of formulas or calculations showing the effect, but the outcome (*p*) needs to be represented as well and *n* and *k* need to be constant	-Descriptive or on basis of calculations showing the effect (in the latter case, *n* and *k* need to be constant)	1
t9.34			Examples:	Examples:	Examples:	
t9.35			-"If it is a matter of replacement, your chances will decrease"	-"$(1/n) \times (1/n)=1/n^2$ $(1/n) \times (1/(n\text{-}1))= 1/(n^2-n)$"	-"If it is a matter of replacement, your chances will decrease"	
t9.36			-"...if you have 10 different cell phones and you need to select one, your chance will be 1 out of 10, if you put the phone back your chance will be 1 out of 10 again, but if you leave it out your chance will increase that you will select the next phone as predicted"	-"1/5 × 1/4 × 1/3=1/60 1/5 × 1/5 × 1/5=1/125"	-"...if you have 10 different cell phones and you need to select one, your chance will be 1 out of 10, if you put the phone back your chance will be 1 out of 10 again, but if you leave it out your chance will increase that you will select the next phone as predicted"	
t9.37	H	Effect of order on probability	-Descriptive or on basis of calculations showing the effect (in the latter case, *n* and *k* need to be constant)	A series of formulas or calculations showing the effect, but the outcome (*p*) needs to be represented as well and *n* and *k* need to be constant	-Descriptive or on basis of calculations showing the effect (in the latter case, *n* and *k* need to be constant)	1

Table 9 (continued)

Represented?	Conceptual tool	Arithmetical tool	Textual tool	PNT
	Examples:	Examples:	Examples:	
	-"If order is important, the chance your prediction will be right will decrease"	-"$(1/n) \times (1/n) = 1/n^2$ $(k/n) \times ((k-1)/n) = (k^2-k)/n^2$"	-"If order is important, the chance your prediction will be right will decrease"	
	-"...If there are 7 runners and you predict the top 3, then the probability is $1/7 \times 1/6 \times 1/5 = 1/210$, but without specifying the positions of specific runners in the top 3 the probability is $3/7 \times 2/6 \times 1/5 = 6/210$..."	-"$1/5 \times 1/4 \times 1/3 = 1/60$ $3/5 \times 2/4 \times 1/3 = 6/60$"	-"...If there are 7 runners and you predict the top 3, then the probability is $1/7 \times 1/6 \times 1/5 = 1/210$, but without specifying the positions of specific runners in the top 3 the probability is $3/7 \times 2/6 \times 1/5 = 6/210$..."	
Maximum number of points				8

References

Anjewierden, A., Kolloffel, B., & Hulshof, C. (2007). Towards educational data mining: Using data mining methods for automated chat analysis to understand and support inquiry learning processes. In C. Romero, M. Pechenizikiy, T. Calders, & S. R. Viola (Eds.), *Applying data mining in e-learning* (pp. 23–32). *Proceedings of the International Workshop on Appling Data Mining in e-Learning (ADML'07).* Crete, Greece, September 2007.

Bell, T., Urhahne, D., Schanze, S., & Ploetzner, R. (2010). Collaborative inquiry learning: Models, tools, and challenges. *International Journal of Science Education, 32*(3), 349–377.

Berthold, K., & Renkl, A. (2009). Instructional aids to support a conceptual understanding of multiple representations. *Journal of Educational Psychology, 101*(1), 70–87.

Cavalli-Sforza, V., Weiner, A. W., & Lesgold, A. M. (1994). Software support for students engaging in scientific activity and scientific controversy. *Science & Education, 78*(6), 577–599.

Clarebout, G., & Elen, J. (2006). Tool use in computer-based learning environments: Towards a research framework. *Computers in Human Behavior, 22*, 389–411.

Clarebout, G., & Elen, J. (2009a). Benefits of inserting support devices in electronic learning environments. *Computers in Human Behavior, 25*(4), 804–810. doi:10.1016/j.chb.2008.07.006.

Clarebout, G., & Elen, J. (2009b). The complexity of tool use in computer-based learning environments. *Instructional Science, 37*, 475–486. doi:10.1007/s11251-008-9068-3.

Clarebout, G., Horz, H., & Elen, J. (2009). The use of support devices in electronic learning environments. *Computers in Human Behavior, 25*(4), 793–794. doi:10.1016/j.chb.2008.07.004.

Cox, R. (1999). Representation construction, externalised cognition and individual differences. *Learning and Instruction, 9*, 343–363.

de Jong, T. (2005). The guided discovery principle in multimedia learning. In R. E. Mayer (Ed.), *The cambridge handbook of multimedia learning* (pp. 215–229). Cambridge: Cambridge University Press.

de Jong, T. (2006). Technological advances in inquiry learning. *Science, 312*, 532–533.

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

de Jong, T., & van Joolingen, W. R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research, 68*, 179–202.

de Jong, T., van Joolingen, W. R., Swaak, J., Veermans, K., Limbach, R., King, S., et al. (1998). Self-directed learning in simulation-based discovery environments. *Journal of Computer Assisted Learning, 14*, 235–246.

Ertl, B., Kopp, B., & Mandl, H. (2008). Supporting learning using external representations. *Computers & Education, 51*, 1599–1608.

Eysink, T. H. S., de Jong, T., Berthold, K., Kolloffel, B., Opfermann, M., & Wouters, P. (2009). Learner performance in multimedia learning arrangements: An analysis across instructional approaches. *American Educational Research Journal, 46*, 1107–1149.

Foos, P. W. (1995). The effect of variations in text summary opportunities on test-performance. *Journal of Experimental Education, 63*(2), 89–95.

Fuchs, L. S., Fuchs, D., Prentice, K., Hamlett, C. L., Finelli, R., & Courey, S. J. (2004). Enhancing mathematical problem solving among third-grade students with schema-based instruction. *Journal of Educational Psychology, 96*(4), 635–647.

Garfield, J., & Ahlgren, A. (1988). Difficulties in learning basic concepts in probability and statistics: Implications for research. *Journal for Research in Mathematics Education, 19*(1), 44–63.

Gerjets, P., Scheiter, K., Opfermann, M., Hesse, F. W., & Eysink, T. H. S. (2009). Learning with hypermedia: The influence of representational formats and different levels of learner control on performance and learning behavior. *Computers in Human Behavior, 25*(2), 360–370.

Gijlers, H., & de Jong, T. (submitted). Facilitating collaborative inquiry learning with shared concept maps and proposition tables.

Gijlers, H., & de Jong, T. (2009). Sharing and confronting propositions in collaborative inquiry learning. *Cognition and Instruction, 27*, 239–268.

Hidi, S., & Anderson, V. (1986). Producing written summaries: Task demands, cognitive operations, and implications for instruction. *Review of Educational Research, 56*, 473–493.

Hiebert, J., & Lefevre, P. (1986). Conceptual and procedural knowledge in mathematics: An introductory analysis. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 1–27). Hillsdale: Lawrence Erlbaum Associates, Publishers.

Janssen, J., Erkens, G., & Kanselaar, G. (2007). Visualization of agreement and discussion processes during computer-supported collaborative learning. *Computers in Human Behavior, 23*(3), 1105–1125.

Jiang, L., Elen, J., & Clarebout, G. (2009). The relationships between learner variables, tool-usage behaviour and performance. *Computers in Human Behavior, 25*(2), 501–509. doi:10.1016/j.chb.2008.11.006.

Kaput, J. J. (1995). Creating cybernetic and psychological ramps from the concrete to the abstract: Examples from multiplicative structures. In D. N. Perkins, J. L. Schwartz, M. M. West, & M. Stone Wiske (Eds.), *Software goes to school: Teaching for understanding with new technologies* (pp. 130–154). New York: Oxford University Press.

Kerr, D. S., & Murthy, U. S. (2004). Divergent and convergent idea generation in teams: A comparison of computer-mediated and face-to-face communication. *Group Decision and Negotiation, 13*(4), 381–399.

Kolloffel, B., Eysink, T. H. S., de Jong, T., & Wilhelm, P. (2009). The effects of representational format on learning combinatorics from an interactive computer-simulation. *Instructional Science, 37*, 503–517.

Kolloffel, B., Eysink, T. H. S., & de Jong, T. (2010). The influence of learner-generated domain representations on learning combinatorics and probability theory. *Computers in Human Behavior, 26*(1), 1–11.

Löhner, S., van Joolingen, W. R., Savelsbergh, E. R., & van Hout-Wolters, B. (2005). Students' reasoning during modeling in an inquiry learning environment. *Computers in Human Behavior, 21*(3), 441–461.

Lou, Y. P., Abrami, P. C., & d'Apollonia, S. (2001). Small group and individual learning with technology: A meta-analysis. *Review of Educational Research, 71*(3), 449–521.

Lund, K., Molinari, G., Séjourné, A., & Baker, M. (2007). How do argumentation diagrams compare when student pairs use them as a means for debate or as a tool for representing debate? *International Journal of Computer-Supported Collaborative Learning, 2*(2), 273–295.

Manlove, S., Lazonder, A. W., & de Jong, T. (2006). Regulative support for collaborative scientific inquiry learning. *Journal of Computer Assisted Learning, 22*(2), 87–98.

Mayer, R. E. (2002). Rote versus meaningful learning. *Theory into Practice, 41*, 226–232.

Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? The case for guided methods of instruction. *The American Psychologist, 59*(1), 14–19.

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

Novak, J. D. (1990). Concept maps and vee diagrams: Two metacognitive tools to facilitate meaningful learning. *Instructional Science, 19*(1), 29–52.

Novak, J. D. (2002). Meaningful learning: The essential factor for conceptual change in limited or inappropriate propositional hierarchies leading to empowerment of learners. *Science & Education, 86*(4), 548–571.

Park, S. I., Lee, G., & Kim, M. (2009). Do students benefit equally from interactive computer simulations regardless of prior knowledge levels? *Computers & Education, 52*(3), 649–655.

Quintana, C., Reiser, B. J., Davis, E. A., Krajcik, J., Fretz, E., Duncan, R. G., et al. (2004). A scaffolding design framework for software to support science inquiry. *Journal of the Learning Sciences, 13*(3), 337–386.

Reiser, B. J. (2004). Scaffolding complex learning: The mechanisms of structuring and problematizing student work. *Journal of the Learning Sciences, 13*(3), 273–304.

Rieber, L. P., Tzeng, S. C., & Tribble, K. (2004). Discovery learning, representation, and explanation within a computer-based simulation: Finding the right mix. *Learning and Instruction, 14*, 307–323.

Rittle-Johnson, B., & Koedinger, K. R. (2005). Designing knowledge scaffolds to support mathematical problem solving. *Cognition and Instruction, 23*(3), 313–349.

Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology, 93*, 346–362.

Saab, N., van Joolingen, W. R., & van Hout-Wolters, B. H. A. M. (2007). Supporting communication in a collaborative discovery learning environment: The effect of instruction. *Instructional Science, 35*(1), 73–98.

Sharma, P., & Hannafin, M. J. (2007). Scaffolding in technology-enhanced learning environments. *Interactive Learning Environments, 15*(1), 27–46.

Sins, P. H. M., Savelsbergh, E. R., van Joolingen, W. R., & van Hout-Wolters, B. H. A. M. (2009). The relation between students' epistemological understanding of computer models and their cognitive processing on a modelling task. *International Journal of Science Education, 31*(9), 1205–1229.

Slavin, R. E. (1995). *Cooperative learning: Theory, research, and practice* (2nd ed.). Boston: Allyn and Bacon.

Strømsø, H. I., Grøttum, P., & Lycke, K. H. (2007). Content and processes in problem-based learning: A comparison of computer-mediated and face-to-face communication. *Journal of Computer Assisted Learning, 23*, 271–282.

Suthers, D. D. (2006). Technology affordances for intersubjective meaning making: A research agenda for cscl. *International Journal of Computer-Supported Collaborative Learning, 1*(3), 315–337.

Suthers, D. D. (2008). Empirical studies of the value of conceptually explicit notations in collaborative learning. In S. Okada, S. Buckingham Shum, & T. Sherborne (Eds.), *Knowledge cartography* (pp. 1–23). Cambridge: MIT Press.

Suthers, D. D., & Hundhausen, C. D. (2003). An experimental study of the effects of representational guidance on collaborative learning processes. *Journal of the Learning Sciences, 12*, 183–218.

Suthers, D. D., Hundhausen, C. D., & Girardeau, L. E. (2003). Comparing the roles of representations in face-to-face and online computer supported collaborative learning. *Computers & Education, 41*(4), 335–351.

Suthers, D. D., Vatrapu, R., Medina, R., Joseph, S., & Dwyer, N. (2008). Beyond threaded discussion: Representational guidance in asynchronous collaborative learning environments. *Computers & Education, 50*(4), 1103–1127. doi:10.1016/j.compedu.2006.10.007.

Swaak, J., & de Jong, T. (1996). Measuring intuitive knowledge in science: The development of the what-if test. *Studies in Educational Evaluation, 22*, 341–362.

Swaak, J., & de Jong, T. (2001). Discovery simulations and the assessment of intuitive knowledge. *Journal of Computer Assisted Learning, 17*(3), 284–294.

Sweller, J. (1989). Cognitive technology: Some procedures for facilitating learning and problem solving in mathematics and science. *Journal of Educational Psychology, 81*, 457–466.

Tarr, J. E., & Lannin, J. K. (2005). How can teachers build notions of conditional probability and independence? In G. A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 215–238). New York: Springer.

Toth, E. E., Suthers, D. D., & Lesgold, A. M. (2002). "Mapping to know": The effects of representational guidance and reflective assessment on scientific inquiry. *Science & Education, 86*(2), 264–286.

Trundle, K. C., & Bell, R. L. (2010). The use of a computer simulation to promote conceptual change: A quasi-experimental study. *Computers & Education, 54*(4), 1078–1088.

van Boxtel, C., van der Linden, J., & Kanselaar, G. (2000). Collaborative learning tasks and the elaboration of conceptual knowledge. *Learning and Instruction, 11*, 311–330.

van der Linden, J., Erkens, G., Schmidt, H., & Renshaw, P. (2000). Collaborative learning. In R. J. Simons, J. Van der Linden, & T. Duffy (Eds.), *New learning* (pp. 37–54). Dordrecht: Kluwer Academic Publishers.

van der Meijden, H., & Veenman, S. (2005). Face-to-face versus computer-mediated communication in a primary school setting. *Computers in Human Behavior, 21*(5), 831–859. doi:10.1016/j.chb.2003.10.005.

van Drie, J., van Boxtel, C., Jaspers, J., & Kanselaar, G. (2005). Effects of representational guidance on domain specific reasoning in cscl. *Computers in Human Behavior, 21*(4), 575–602.

Van Meter, P., & Garner, J. (2005). The promise and practice of learner-generated drawing: Literature review and synthesis. *Educational Psychology Review, 17*, 285–325.

Van Meter, P., Aleksic, M., Schwartz, A., & Garner, J. (2006). Learner-generated drawing as a strategy for learning from content area text. *Contemporary Educational Psychology, 31*, 142–166.

Wilensky, U. (1995). Paradox, programming, and learning probability: A case study in a connected mathematics framework. *Journal of Mathematical Behavior, 14*, 253–280.

Wouters, P., Paas, F., & Van Merriënboer, J. J. G. How to optimize learning from animated expert models? In *Annual meeting of the AERA, Chicago, IL, USA, April 2007.*