Multimodal representations in collaborative history learning

Netherlands Organisation for Scientific Research

The research reported in this dissertation was supported by the Netherlands Organisation for Scientific Research (project no. 411-01-004).

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

ISBN: 978-90-393-4567-2

Printed by PrintParntners Ipskamp.

Cover illustration: Wim Euverman, Utrecht.
Illustrations: p. 73, Figure 5.2; p. 84, Figures A.1. & A.2.; p. 85, Figure A.3.; p. 96, Figure C.3. & C.4.: Wim Euverman, Utrecht.

Multimodal representations in collaborative history learning

Multimodale representaties bij samenwerkend geschiedenis leren

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor
aan de Universiteit Utrecht
op gezag van de rector magnificus, prof.dr. W.H. Gispen,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen
op woensdag 20 juni 2007 des middags te 2.30 uur

door

Maaike Elizabeth Prangsma
geboren op 22 juni 1974 te Apeldoorn

Promotoren: Prof. dr. G. Kanselaar
Prof. dr. P.A. Kirschner
Co-promotor: Dr. C.A.M. van Boxtel

Dit proefschrift werd (mede) mogelijk gemaakt met financiële steun van NWO.

Contents

1. Introduction

We cannot – in the 21st century – imagine a school textbook not containing many different types of illustrations. Modern schoolbooks are packed with pictures, tables, graphs, and diagrams – in addition to texts. Visual or multimodal representations are seen as being more than just motivational or illustrative, but rather as being instrumental in encouraging meaningful and deep learning. When explaining my research to non-educationalists the most common response is: “Of course pictures make learning easier and better!” However, this opinion was not always predominant, and pictures were not always taken for granted as an ingredient of learning materials. The question, which has been researched since the 1960s, is: Do these pictures, tables, graphs, et cetera really do what we think they do and why?

The main topic of this thesis is multimodal representation used in learning. The aim is to determine whether using multimodal representations in history learning tasks makes a difference for the learning outcomes achieved and/or the learning process carried out.

1.1 Theories on learning with multimodal representations

Multimodal representations combine two or more modes of representation, for example, a verbal representation in the form of a text with one or more types of visualisation (schematic and/or depictive). Combining text and visualisations requires translating visual information into verbal information and vice versa, and then relating them to each other. These multimodal representations allow the learner to focus on different aspects of the topics being presented and on connections between those topics, thus promoting deep learning (Ainsworth, 1999). Positive effects of using multimodal representations in learning materials have been found within the domain of science and technology and in the context of individual use of presented multiple representations (Mayer, 2001).

There is a large body of research on learning with visualisations going back to the 1960s, where visualisations usually consisted of schematic (e.g., graphs, tables, flowcharts) or depictive representations (e.g., realistic drawings, photos) as opposed to verbal representations (e.g., stories). In the studies discussed in this thesis, multimodal representations combine verbal representation (textual) with one or more types of visualisation (schematic and/or depictive). Although research on the value of visualisations has continued in the intervening years, few attempts have been made to describe and compare different types of visualisations. In an empirical study, Lohse, Biolsi, Walker, and Rueter (1994) analysed the way in which multimodal representations and visualisations were perceived. Their study describes different types of visualisations in a taxonomy of visualisations. This framework used 10 scales to describe 60 different visualisations, resulting in 11 categories of visualisations, such as network charts, cartograms, process diagrams, structure diagrams, and time charts. An overview is shown in Table 1.1. Although the representations in their study were not necessarily educational in nature, their classification scales show how people react to different types of representations, giving insight into the different functions of these visualisations. Pictures, for example, were seen to convey an integrated piece of information in its entirety, but not to convey a great deal of information, whilst tables were judged to present individual pieces of information.

Table 1.1. Definitions for types of visual representations (from Lohse, Biolsi, Walker, & Rueter, 1994) and examples.

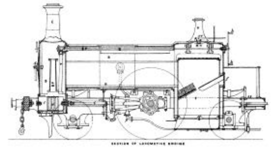
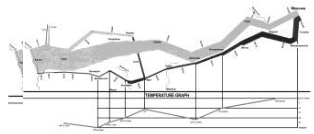
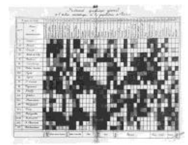
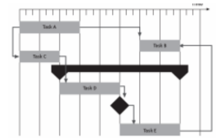
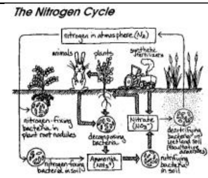
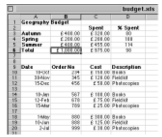
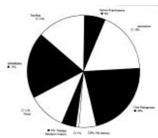
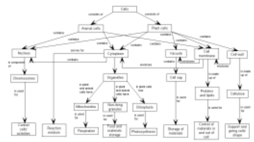
Types of visual representations	Definitions	Examples	
Structure diagrams	A static description of a physical object. The spatial data expresses the true coordinate dimensions of the object.	Cross-sectional view of an engine	
Cartograms	Spatial maps that show quantitative data.	Flow map of Napoleon's march to Moscow	
Maps	Symbolic representations of physical geography. Depict geographic locations of particular features using symbols or lettering.	Topographic maps	
Graphic tables	Tables that use graphic encoding – e.g., shading – of numerical information.	Matrix of chalet types and facilities in a holiday park brochure	
Icons	Impart a single interpretation of meaning for a picture. Used when the meaning is apparent to the target audience.	Traffic signs	
Time charts	Emphasise temporal data.	Gantt chart Timeline of WW II	

Table 1.1. (Continued)

Types of visual representations	Definitions	Examples	
Process diagrams	Describe interrelationships and processes associated with physical objects. The spatial data expresses dynamic, continuous, or temporal relationships among objects.	The nitrogen cycle The Krebs cycle	
Photo-realistic pictures	Realistic images of an object or scene.	'The Night Watch' by Rembrandt	
Tables	An arrangement of words, numbers, signs, or combinations of them to exhibit a set of facts or relationships in a compact format.	Spreadsheet budget	
Graphs	Encode quantitative information using position and magnitude of geometric objects.	Pie chart	
Network charts	Show the relationships among components. Symbols indicate the presence or absence of components. Correspondences among the components are shown by lines, arrows, proximity, similarity, or containment.	Organizational chart	

Many of the visualisations discussed by Lohse et al. (1994) are multimodal in nature. A good example of this multimodality is a cartogram, which contains a map showing the contours of an area, schematic information about the location of cities, verbal information – such as names of places and countries – and information about the spread or distribution of specific phenomena – such as industry or religion. The multimodal representations used in the studies described here are limited to cartograms, network charts, structure and process diagrams (incorporating pictures), and coordinating time charts (timelines).

Multimodal representations are used for different types of learning tasks and goals – both in research and in actual learning settings. They are used, for example, to support problem solving (e.g., Cox & Brna, 1995; Zhang, 1997), where they can serve cognitive offloading. Multimodal representations can also support memory and concretisation in the acquisition of declarative knowledge (e.g., Fischer, Bruhn, Gräsel, & Mandl, 2002). In reading comprehension, multimodal representations can help with text interpretation and focusing (e.g., Carney & Levin, 2002), and in collaborative learning, multimodal representations can function as a common referent (e.g., Munneke, Van Amelsvoort, & Andriessen, 2003).

Research has shown that multimodal representations can contribute positively to learning processes and learning outcomes. However, the results are not unequivocal. Most research on learning with multiple representations is based on Paivio's Dual Coding Theory and on Mayer's Cognitive Theory of Multimedia Learning (Paivio, 1991; Mayer, 2001). Dual Coding Theory (DCT) assumes that information is processed through one of two separate but interrelated channels: the verbal channel or the visual channel. It predicts that adding pictures to text will benefit learning in most cases, as pictures are more likely to be processed both verbally and visually. This will result in more elaborate encoding, and the learner is provided with more retrieval cues (Paivio, 1991).

Table 1.2. Definitions for principles of multimedia learning (from Mayer, 2001).

Principles	**Definitions**
Multimedia	Students learn better from words and pictures than from words alone.
Spatial contiguity	Students learn better when corresponding words and pictures are presented near rather than far from each other on the page or screen.
Temporal contiguity	Students learn better when corresponding words and pictures are presented simultaneously rather than successively.
Coherence	Students learn better when extraneous words, pictures and sounds are excluded rather than included.
Modality	Students learn better from animation and narration than from animation and on-screen text.
Redundancy	Students learn better from animation and narration than from animation, narration and equivalent on-screen text.
Individual differences	Design effects are greater for low-knowledge learners than for high-knowledge learners and for high-spatial learners than for low-spatial learners.

Mayer's Cognitive Theory of Multimedia Learning (CTML) builds on DCT and makes three main assumptions: (1) information is processed through two separate subsystems for verbal and nonverbal information, (2) meaningful learning involves conscious processing, and (3) there is a limit to the capacity of working memory (Moreno & Valdez, 2005). The assumptions lead to a set of seven principles for multimedia learning, shown in Table 1.2 (Mayer, 2001), many of which are based upon Cognitive Load Theory (Sweller, 1988; Sweller & Chandler, 1991; Van Merriënboer & Sweller, 2005). In his research, Mayer has found substantive evidence to support his theory (Mayer, 2003; Mayer & Sims, 1994).

However, although DCT and CTML and other previous research on learning with multimodal representations offer valuable insights, they do have their limitations, most notably in (1) their generalizability from the technical or natural science domains to other domains, (2) their tendency to focus on presented representations rather than construction of representations by learners, and (3) their focus on individual settings and learning outcomes rather than on collaborative settings and learning processes.

1.2 Limitations of earlier research on multimodal representations

1.2.1 Multimodal representations in different domains

The topics of the materials used in most research have been drawn from the natural sciences, such as biology or physics, where specific concrete concepts are combined with very clear-cut depictive representations like the ones shown in Figures 1.1a and 1.1b. However, not all domains are typified in such a way. Thus, the findings in one domain do not necessarily generalise to another. Research attempting to replicate Mayer's CTML work in the domain of educational sciences and pedagogy (De Westelinck, Valcke, De Craene, & Kirschner, 2005) did not confirm Mayer's multimedia principle and spatial contiguity principle, and De Westelinck et al. propose that different fields of knowledge raise different possibilities for the use of multimodal representations.

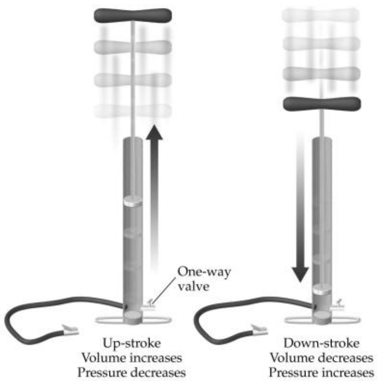


Figure 1.1a. Operation of a manual bicycle pump (from Tro, 2003).

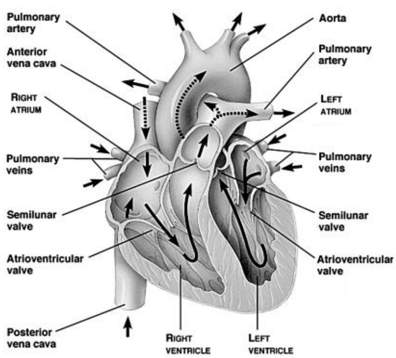


Figure 1.1b. The human heart (from Vogt, 1999).

Multimodal representations in the natural sciences are characterised by a consistent visual language – e.g., arrows, shapes in flowcharts or circuits – and the information represented often deals with 'How things work' by showing the underlying mechanisms. Examples of this are topics like volcanoes, engines, or lightning.

Information types in a humanities domain like history, on the other hand, are very different from this. In the school subject of history, prior to thinking or reasoning historically, students need to construct a chronological frame of reference that they can use to investigate, describe, explain and evaluate historical and contemporary phenomena (Van Drie, 2005). A prerequisite for remembering and applying such a frame of reference is that students can visualise the historical phenomena that are part of the chronological overview to be acquired. These may include events and people, developments, structures or systems, and underlying themes (Leinhardt, 1993). Research has shown that students have particular difficulty forming a notion of complex historical developments and structures (Carretero, Asensio, & Pozo, 1991; Husbands, 1996; Van Boxtel, 1994). Both temporal and causal relationships are important in this. Visual and multimodal representations, such as pictures, animations, timelines, matrices and (causal) diagrams can be used to represent the different types of historical phenomena and relations.

Finally, though historical and societal contextualisation is necessary in almost all domains, it is particularly import in a domain such as history. The understanding and interpretation of original realistic or depictive visualisations (e.g., paintings or drawings of everyday scenes, historical events, landscapes, or cartoons; Sauer, 2003) is strongly influenced by an understanding of the historical, societal and even artistic conventions of the time-period in which they were produced. The use of perspective, though considered normal now, was not present in historical paintings until the Late Middle Ages. Specific garb and the way that it was worn had special meaning in the period that Rembrandt painted. Not knowing or understanding such conventions makes it almost impossible to properly understand and interpret such depictions. In addition, an original picture may have been heavily influenced by the interpretation of its maker, making the source unreliable, and often there simply are no original pictures, because paintings of unpopular topics were rarely commissioned. Multimodal representations constructed especially for educational purposes – for example, a reconstruction of a Roman army camp or of a medieval village – can aid the reader/viewer/learner in such situations by focusing on certain aspects and eliminating distracting details.

1.2.2 Activity continuum: presented vs. construction

Most research on learning with visual or multimodal representations deals with presented representations. Often, the participants are instructed simply to ‘study’ the materials, and not required to perform any other activity, such as labelling parts of a picture. These types of tasks can be viewed as different positions along a continuum, with representations constructed by others and presented to the learners at one end, and representations constructed by the learners themselves (e.g., drawings) at the other end (Van Meter & Garner, 2005). Somewhere in between those two extremes on the continuum are representations where some parts are provided to the learners and some parts are constructed by the learners – for example when learners have to sequence pictures that represent the components of a process. Both of the mainstream theories, DCT and CTML are primarily based on research with presented representations.

The mediating function of multimodal representations is determined – among others – by the nature of activities with the representations (Peeck, 1993). Most research has paid little attention to activities of constructing or adapting multimodal representations (Scaife & Rogers, 1996). Still, theory seems to suggest that assembling and constructing multimodal representations – as opposed to simply presenting them – more strongly encourages articulation of ideas and content, discussion, and deep processing (Cox, 1999). Stern, Aprea and Ebner (2003) compared reasoning with ready-made representations (linear graphs) with reasoning with self-constructed representations in the domain of economics and found that self-constructed representations resulted in more elaborated and organised knowledge structures. Van Amelsvoort, Andriessen, and Kanselaar (in press) found that constructed diagrams deepened and broadened argumentative discussion more than presented diagrams.

Some studies on collaborative learning with multimodal representations in small groups have shown that new domain knowledge can be built by articulating or representing ideas and by making connections in the shape of concrete, visual structures in concept maps (Van Boxtel, Van der Linden & Kanselaar, 2000; Roth & Roychoudhury, 1994). By comparing interaction processes and learning outcomes of learning activities with different types of representations we can gain insight into the possibilities, and even more so the limitations of each of the representational forms in (collaborative) history learning. The different types used in the studies in this thesis are: making a timeline, completing a storyboard (process diagram), completing a causal map (network chart), describing a historical image (structure diagram), and completing a cartogram.

1.2.3 Individual vs. collaborative settings

The use of multimodal representations has mainly been studied in individual settings (Mayer & Chandler, 2001; Ainsworth, 1999; Schnotz, 1993; Larkin & Simon, 1987). In these studies multimodal representations are often seen as cognitive thinking tools: tools for remembering, thinking and problem solving (Jonassen, Reeves, Hong, Harvey, & Peters, 1997). Less research has been done in collaborative settings. In studies on the facilitating role of multimodal representations in collaborative learning environments, multimodal representations are often described as communicative tools (Reimann, 2003; Teasley & Rochelle, 1993; Suthers & Hundhausen, 2001; Erkens, Kanselaar, Jaspers, & Schijf, 2001; Van Drie, Van Boxtel, Jaspers, & Kanselaar, 2005). It is thought that the collaborative multimodal product under construction is incorporated in the co-construction of knowledge, because pupils refer to different elements of the product and because it gives them a shared frame of reference for exchanging and negotiating meaning (Roth & Roychoudhury, 1994). Concept mapping, for example, engages students in discourse on relevant conceptual relationships. The required group product makes students focus on pivotal principles in the domain, and thus stimulates abstract discussion. Collaborative concept mapping is an open task with no predetermined answers, provoking negotiation of meaning. The product serves as a visible representation that can facilitate discourse on abstract concepts and relationships. Recent research on collaborative construction of representations has rendered positive results for both learning processes and learning outcomes (Suthers & Hundhausen, 2003; Van Drie, Van Boxtel, Jaspers, & Kanselaar, 2005; Van Boxtel, Van der Linden, & Kanselaar, 2000).

In addition to these advantages, collaborative learning makes it possible for researchers to study learning processes through analysis of student discourse. In many studies dealing with co-construction of visualisations, the question asked is whether a difference in type of visualisation will result in a difference in learning activities. To this end, research on learning through co-construction of visualisations has focused on the communicative functions (Suthers & Hundhausen, 2003), on the activities in the software (Bodemer, Ploetzner, Bruchmüller, & Häcker, 2005), on procedures and task management (Erkens, Jaspers, Prangsma, & Kanselaar, 2005), or on the content (Van Drie, 2005; Fischer, Bruhn, Gräsel, & Mandl, 2002).

1.2.4 Integrating multiple multimodal representations

Single units of the curriculum usually contain different types of information. Certain types of information are best presented through a specific type of representation. For example, the passing of time is more easily represented in a timeline than in a structure diagram. A single multimodal representation may be insufficient to adequately represent all aspects of a particular domain or topic, or the representation may become too complicated to interpret. De Jong et al. (1998) give three reasons for introducing more than one type of representation. First, multiple representations allow the tuning of the domain information and the representation. A second rationale is the idea that multiple representations will result in more flexible knowledge. A final reason is based upon the assumption that a specific order of representations will facilitate the learning process. For example, in order to understand a complex system, students can be provided with pictorial representations of the individual components of the system followed by a structure diagram showing how the system works as a whole. Figure 1.2 shows an example from research by Van der Meij and De Jong (2003) on multiple physics representations integrated in a single simulation. Manipulating one representation resulted in a change in the other representations. These examples indicate that the use of multiple representations calls for coordination of the information in the different representations. The relations between the different representations are important for acquiring a consistent conceptual frame that can be reused in new situations. However, research on the use of different representations shows that this coordination does not occur spontaneously (Ainsworth, 1999).

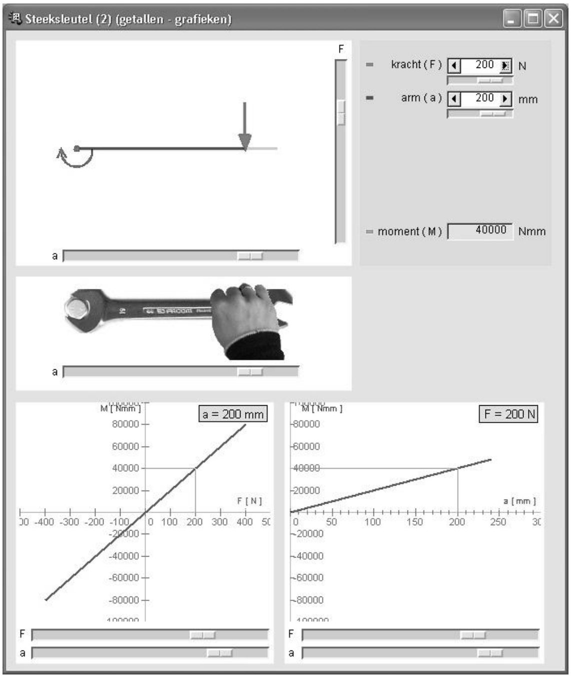


Figure 1.2. Multiple representations in a single simulation (from Van der Meij & De Jong, 2003).

Acquiring a chronological frame of reference calls for visualization and conceptualization of specific periods in history and of relationships between aspects of these periods. Research has shown that students have trouble developing a consistent 'chain of events and developments'; the result is that they cannot remember the chronological overview or reuse it, causing confusion of phenomena and concepts (Beck & McKeown, 1994; Leinhardt, Stainton & Virji, 1994). The timeline is a representation type that is typical for history that functions as a tool to represent time and give a chronological overview. However, a timeline with dates or periods and textual descriptions of historical phenomena only visualises temporal relationships. It does not show the underlying causal relationships between the events and phenomena it contains. Moreover, a timeline only gives limited support in visualizing and grasping the character of an era. Other representations (such as historical pictures, animations, matrices comparing periods, numerical information and causal networks) might combine with a timeline to facilitate visualization and construction of relationships, requiring explicit linking of the different representations. The historical phenomena being visualised need to be contextualised in time. Aspects of continuity and change cannot be recognised and causal relationships cannot be drawn until the temporal relationships are clear. In addition, causal relationships do not just require insight in temporal relations, but also in the types of phenomena to be explained. Connecting different types of information is thus crucial in history learning. Integrating different multimodal representations can support this, for example through a timeline consisting of elements like pictures, verbal descriptions, and arrows indicating causal relations. The studies in this thesis look at the possible added value of a method in which different types of representations are related to each other: a timeline incorporating verbal, depictive and schematic representations.

1.3 Problem definition and research questions

The central issue in this research is how multimodal representations can support verbalization and co-construction of meanings and relations, and thus contribute to the attainment of domain specific conceptual knowledge. The studies in this thesis look at learning tasks with multimodal representations, and at integrating different representations in a timeline, and contrast these with learning tasks with textual representations. In other words: How can learning *activities* with *different* multimodal representations *in relation to each other* contribute to meaningful learning? The central question addressed in this thesis can now be specified as follows:

> How does making and connecting different types of multimodal representations affect the collaborative learning process and the acquisition of a chronological-conceptual frame of reference in 12 to 14-year-olds enrolled in pre-vocational education?

The research questions derived from this central question are:

1. What are the effects of the general type of co-constructed representation – textual, multimodal, or integrated multimodal – on the acquisition of a chronological-conceptual frame of reference? (Chapter 2)
2. What are the effects of the collaborative completion and construction of integrated multimodal representations versus textual representations on the content of the student dialogue? (Chapter 3)

3. Are there any differences in domain-specific content of the student dialogue between collaborative completion and construction of different types of multimodal representations – structure diagram, network chart, process diagram, or cartogram? (Chapter 3)
4. What are the effects of combining text and different types of visualisations – abstract and concrete – in collaborative history tasks on learning historical phenomena and concepts? (Chapter 4)

1.4 Outline of the three studies in this thesis

All studies in this thesis worked with 12 to 14-year-old participants in the first year of pre-vocational secondary education (VMBO). A majority of Dutch pupils in secondary school (some 60%) attend this level of schooling. The language proficiency of these pupils is often lower than desired. History as a school subject is part of the compulsory curriculum only for the first two years for these pupils, so there is little time for developing a chronological frame of reference.

The topics of all learning materials used in the three studies reported on here were all taken from the Early Middle Ages. This period – 500 to 1000 AD in (Western) European history – was selected for several reasons. First, the period includes a wide range of types of phenomena that are dealt with in the history curriculum. In addition, the existing curriculum for this period includes different types of developments and social structures that are closely related. It marks a turning point in Western European history with the decline of the Roman Empire and the subsequent development of manorialism (‘hofstelsel’), and these two developments can only be understood in relation to each other. A number of important socioeconomic, political and religious changes took place during the Early Middle Ages. The period includes some very abstract concepts, and life during this period was in many ways very different from our pupils’ lives. On top of that, there is very little original visual material available from the Early Middle Ages that can help pupils in shaping a notion of this period.

1.4.1 Study 1 – Effects of multimodal representations: learning outcomes

This study answers the first research question. Based on the literature on learning with multimodal representations and what is important for the acquisition of a chronological-conceptual frame of reference in history we expected that the collaborative completion and construction of multimodal representations would contribute more to learning than the collaborative completion and construction of textual representations. Moreover, we expected that integration of such multimodal representations in a timeline would strengthen this effect. Chapter 2 presents the individual learning results of an experiment in which three types of tasks were compared: textual tasks, multimodal tasks, and multimodal tasks integrated in a timeline. The aim of this study was to determine the differences (if any) in learning outcomes, both in the short term and in the long run. In addition, the chapter discusses an exploratory analysis of the dialogues of two dyads working on the timeline tasks.

1.4.2 Study 2 – Effects of multimodal representations: learning processes

This study answers research questions 2 and 3. The first study shows that there were some short term differences in learning outcomes between the textual and timeline conditions. This raises the question – answered in Study 2 – whether the content quality of the learning process might mediate this in the timeline condition. Also, the preliminary exploratory analyses of pupil discourse discussed in the report on Study 1 seem to suggest substantial differences in dialogue quality between tasks within the timeline condition, even though all four tasks were designed to elicit productive interaction. This is explored further in Study 2. Student dialogues were recorded during data collection for Study 1. The utterances were coded as procedural, social or content utterances. Content utterances were analysed further for level (core vs. auxiliary), integration of modes, and concept use. Study 2 is reported in Chapter 3.

1.4.3 Study 3 – Effects of concrete and abstract visualisations

The multimodal tasks in Studies 1 and 2 combined abstract and concrete visualisations. As a result, little can be said about differences between effects of abstract, concrete and combined representations. In addition, Study 1 seemed to indicate that the tasks that most obviously combined concrete and abstract visualisations (pictures and diagrams) elicited more content-related dialogue. Study 3 answers the fourth research question, and describes a follow-up experiment that deals with differences in learning results between different types of multimodal representations: abstract, concrete and a combination of abstract and concrete. The short term and long term learning outcomes for the three multimodal conditions were also compared to those for a textual condition. As learning history requires understanding both abstract and concrete phenomena, concepts, and relations, we expected that a combination of abstract and concrete visualisations would have the greatest potential for enhancing learning. Study 3 is reported in Chapter 4.

A general discussion of the studies described in this thesis can be found in Chapter 5.

2. Developing a 'big picture': Effects of collaborative construction of multimodal representations in history[1]

Many pupils have difficulties with the abstract verbal information in history lessons. In this study we assessed the value of active construction of multimodal representations of historical phenomena. In an experimental study we compared the learning outcomes of pupils who co-constructed textual representations, multimodal representations, or multimodal representations integrated in a timeline. Eighty-five pupils in pre-vocational secondary education, aged 12 to 14, worked in dyads on a series of four history tasks. All pupils took a pre-test, post-test and retention test. Results show that working on multimodal representations integrated in a timeline leads to higher short-term results than co-constructing textual representations. Dialogue analyses for two dyads working in the condition with multimodal representations integrated in a timeline indicate that the extent to which pupils verbally integrate textual and visual information differs for the four different tasks.

2.1 Introduction

Think back to your history classes in secondary school. Your knowledge of secondary school history is likely to be fragmentary and unstructured. Shemilt (2000) reports in an evaluation of the Schools History Project in Great Britain: "few fifteen year-olds are able to map the past; even fewer can offer a coherent narrative [….] for many, the 'event-space' within which [historical narrative frameworks] form and grow is *incoherent* and lacking in order or meaning" (p. 86). Lee (2005), who interviewed students about historical change, reports that the students did not show convincing signs of access to an overall framework of the past.

Acquiring a chronological frame of reference is vital in learning and understanding history. However, many pupils have difficulties acquiring a coherent overview of significant historical events and developments, and they confuse different phenomena and the concepts used to describe them. For example, any ruler is easily called a king, and his house a castle, even when these are anachronisms for a given period and location. Many of the phenomena in history are abstract in nature, and they are related in complex ways. Think, for example, of forms of government, such as democracy.

The prominence of abstract verbal information in history can be a problem for relatively weak readers in lower secondary education (Hacquebord, 2004). These pupils are expected to benefit particularly from efficient use of multimodal representations. Much of history can be visualised through different types of representations, such as

[1] Prangsma, M. E., Van Boxtel, C. A. M., & Kanselaar, G. (in press). Developing a 'big picture': Effects of collaborative construction of multimodal representations in history. *Instructional Science*.

pictures, cartograms, process and structure diagrams, network charts, and, naturally, timelines. The aim of this study is to assess the value of *active construction* of multimodal representations in supporting learners trying to acquire knowledge of historical phenomena in a chronological frame of reference.

2.1.1 Learning with multimodal representations

Research has given us insights into conditions for effective use of pictorial representations in addition to verbal ones, as well as into the processes of mental model construction through multimedia learning. Combining text and visualisations requires translating visual information into verbal information and vice versa, and then relating them to each other. These multimodal representations allow the learner to focus on different aspects of the topic to be tackled and on connections between topics and aspects, thus promoting deep learning (Ainsworth, 1999). Positive effects have been found within the domain of science and technology and in the context of individual use of presented multiple representations. Most research on learning with multiple representations is based on Paivio’s Dual Coding Theory and on Mayer’s Cognitive Theory of Multimedia Learning (Paivio, 1991; Mayer, 2001).

Dual Coding Theory (DCT) assumes that information is processed through one of two channels: the verbal channel or the visual channel. It predicts that adding pictures to text will benefit learning in most cases, as pictures can be processed both verbally and visually. This will result in more elaborate encoding, and the learner is provided with more retrieval cues (Paivio, 1991). Mayer’s Cognitive Theory of Multimedia Learning (CTML) strongly builds on DCT and makes three main assumptions: (1) information is processed through two separate subsystems for verbal and nonverbal information, (2) meaningful learning involves conscious processing, and (3) there is a limit to the capacity of working memory (Moreno & Valdez, 2005). The assumptions lead to a set of seven principles for multimedia learning: spatial and temporal contiguity, coherence, modality, redundancy, individual differences and the multimedia principle (Mayer, 2001; see Table 1.2 in Chapter 1). In his research, Mayer has found substantive evidence to support his theory (Mayer, 2003; Mayer & Sims, 1994).

However, although DCT and CTML offer valuable insights, they do have their limitations, most notably in (1) their applicability to different types of representations, (2) their generalizability to different domains, and (3) the tendency of research to focus on presented representations rather than construction of representations by learners. These three limitations will now be briefly discussed.

First, DCT and CTML have been tested on a limited number of different types of representations, mainly process diagrams showing ‘How things work’ (Mayer, 2003; Moreno & Valdez, 2005; Lohse, Biolsi, Walker, & Rueter, 1994). Cox (1999) remarked that a diagram is not always worth ten-thousand words, because its worth depends on the type of diagram, which meanings it represents, who produced or uses it, and the nature of the task. Schnotz and Bannert (2003) suggest that the structure of depictive representations (as opposed to descriptive representations) directly influences the structure of the mental model constructed from it. In a study with different types of graphics they found that the structure of the representation offered to learners was reflected in the structure of their mental models of the topic. This implies that depictive representations need to be chosen with extreme care to match the class of phenomena, task and the intended mental model. It seems reasonable, then, to make further distinctions between different subtypes of depictive representations. Jones, Pierce, and

Hunter (1988, 1989) present a taxonomy of types of visual representations corresponding to different types of text structures: spider map, series of events chain, continuum/scale, compare/contrast matrix, problem/solution outline, network tree, fishbone map, human interaction outline and cycle. Lohse et al. (1994), on the other hand, developed a structural classification on the basis of ten scales of characteristics – e.g., concrete-abstract, spatial-nonspatial – which resulted in eleven types of visual representations: structure diagrams, process diagrams, maps, cartograms, tables, graphic tables, pictures, icons, time charts, network charts, and graphs. Their taxonomy is research-based, practical and exhaustive and we will use it in our description of the representations in our research (see Table 1.1 in Chapter 1).

Second, previous studies have been limited mainly to the domains of secondary school mathematics and physics. Research replicating Mayer's work in the domain of Educational Sciences and Pedagogy (De Westelinck, Valcke, De Craene, & Kirschner, 2005) did not confirm Mayer's multimedia principle and spatial contiguity principle, and De Westelinck et al. propose that different fields of knowledge raise different possibilities for the use of multimodal representations. The role of multimodal representations for history learning remains largely unclear.

Third, most of the research concerned with learning with multiple representations focuses on presented representations rather than on active construction of representations by learners. The mediating function of multiple representations is determined – among other things – by the nature of the activities elicited by the representations (Peeck, 1993). Research, however, has paid relatively little attention to activities of individual and collaborative construction or adaptation of multimodal representations (Scaife & Rogers, 1996). Current trends in the field of learning and instruction stress the importance of active knowledge construction and collaborative learning. Cox (1999) stated that the process of translating information from a linguistic representation to a visual representation might be more effective than translation from one representation to another within the same modality. This idea is consistent with DCT (Paivio, 1991). A meta-analysis by Horton, McConney, Gallo, Woods, Senn, and Hamelin (1993) reports a modest positive effect of construction of concept maps on student achievement. Several researchers suggest that multiple representations support deeper understanding when students integrate information from different types of representations (e.g., Ainsworth, 2006). Bodemer, Ploetzner, Bruchmüller, and Häcker (2005) found that asking students to actively relate textual components to components of a visualisation had a beneficial effect on learning when the learning material was particularly difficult and complex. However, most studies on learning with multiple representations do not discuss the extent to which learners actively relate textual and visual information.

In addition to fostering knowledge construction, multimodal representations can also function as communicative support in collaborative learning (Reimann, 2003). Roth and Roychoudhury (1994) argue that concept mapping as a collaborative activity encourages communication and negotiation of meaning. Concept mapping engages students in discourse on relevant conceptual relationships. The required group product forces students to focus on pivotal principles in the domain, and thus stimulates abstract discussion. Collaborative concept mapping is an open task with no predetermined answers, and this provokes negotiation. The product serves as a visible representation that can facilitate discourse on abstract concepts and relationships. Students can refer to the concept labels and the propositions in the emerging representation while verbalizing their ideas and negotiating meaning. Recent research on collaborative construction of representations has rendered positive results for both learning processes and learning outcomes (Suthers & Hundhausen, 2003; Van Drie, Van Boxtel, Jaspers, & Kanselaar, 2005; Van Boxtel, Van der Linden, & Kanselaar, 2000).

2.1.2 The potential of multimodal representations for the acquisition of a chronological frame of reference

Our research focuses on a domain that is relatively unexplored regarding research on learning with multimodal representations: the domain of history. A chronological frame of reference is the knowledge base that is used when reasoning about the past. It consists of knowledge about: (1) historical phenomena, (2) temporal and causal relations, and (3) concepts describing phenomena and relations. Research by Beck and McKeown (1994) has shown that pupils have difficulty developing a coherent chain of events, and that the schemas pupils use are too general to offer ready slots to fit the specific information that they might have gleaned. In addition, the specific information is too sparse to be useful in connecting it to more general information. Furthermore, pupils have particular difficulty forming a notion of complex historical developments and structures (Husbands, 1996; Carretero, Asensio, & Pozo, 1991). Making these developments, structures, temporal and causal relations visible through pictures and diagrams can render abstract phenomena and relations more explicit.

The first component of a chronological frame of reference consists of different historical phenomena: the events, structures and themes of an era. Leinhardt (1994) makes a distinction between different types of phenomena that are central to instructional explanations in history classes: events, structures, themes and metasystems. Events are narratives of the actions of people and institutions, limited in time and space, such as civil wars. Structures are the more constant social elements with descriptive features, for example rural or industrial society. Themes are the clarifying notions at the core of historical understanding of people and nations over time, such as tensions between North and South in the United States. Metasystems include the metacognitive tools of history, for example analysis and perspective taking. Events, structures and themes are specific classes of historical phenomena that may require different types of representations. Events, for example, are often represented by narratives, such as a narrative of the rise and fall of the Roman Empire or a narrative of the French Revolution (Husbands, 1996). Such narratives can be textual, but they can also be visually represented, for example in a timeline or a comic strip.

The second component of a chronological frame of reference is knowledge of relations between historical phenomena, both temporal and causal. Temporal relations can be represented by a timeline (Hoodless, 1996). Constructing timelines can help to sequence events, and to develop awareness of duration and 'key dates' or landmarks (Stow & Haydn, 2001). Dawson (2004) emphasises both the active construction of timelines (instead of looking at completed ones) and the inclusion of images rather than just words and dates. However, a timeline with dates or periods and textual descriptions of historical phenomena only visualises temporal relationships. It does not show the underlying causal relationships between the events and phenomena. Other representations (such as historical pictures, animations, matrices comparing periods, and causal diagrams) might combine with a timeline to facilitate visualization of phenomena and casual relationships. Spatial representations can make it easier to understand relations (Larkin & Simon, 1987), and pictures may elucidate historical figures, situations and landscapes. Such a combination of diagrams and pictures is also in line with Friedman (1982), who found that in history, children under 10 preferred to work with verbal or pictorial representations over spatially organised diagrams, such as timelines. Combining diagrams, pictures and text requires explicit linking of the different representations. The historical phenomena being visualised need to be contextualised in time. Continuity, change and causality cannot be recognised unless the temporal relationships are clear. In addition, causal relationships do not just require insight into temporal relations, but also into the types of phenomena to be explained. Connecting different types of information (that are represented in different representational formats) may thus be crucial in history learning.

The third component of a chronological frame of reference is knowledge of concepts used to describe phenomena and relations. Using historical terminology is an important part of history learning, and it involves both methodological concepts such as change, continuity and causes, as well as substantive concepts, such as feudalism and Enlightenment. Understanding the big picture requires generalisation through a range of abstract concepts (Hunt, 2000). Domain specific concepts are tools to question, think about, describe, analyse, synthesise and discuss historical phenomena (Van Boxtel & Van Drie, 2004). Therefore, historical concepts are an important component of multimodal representations used to display historical phenomena and relations. In line with the spatial contiguity effect (Mayer, 2003), it can be expected that students learn more deeply when relevant historical concepts are placed near the corresponding pictures.

2.1.3 Problem definition

The focus of our research is on the active *construction* of multimodal representations in *collaborative learning tasks* in *history*. We address the following question: What are the effects of the type of constructed representation on the acquisition of a chronological-conceptual frame of reference? We expect to find that construction of a combination of visual and textual representations has a positive effect on the acquisition of a chronological-conceptual frame of reference, and that construction of a combination of visual and textual representations integrated in a timeline has an even larger positive effect, compared to construction of textual representations. The hypotheses are in line with the idea that active construction of multimodal representations and the integration of verbal and pictorial information helps pupils to develop a 'big picture' of historical phenomena that can be more easily remembered and transferred to new tasks.

2.2 Method

2.2.1 Participants and design

The pupils in our study (*N* = 143) were from six different first year classes in three different schools, with one history teacher for each school (pupils aged 12 to 14). The classes were all in pre-vocational secondary education (VMBO), which a majority of Dutch pupils in secondary school (some 60%) attend. The language proficiency of these pupils is relatively low (Hacquebord, 2004). History as a school subject is part of the compulsory curriculum only for the first two years for these pupils, so there is little time for developing a chronological frame of reference.

We conducted an experiment with dyads working in one of three conditions: textual representations (Text); multimodal representations (Visual); and multimodal representations integrated in a timeline (Timeline). Scores for the multiple-choice section of the pre-test were used for heterogeneous dyad composition: the pupils were divided into three groups with low, intermediate and high scores. Pupils from the group with the lowest scores and pupils from the group with the highest scores were teamed up with pupils from the group with intermediate scores. In classes where the intermediate scoring pupils outnumbered the low and high scoring ones together, intermediate-intermediate dyads were also formed. Due to absence of a number of pupils at the start of the first task lesson, some classes did not contain enough intermediate scoring participants, so several low-low and high-high dyads were also formed. In total, the final sample contains 4 participants from low-low dyads (2 in the Visual condition and 2 in the Timeline condition), and 4 participants from high-high dyads (2 in the Text condition and 2 in the Visual condition). All dyads, except two, consisted of either two girls or two boys.

Given the space, time and attention required for the different conditions we chose to assign the dyads within each class randomly to one of two (instead of all three) conditions. In four out of six classes, the two conditions were put in different classrooms. We made sure that low/intermediate and high/intermediate scoring dyads were evenly distributed over the two conditions. All conditions were given the same amount of time to finish the tasks.

Due to a number of causes, 41% of the original sample had to be discarded: 32% of the Text condition, 49% of the Visual condition, and 37% of the Timeline condition. Pupils or dyads were discarded if one or more of the following reasons applied to them: absence during one or more task lessons; absence during the pre-test or the post-test; one or more tasks were missing or not finished; working individually (due to odd numbers). Pupils were included in the final sample in the following cases: if only the last task was partly missing; absence during the retention test; dyad partner absent during only one lesson. The high proportion of discarded dyads was mainly due to loss of concentration among pupils in school A, who were participating in Ramadan (a Muslim religious festival) during the period the experiment took place. As a result, few dyads in this school managed to finish all tasks. As a result, there is less diversity in the ethnic background of the entire sample (about 9% with a foreign background) than in the general population of Dutch pre-vocational secondary education (24%).

The final sample sizes are shown in Table 2.1. Although the participants worked in dyads, the table shows some odd totals. This is because participants were individually assessed, and from some dyads only one of the partners could be included for the reasons mentioned above.

Table 2.1. Number of pupils per condition for each school/teacher in final sample.

School/teacher	**Text**	**Visual**	**Timeline**	**Total**
A	3	1	0	4
B	12	24	10	46
C	15	6	14	35
Total	30	31	24	85

2.2.2 Experimental tasks

The experimental tasks were based on our experiences with a pilot study (see further). Working in pairs, pupils carried out a series of four tasks on the Early Middle Ages during three consecutive history lessons. The period of 500 to 1000 AD of Western European history – the Early Middle Ages – was selected for several reasons. First, the period includes the full range of phenomena that are dealt with in the history curriculum. Also, the period is representative of the difficulty level of, and the types of relations in other periods in the history curriculum. In addition, the existing curriculum for this period includes different types of developments and social structures that are closely related. It marks a turning point in Western European history with the decline of the Roman Empire and the subsequent development of manorialism, and these two developments can only be understood in relation to each other. A number of important socioeconomic, political and religious changes took place during the Early Middle Ages. The period includes some very abstract concepts, and life during this period was in many ways very different from our pupils' lives. On top of that, there is very little original visual material available from the Early Middle Ages that can help pupils in shaping a notion of this period. The specific task content was chosen on the basis of the 2001 report of the (Dutch) Committee of History and Social Studies that proposes ten eras with their specific aspects for the history curriculum in Dutch schools. In this proposal, the Early Middle Ages are called "the time of monks and knights", and its specific aspects are the spread of Christianity, manorialism, and the rise and spread of Islam. The four tasks each had different types of content.

Pupils in all conditions worked in dyads and started out by reading the same accompanying texts. These were two to three pages in length for each task (of which there were four, see further), including appropriate illustrations. These illustrations were not actively used in the construction task that followed, but were provided to improve basic understanding of the text for all conditions. Important concepts were printed in a bold typeface, and participants were encouraged to use these concepts in their answers. The texts were available to the students throughout the rest of the experiment.

The tasks that were used in the Visual and the Timeline condition were designed according to Mayer's (2003) principles for multimedia learning. According to these principles students learn more when words and pictures are combined (multimedia effect), when extraneous material is excluded (coherence), and when words are placed near a corresponding picture (spatial contiguity). The representations for the four tasks can be categorised as follows according to the taxonomy by Lohse et al. (1994): (1) process diagram (decline of the Roman Empire; see Appendix A Figure A.1), (2) network chart (effects of the fall of the Roman Empire; see Appendix A Figure A.2), (3) structure diagram (manorialism; see Appendix A Figure A.3), and (4) cartograms (spread of Christianity and Islam; see Appendix A Figure A.4). Appendix A shows the completed answer sheets for all four tasks in the Visual and Timeline conditions. All pictures used for the answer sheets were made especially for the experiments by a professional illustrator. The picture for the third task was in full-colour, while for the other tasks we used black-and-white drawings. The task products in the Timeline condition were linked to each other in a large timeline the size of two sheets of flip chart paper (about 60x140 cm). The tasks in the Text condition covered the same content, but only textual answers were required, and no pictures were provided on the answer and instruction sheets. An overview of the tasks is shown in Table 2.2.

Table 2.2. Contents of the collaborative tasks and corresponding activities in the experimental groups.

	Conditions		
Knowledge type and contents	**Text (control group)**	**Visual group**	**Timeline group**
Task 1 Development: The decline of the Roman Empire	Fill-in-the-blank Ordering sentences Summary question	Completing a storyboard: - Ordering pictures - Adding concepts and captions Summary question	Completing a storyboard: - Ordering pictures - Adding concepts and captions Summary question Adding text and colour to timeline
Task 2 Causal relations: The effects of the fall of the Western Roman Empire	Fill-in-the-blank Ordering sentences Summary question	Completing a causal network: - Choosing pictures - Adding concepts and captions Summary question	Completing a causal network: - Choosing pictures - Adding concepts and captions Adding arrows to timeline Summary question
Task 3 Structure: Manorialism	Answering questions Summary question	Describing a historical image: - Answering questions linked to details in the picture Summary question	Describing a historical image: - Answering questions linked to details in the picture Summary question Adding text, colour and arrows to timeline
Task 4 Development: Religion: spread of Christianity and Islam	Fill-in-the-blank	Completing two cartograms: - Fill-in-the-blank in tags linked to specific points on the maps - Colouring the spread of religions in the maps Summary question	Completing two cartograms: - Fill-in-the-blank in tags linked to specific points on the maps - Colouring the spread of religions in the maps Summary question Adding text and colour to timeline

The tasks were concluded with a summary question, such as: “Finish the following sentence with two causes and use your answers above and the text: The Roman Empire disappeared because … and because … “. The summary question was accidentally left out in Task 4 of the Text condition.

The pilot study mentioned at the beginning of this section ($N = 22$) was done to determine how much time was needed, and to try out the tests and the tasks in two different experimental conditions (Text and Timeline). On the basis of the pilot study, adequate time was determined as 3 lessons for four tasks. One task – a compare-and-contrast matrix – was discarded because it was too textual in nature. The task instructions and texts were also revised slightly on the basis of the pilot study.

2.2.3 Tests

The pupils were given the same individual knowledge test three times: a pre-test, a post-test, and a retention test. We included a retention test to investigate long-term effects. The test consisted of three parts that were administered separately to prevent questions giving away each other's answers. To ensure coverage of the learning goals, the test was based on a test matrix, which in turn was based on a matrix of the learning goals. The test questions are shown in Appendix B. Part A was a free association spider on the Early Middle Ages (see Appendix B Part A). Part B consisted of open questions, including both textual and visual questions (e.g., with pictures, maps, or timelines; see Appendix B Part B). Part C consisted of different types of multiple-choice questions (e.g., multiple response questions, choosing which came first in time; see Appendix B Part C). A total of 11 items were devoted to questions on chronology: 3 items in Part B, and 8 items in Part C. An overview of the test is shown in Table 2.3.

Table 2.3. Overview of the test and its parts.

Test unit	No. of items	Maximum score	Response format
A	1	no maximum	Association spider
B	8	8	Short answer
C	18	18	Multiple-choice
Total	27	no maximum	

We used Cronbach's alpha to determine test reliability. Table 2.4 shows Cronbach's alpha for the pre-test, post-test and retention test. Prior knowledge was low, so pupils had to resort to guessing on the pre-test, resulting in low homogeneity. Cronbach's alphas for the post-test and retention test were not high but were still considered acceptable. One item from part B was excluded from further analyses, because it had zero variance in the pre-test, and little variance in the post-test and retention test. This item asked about the Arab Empire, while most pupils seemed unable to disconnect the concept *empire* from the Romans. Table 2.4 shows the test reliability after deleting this item.

Table 2.4. Results of reliability analysis for pre-tests, post-tests, and retention tests.

Test	**Cronbach's Alpha (standardised item alpha)**
Pre-test ABC	.31
Post-test ABC	.58
Retention test ABC	.66

The three parts were scored on the basis of the task text content. No interrater reliability was calculated for Part A, as the percentage of agreement on 30 randomly chosen spider tests was 100%. The interrater reliability between two raters for Part B on 74 randomly chosen tests from different classes (28% of the total number of scored Parts B) was .89 (Cohen's kappa).

2.2.4 Setting and procedure

The pre-tests were administered four or seven days before the start of the experiment. The total time taken by pupils varied from 30 to 40 minutes. Pupils were given brief instructions about the research study and the tasks both during the pre-test lesson and at the start of the experiment. For all classes, the experiment was started in the first history lesson after administering the pre-test.

The classes were divided over two classrooms by condition. One of the researchers monitored one condition, and an assistant monitored the other condition, while the teacher switched classrooms from time to time. Apart from a short introduction before splitting up in conditions, there was no classroom instruction, nor did the teacher, researcher or assistant give feedback on content, only on completion ("Is it finished, yes or no?").

The participants started each task by reading a text on the task topic; the same text was used for all conditions. After reading, the dyads were given the task sheets and instructions. When a dyad had completed one task, they were given the next. The pupils had three lessons (about 150 minutes) to finish all four tasks. The conversation by the dyads was audio taped with recorders placed on their desks.

The post-test was administered either at the start of the first history lesson after rounding off the experiment, or the next day at the start of a lesson by the class mentor. The retention test was administered 33 to 49 days after the post-test (M = 37). Between the post-test and the retention test, regular classes were taught about the Early Middle Ages, but not about the specific topic of the experiment.

2.2.5 Analysis of student dialogue

From the final sample of 85 participants, two Timeline dyads (4 participants) were selected for a closer look at the discourse, in particular to get an idea of the occurrence of integration of text and representations during work on the different tasks. The dyads were selected from all Timeline dyads from the final sample. The chosen dyads had the most complete protocols (i.e., protocols were available for all four tasks). Also, the participants in these two dyads showed a strong increase in their scores between pre-test and post-test, so we expected to find indicators of relating textual to visual information in the student dialogues. Dyad A consisted of two boys: Allan (pre-test score = 14, post-test score = 23, retention test score = 20) and Adrian (pre-test score = 14, post-test score = 24, retention test score = 26). Dyad B consisted of two girls: Bridget (pre-test score = 7, post-test score = 14, retention test score = 11) and Betty (pre-test score = 12, post-test score = 27, retention test score = 20).

For each of the two dyads selected, the dialogues were typed out, coded and analysed with utterances as the unit of analysis. The protocols were analysed in several steps. First, the utterances were coded for their basic topic: Content, Procedural, Social or Other utterances. The focus in coding and analysing the interaction processes was on the content related part of the discourse. Content utterances included utterances about historical phenomena and relations, about pictures, or about the answers to be given on the answer sheet. Procedural utterances referred to physical characteristics of the task materials or to the spelling of the answer, or they were utterances for regulating the collaboration or the behaviour of the partner. Social utterances included all utterances by the dyad partners that were irrelevant to the task. The category Other included utterances by other participants, by the teacher, or by the experimenter. Unintelligible utterances were also coded as Other. The Procedural, Social and Other utterances were not investigated further, but served as a valuable context for interpreting the content utterances. Examples of these categories are shown in Table 2.5.

Table 2.5. Examples for the codes used for coding the dialogues: Social, Procedural, and Content utterances.

Code	Examples
Social utterances	"Are you going to the party on Friday?" "You know who called last night?"
Procedural utterances	"It's your turn now." "Do you have sticky tape?"
Content utterances	"There was little trade" "There was less trade" "The bridge has collapsed" "When the Romans left, bridges collapsed" "Trade" "Viking boats are cool"

The next step was to indicate which utterances were passages taken directly from the text, the instructions, or the answer sheet. As the participants did not originate these utterances, they were not coded within the content category. Again, these utterances provided a context for interpreting the content utterances. The steps thus described were also used to code and analyse a larger set of 20 dyad dialogues of Task 2 (see Chapter 3). The interrater reliability between two coders was calculated for four dialogue protocols (totalling 1060 utterances) and amounted to .74 (Cohen's kappa).

Table 2.6. Dialogue fragments with integrative utterances.

Speaker	**Relating text and timeline: Colouring and labelling the timeline (Task 1)**	**Integration**
Allan	Colour it yellow. The part that goes with Antiquity	0
Allan	Of all Antiquity it is	1
Adrian	But this is the Early Middle Ages and Antiquity is this	1
Allan	Then that is from 500 … from 400 to 500 I think	1
Adrian	Yes	0
Speaker	**Relating text and pictures – Describing and sequencing historical phenomena (Task 1)**	
Allan	Look, the Romans	1
Allan	This and this I think	1
Allan	But it is all …	0
Adrian	But here it is split, that is almost at the end	1
Adrian	And here is the fighting	0
Adrian	And that comes then too, and I think it is like this	1
Speaker	**Relating text and pictures – Describing causal relations between depicted phenomena (Task 2)**	
Betty	Um, but is this an effect of that, um that the effect of that?	1
Betty	Um, but this is actually, this is actually the effect of this one.	1
Betty	Because there was no protection anymore …	0
Teacher	Yes	0
Betty	… there was looting.	0
Speaker	**Relating text and geographic map – Colouring and labelling the map (Task 4)**	
Betty	Here it says you have to colour, um, red that were Christian during the Early Middle Ages. That's this, right?	1
Bridget	That is … Christian part	1

Finally, the content utterances were coded for integration of visual and textual information, which was defined as relating information from the task text to (a part of) the multimodal representations on the answer sheets (e.g., schemas, pictures, the timeline itself). The dialogues were coded by the two experimenters, and any disagreements were discussed and decided upon by them. Some dialogue fragments with integrative utterances are shown in Table 2.6.

2.2.6 Hypotheses

We compared the learning outcomes of pupils who co-constructed textual representations, multimodal representations, and multimodal representations integrated in a timeline. The multimodal representations that pupils constructed, integrated historical concepts, historical phenomena and relations (the components of a chronological-conceptual frame of reference). The considerations above lead to the following hypotheses. First, the pupils who co-construct multimodal representations that combine pictures, diagrams and historical concepts will gain more historical knowledge than pupils who co-construct textual representations, because information will be processed both verbally and visually, and verbal and visual information will be integrated. Second, pupils who integrate different multimodal representations into an overall representation (a timeline) will gain more historical knowledge than pupils who co-construct textual representations and pupils who co-construct multimodal representations without integration, because in this condition temporal relations are also visualised. Third, these differences will endure over a longer time span.

2.3 Results

2.3.1 Tests

We present the results of our analyses of pupils' performance in the three conditions. Table 2.7 shows the mean scores and standard deviations on the pre-test, post-test and retention test in the three conditions. Pre-test data were examined to identify initial differences in prior knowledge scores between the conditions. An ANOVA indicated that the three conditions did not differ significantly from each other in their mean pre-test scores ($F(2, 82) = .32$, $p = .72$). Although the pre-test turned out to have problems of reliability, the three conditions did not differ significantly on pre-test results. There is no obvious reason to suspect that the groups differed at pre-test.

Table 2.7. Means and standard deviations for test scores in the three conditions.

	Text			**Visual**			**Timeline**		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
Pre-test	30	12.13	2.56	31	11.55	2.71	24	11.79	3.32
Post-test	30	15.83	4.08	31	17.58	4.40	24	19.21	5.32
Retention test	30	14.87	4.54	29	15.17	4.44	22	16.09	5.49

Before comparing the test results of the three conditions, we investigated a preliminary question: Did the participants actually learn something from the tasks? For each condition, we checked this with paired samples t-tests between pre-test, post-test and retention test. The paired samples t-tests showed that in each condition the pupils' score increased significantly from pre-test to post-test ($p = .00$) and from pre-test to retention test (p = .00). There was a significant decrease from post-test to retention test for both the Visual ($t(31)$ = 4.34, p = .00 (one-tailed)) and the Timeline condition ($t(21)$ = 3.23, p = .00 (one-tailed)), but not for the Text condition ($t(29)$ = 1.79, p = .08 (one-tailed)).

As we found no significant differences between the conditions on the pre-test, we went on to check for differences between conditions on the post-test and retention test scores. The ANOVA for the post-test showed a significant difference between the three conditions ($F(2, 82)$ = 3.66, p = .03). Post-hoc analysis (Bonferroni) showed that this is attributable to the significant difference between the Timeline and the Text condition, with a mean difference of 3.38 (p = .03). The Timeline group performed significantly better on the post-test than the Text group. On average, the scores of the Visual group on the post-test were 1.75 higher than the Text group scores, but this difference was not significant (p = .58). We did not find significant differences between conditions for the retention test scores, which means there is no difference in long-term effects ($F(2, 78)$ = .43, p = .65).

We suspected that the difference in post-test scores between the Text and Timeline groups might be attributable mainly to questions related to temporal relations – that entailed such questions as placing events on a timeline and indicating the correct chronological order of events – so we did an exploratory analysis for the section of the test with the 11 time-related questions (marking an event on a timeline or ordering events and phenomena chronologically). After all, there were no significant differences for the Visual group, and that leaves the built-in emphasis on temporal relations as the main distinguishing factor between the Text and Timeline tasks. The ANOVA results for the time-related questions in the post-test – taken both from part B and part C – again showed a significant difference between conditions ($F(2, 82)$ = 4.71, p = .01), and again, post-hoc tests (Bonferroni) showed a significant difference only between the Text and Timeline groups (p = .01) with a mean difference of 1.45. These results suggest that the difference between the Text and Timeline groups might be attributable to a better chronological frame of reference in the latter group. After all, the goal was for the timeline to encourage development of this framework, and these questions were designed to measure this. Again, the difference disappears in the long run ($F(2, 78)$ = .26, p = .77). Analyses of just the non-time-related questions shows no significant differences between conditions (post-test: $F(2, 82)$ = 1.89, p = .16; retention test: $F(2, 78)$ = .46, p = .64).

Although the participating teachers (there was only one teacher per school) had pupils from all conditions, and pupils and dyads were carefully (but not completely randomly) assigned to conditions, we decided to check whether the distribution of conditions over classes could have interfered with the results. We used independent samples t-tests and found that the significances and directions of the significances for the separate classes were the same as for the entire sample.

We found no significant differences, then, between the Visual and the Text conditions, nor between the Visual and Timeline conditions. On the other hand, we did find a significant difference in time-related questions between the Text and Timeline conditions. It seems then that the visual support with pictures and schemas might not be the distinguishing factor we expected, whilst the combination of pictures and schemas with specific visualisation of time does seem to make a difference, at least in the short run.

2.3.2 Products and process

To gain more insight into the learning processes, we first took a closer look at the use of the multimodal representations in the group products to see how the pupils in the Visual and Timeline conditions dealt with the pictures incorporated in them. The products revealed that especially in the first and fourth tasks some dyads came up with interpretations that did not match their intended meaning or the information in the text. In the first task the pupils completed a storyboard about the decline of the Roman Empire by ordering drawings and adding concepts and captions. One of the drawings in this task showed Roman soldiers who walk away from a ruin (see Figure 2.1a). Some of the dyads did not relate this drawing to the departure of Roman soldiers from the provinces back to Italy after the Western-Roman Empire fell. These dyads came up with descriptions such as “the armies revolted”, “wandering of nations”, and “they had built roads and bridges”. One dyad described the drawing that showed the division of the Roman Empire as showing that “the Romans had conquered almost everything they wanted to conquer”. One picture contained a map of the Eastern Roman Empire with its emperor on the right, and an empty space and an empty chair on the left (a representation of the fact that there was no Western Roman Empire anymore, see Figure 2.1b). One dyad associated this picture with trade and wrote the following caption: “There were large distances between countries, so for trade as well”.

Figure 2.1a. Picture from Task 1. Departure of the Roman Army from Western Europe.

Figure 2.1b. Picture from Task 1. The Western Roman Empire has disappeared, leaving an empty emperor’s chair, whilst the emperor of the Eastern Roman Empire is still in position.

In the fourth task it appeared that many pupils found it difficult to colour the spread of Christianity and the spread of Islam on a map. Pupils were asked to first colour the spread of Christianity until 500 AD in map A, and then colour the spread of Christianity from 600 to 1000 AD and the spread of Islam from 600 to 1000 AD in map B. The pupils were provided with a map that combined these changes. Many dyads did not execute the task as it was intended. Some of the dyads simply copied the colours of the given map both in map A and map B – which was the wrong answer – others coloured more than was asked for in the instruction or did not colour the correct parts of the map. Thus, it seems that pupils had difficulties reading a complex map and using it to construct two new maps. We may have over-estimated pupils' map reading and construction skills.

Our premise was that the Timeline condition would do better than the Text condition on the tests because these participants had to integrate the textual and verbal information within the tasks. In addition, they had to integrate the information from the different tasks into a larger whole – a timeline. Table 2.8 shows the results of a closer examination of the number of utterances by task for each of the two dyads, as well as the percentage of integrative utterances in the sections of the dialogues where the participants are working on the timeline itself – as opposed to just one task sheet. The differences between the two dyads might be explained by their pre-test levels: Dyad A consisted of two intermediate scoring pupils, whilst Dyad B consisted of an intermediate and a low scoring pupil.

A closer look at integration of text and representations in Timeline dyads A and B showed that Task 1 – which involved completing a process diagram on the decline of the Roman Empire by sorting pictures and adding text – elicited the most integrative utterances in both dyads (45 and 37, respectively). On the other hand, Task 3 – which involved adding captions to parts of a structure diagram on manorialism – elicited very little integration (12 and 3, respectively), and most or even all (Dyad B) of these integrative utterances are found when the participants are occupied with linking Task 3 to the timeline (11.11% and 9.68% of all content utterances, respectively).

Table 2.8. Total number of utterances by task for each dyad, and the percentage of integrative utterances in the timeline sections of the dialogues.

	Task 1 Storyboard		**Task 2 Causal network**		**Task 3 Image**		**Task 4 Cartogram**	
	Dyad A	**Dyad B**	**Dyad A**	**Dyad B**	**Dyad A**	**Dyad B**	**Dyad A**	**Dyad B**
Total number of utterances	322	194	218	98	158	58	265	187
Procedural utterances	72	50	56	33	41	21	61	49
Social utterances	53	4	53	8	36	6	58	7
Content utterances	197	140	109	57	81	31	146	131
Within content utterances:								
Integration frequency[1]	45	37	28	14	12	3	29	13
Integration in timeline frequency[2]	11	5	0	0	9	3	0	2
Integration %[3]	22.84	26.43	25.69	24.56	14.81	9.68	19.86	9.92
Integration in timeline %[4]	5.58	3.57	.00	.00	11.11	9.68	.00	1.53

[1] Integration frequency: The total number of content utterances that show integration.
[2] Integration in timeline frequency: The total number of content utterances that show integration and refer to the timeline.
[3] Integration %: The percentage of content utterances that show integration.
[4] Integration in timeline %: The percentage of content utterances that show integration and refer to the timeline.

Since we found a significant difference between the Text and Timeline conditions on the post-test for the timeline-related test questions, we expected to find a substantive amount of integration during the part of the task that involved the overall timeline. Table 2.8 also shows the frequency and percentage of integrative utterances in the timeline sections of the dialogues. The numbers differ widely between the tasks. Whilst for Task 3 almost all of the little integration we found concerns the timeline, for Task 4 we found very little timeline integration, and for Task 2 we found no timeline integration utterances at all. Again, Task 1 shows slightly higher numbers than the other three tasks. It is possible that the combination of the type of representation, the types of activities, and the type of timeline activities involved specifically in Task 1 – a process diagram, sorting pictures and colouring and labelling the timeline – encouraged integration more than the representation types and activities of the other three tasks.

2.4 Discussion

The aim of this study was to investigate the effects of the type of constructed representation on the acquisition of a chronological frame of reference. We compared the learning outcomes of pupils who – after reading a text on the Early Middle Ages – worked in one of the following conditions: (1) co-construction of textual representations; (2) co-construction of multimodal representations; and (3) co-construction of multimodal representations integrated in a timeline. It was hypothesised that (a) pupils who co-constructed multimodal representations would gain more historical knowledge than pupils who co-constructed textual representations, (b) pupils who integrated different multimodal representations into an overall timeline would gain more historical knowledge than pupils who co-constructed textual representations and pupils who co-constructed multimodal representations without integration, and (c) these differences would endure over a longer time span.

Our hypotheses were partially confirmed. The pupils who integrated the multimodal representations in a timeline outperformed the pupils in the textual representation condition on the post-test, but not on the retention test. Considering the fact that the pupils in the Timeline condition scored higher particularly on the time-related items of the test (marking an event on a timeline or ordering events and phenomena chronologically), we assume that the timeline representation was instrumental in achieving the higher learning outcomes, because it made more temporal relations explicit. The temporal relations specified in the Timeline condition not only entailed the relations between the events – such as plundering Vikings and migrating peoples – that make up a specific development, but also the temporal relations between these developments – such as the disappearance of the Western Roman Empire and the rise of manorialism. An explanation for the disappearance of this effect in the long run (as measured with the retention test), might be that knowledge about temporal relations is more difficult to retain. Furthermore, a timeline that is studied or constructed may be difficult to remember and to reconstruct from memory in a new task. A suggestion for further research is to examine the influence of different representational formats on long-term learning effects as well as on the learning process.

Contrary to our expectations pupils who co-constructed the multimodal representations (the Visual condition) did not show significantly higher scores on the post-test and retention test than pupils in the textual representation condition. This may suggest that collaborative construction of multimodal representations without integration in a timeline does not help pupils more in developing a ‘big picture’ of historical phenomena (events, structures, and themes) than the construction of textual representations. However, we need to be careful in drawing this conclusion. First, although the multimodal representation group did not show significantly higher post-test scores than the textual representation group, the means point strongly in the direction of our hypothesis. Second, although the pictures were meant to give a more concrete representation of abstract phenomena and abstract relations, a closer look at the group products revealed that some dyads had difficulty with understanding some of the pictures that were used, as well as with reading and producing maps showing historical developments. De Westelinck et al. (2005) found the same problem with the iconic sign system used in their research. A possible drawback of pictures in comparison with texts might be that it is more difficult to predict the kind of associations that are evoked. This also makes it difficult to control construction activities with these representations. Some situations and developments in history are difficult to represent with a low degree of ambiguity in a drawing or diagram and might be better represented by means of a story, film or a computer animation. This is a problem that springs from the unclear semiotics of the domain of history. Even when the context of a picture is known, for example the Early Middle Ages, pictures are often open to multiple interpretations. Thus, misinterpretation of the pictures may have resulted in inaccurate or inappropriate mental representations. The natural ambiguity of the domain also makes it hard for students to learn to understand the forms of representation – the importance of which is underlined by Ainsworth (2006): there are no fixed formats in history. Further research can shed light on the associations that are evoked by different types of pictures.

When we look at Mayer’s seven principles for multimedia design (2001), we find that only one is partially confirmed by our data: the multimedia principle. This principle states that a combination of text and visual representations results in higher learning outcomes than just text. Simply a combination of text and visual representations did not have the effect predicted by Mayer’s principle, but we did find that a combination of text and visual representations integrated in a wider overview had more positive effects than just text. However, the circumstances of Mayer’s experiments have so far been quite different from those in our research: (a) his experiments were very short, whereas ours lasted four sessions, (b) his participants were college students, whereas ours were pupils in pre-vocational secondary education, (c) his experiments took place mainly in lab settings, whereas ours were set in normal classrooms, embedded in the curriculum, and (d) the topics of his tasks were very physical in nature, whereas ours were quite abstract and more difficult to represent, and (e) his tasks involved presented representations, whereas ours required construction activities by the learners. The principles of CTML might not work in quite the same way under these different circumstances, as was the case in the study by De Westelinck et al. (2005).

The dialogue analyses seem to support the idea that learning with multimodal representations is to a large extent mediated through active relating of information from different representations (in this study: textual and visual information). The discussions by the two dyads – both of which have learned relatively a lot from the tasks – show quite a few utterances that reflect active integration of textual and visual information. However, the analyses also showed large differences between tasks. Although all four tasks were designed to encourage the pupils to relate and integrate the information from the text and the visual representations, it seems that Task 1 (completing a storyboard about the fall of the Roman Empire) elicited much active integration, whilst Task 3 (describing an image representing manorialism) elicited little integrative utterance. Possibly, Task 1 stimulated active integration more strongly because the integration was more scaffolded: first the pictures are discussed using historical concepts from the text, and then the pictures are put in the correct order. Both activities require that textual information (historical concepts, causal and temporal relations) are related to visual information. In Task 3, pupils answered questions about elements of an image, to which the correct answers were concepts from the text. Thus, it was perhaps not strictly necessary to connect the text to the image, and connecting it to the questions was sufficient, even though the questions did refer to the image. Out of the four tasks offered, Task 3 required the least varied activity from the learners, and fewer relations had to be drawn within the representation than in the other three tasks. It is possible that this task was too easy, that too much information was given away by the guiding questions, or that the picture was not really needed to find the answers to these questions. Future research should focus on the extent to which pupils working on multiple representations connect the information from the different representations – in different task types, representation types, and domains – and how the occurrence or non-occurrence of such connecting activities is related to learning outcomes.

The multimodal representations that pupils worked on in this study combined visualisations of historical phenomena (e.g., the split of the Roman Empire, trade by barter) and relations (causal, temporal). We cannot draw conclusions about the effects of such a combination from the results of this study. A follow-up study – described in Chapter 4 of this thesis – investigates whether representations that combine visualisation of both phenomena and relations can contribute more to the acquisition of historical knowledge than visualisation of only the phenomena or the relations. In conclusion, there are indications in our study that integration and visualisation of historical phenomena, concepts, and causal and temporal relations helps pupils to acquire knowledge of an era.

3. Multimodal representations in collaborative history tasks: A look at student dialogues

Multimodal representations are representations containing a combination of text and schemas and/or pictures. According to the Cognitive Theory of Multimedia Learning such representations can be powerful learning tools. The study described here approaches this theory from the domain of history in co-construction tasks. In an experimental study, the dialogues of pupils who co-constructed either textual representations or multimodal representations integrated in a timeline were compared. The participants were 12 to 14-year-old pupils in pre-vocational secondary education who worked in dyads on a series of four history tasks. Dialogue protocols of the taped student conversations were analysed. The results show that integrated multimodal representations do – to some extent – lead to more discussion about domain content as well as about procedural issues than working with textual representations and resulted in higher scores on a post-test. Furthermore, a comparison between the dialogues for different integrated multimodal tasks showed that the task involving completion of a storyboard or process diagram elicited a higher proportion of content utterances, use of more historical concepts, and a higher proportion of content utterances with reference to visual elements than the other three tasks. In addition, the results show that reference to schemas and/or pictures as well as using more domain-specific concepts correlate positively with learning outcomes.

3.1 Introduction

Multimodal representations combine two or more modes of representation, for example, a verbal representation with one or more types of visualisation (schematic and/or depictive). While research has been done on the effects of learning with multimodal representations – in particular visualisations presented alongside text – there has been less research on the learning processes that occur when using multimodal representations in tasks as reflected, for example, in content analysis of discourse between learners using these representations. This study sheds light on differences in the learning process between multimodal and verbal tasks at the level of task content, as well as on differences in the learning process when using different types of multimodal tasks.

There is a large body of research on learning with multimodal representations (Ainsworth, 1999; Larkin & Simon, 1987; Mayer & Chandler, 2001; Schnotz, 1993). In these studies multimodal representations are seen as cognitive thinking tools – tools for remembering, thinking and problem solving (Jonassen, Reeves, Hong, Harvey, & Peters, 1997) – and the focus is primarily on the learning outcomes that these representations bring about. These learning outcomes are explained by Dual Coding Theory (Paivio, 1991), Cognitive Theory of Multimedia Learning (Mayer, 2001) and/or Cognitive Load Theory (Kirschner, 2002; Sweller, Van Merriënboer & Paas, 1998). In recent research, two topics related to learning with multimodal representations have gained extra attention, namely: 1) the effects of constructing different forms of multimodal representations on learning processes and outcomes, and 2) the learning and interaction processes that are provoked and supported by the construction of multimodal representations by the learners themselves. We will first elaborate on these topics below.

Cox (1999) argues that constructing external representations can be beneficial for learners. Several studies have shown positive learning effects for one specific type of task in which a multimodal representation is constructed, namely a concept map or knowledge map (e.g., Horton, McConney, Gallo, Woods, Senn, & Hamelin, 1993; O'Donnell, Dansereau, & Hall, 2002; Van Boxtel, Van der Linden, & Kanselaar, 2000). Concept maps are defined as diagrams indicating interrelationships between concepts and representing conceptual frameworks within a specific domain of knowledge (Novak, 1990). Figure 3.1 shows an example of a concept map. In order to construct a concept map, students have to think about and describe both concepts that are instrumental to a specific situation and the relationships between those concepts. By doing this, concept mapping helps students focus on the macrostructure of the content, stimulates elaboration and can provide multiple retrieval paths for accessing knowledge.

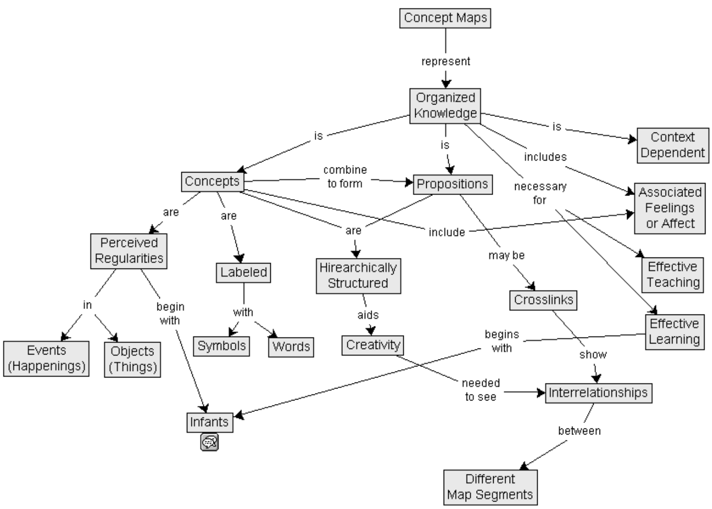


Figure 3.1. Example of a concept map (from http://en.wikipedia.org/wiki/Image:Conceptmap.gif).

The higher level of activity required from students in construction of representations leads to different learning processes. According to Bodemer, Ploetzner, Bruchmüller, and Häcker (2005), students often remain rather passive when they only have to look at a multimodal representation. The researchers argue that active integration of textual and pictorial information by the students themselves is essential for successful learning. They conducted an experimental study in which the students had to relate textual and pictorial information about the working of a tire pump by actively dragging and dropping captions into a drawing of a tire pump on a computer screen. This active integration significantly improved learning. Van Meter and Garner (2005) give examples of tasks in which students construct visual representations, such as organizing or sequencing, given component pieces of a representation.

These ideas about the level of activity and active integration of representations led to an experimental study that was conducted to assess the value of active construction of multimodal representations of historical phenomena (see Chapter 2). Students constructed or completed multimodal representations for different historical phenomena and developments in the Early Middle Ages along with a coordinating timeline. Then, the learning outcomes of pupils who co-constructed textual representations were compared with learning outcomes of pupils who co-constructed multimodal representations integrated in a timeline. Results showed that working on multimodal representations integrated in a timeline led to higher learning outcomes than co-constructing textual representations. The differences in the content of the discourse of collaborating students in these two conditions were assumed to have contributed to this outcome.

Studies on the construction of multimodal representations in collaborative settings that include discourse analysis give some insights into the learning *processes* with multimodal tasks (Bodemer, Ploetzner, Bruchmüller, & Häcker, 2005; Suthers & Hundhausen, 2003; Van Boxtel, Van der Linden, Roelofs, & Erkens, 2002). These studies showed that multimodal representations can facilitate discourse – or articulation – at the content level, thus encouraging the use of domain-specific language, and ultimately knowledge building. Discourse analysis in these studies has focused on the communicative functions (Suthers & Hundhausen, 2003), on the activities in the software (Bodemer et al., 2005), on procedures and task management (Erkens, Jaspers, Prangsma, & Kanselaar, 2005), and on topic content (Fischer, Bruhn, Gräsel, & Mandl, 2002; Roth & Roychoudhury, 1994; Van Boxtel, Van der Linden, & Kanselaar, 2000; Van Drie, Van Boxtel, Jaspers, & Kanselaar, 2005).

In addition to construction, a second topic that has gained attention in research on multimodal representations is the working of different types of multimodal representations. In multimodal representations, verbal representations can be combined with schematic representations (e.g., concept maps), depictive representations (e.g., realistic drawings) or both (e.g., concept maps that include drawings). Particular kinds of multimodal representations fit particular kinds of content or domains. Schnotz and Bannert (2003) pointed out that the form of the visualisation affects the structure of the mental representations. Lohse, Biolsi, Walker, and Rueter (1994) developed a taxonomy of 11 different categories of representations. Each representation type has its own strengths and weaknesses in representing specific types of information. It remains unclear, however, how the different forms affect the learning processes. An earlier study on history learning tasks (see Chapter 2) analysed the discourse of two dyads that each constructed a process diagram in the form of a storyboard about the decline of the Roman Empire, a network chart incorporating pictures showing the effects of the fall of the Western Roman Empire, a structure diagram where learners labelled an image about manorialism, and cartograms showing the spread of Christianity and Islam. The study focused on the use of historical concepts and content utterances about historical phenomena and relations. Concepts serve as building blocks to describe phenomena and relations vital in learning and understanding history (Carretero, López-Manjón, & Jacott, 1997). Both temporal and causal relations are important in understanding complex historical phenomena. Furthermore, Chapter 2 of this thesis looked at the amount of integration of textual and pictorial information. Preliminary exploratory analyses of pupil discourse suggested substantial differences in dialogue quality between tasks, even though all four tasks were carefully designed according to the principles of the Cognitive Theory of Multimedia Learning, to elicit productive interaction and meaningful learning.

3.1.1 Problem definition

This study examines whether collaborative construction and completion of multimodal representations can encourage students to describe and explain historical phenomena using historical concepts. It looks at student dialogues and concentrates primarily on what actually occurs at the content-related level, as this is where pupils most obviously work on acquiring the building blocks of a domain. Through verbalisation of the thought processes, light can be shed on the processes involved in learning with multimodal representations.

The main question that this study tries to answer is: What are the effects of the collaborative completion and construction of integrated multimodal representations versus textual representations on the content of the student dialogue? The study compares the learning processes of pupils who co-constructed either textual representations or multimodal representations integrated in a timeline. The multimodal representations used joined historical concepts, phenomena and relations (i.e., the components of a chronological-conceptual frame of reference).

On the basis of the literature one would expect the collaborative dialogues to be different for different types of tasks (i.e., textual versus integrated multimodal) in several ways. The pupils who co-construct integrated multimodal representations that combine pictures, diagrams and text in a timeline should elaborate more on the (domain-specific) content of the tasks than pupils who co-construct textual representations, because the information in the visual elements needs to be verbalised and then related to the verbal information given to them (Teasley & Roschelle, 1993). In turn, verbalising knowledge and information about abstract historical phenomena and the relations between them should be easier when learners can make use of multimodal representations because it gives learners a common referent: They can point out parts of the representation they are constructing to indicate what they are referring to, and the task product provides a joint workspace that visualises what knowledge has already been co-constructed.

In addition to the main question, this study explores two related questions, one about differences between different types of multimodal tasks, and another about the relation between dialogue and learning outcomes for pupils working with multimodal representations.

With respect to the former, this study explores whether different multimodal tasks have different effects on the content of the student dialogue. In other words: Are there differences in the content of the student dialogue between collaborative completion and construction of four different types of multimodal representations, namely structure diagram, network chart, process diagram, or cartogram? Although differences between multimodal tasks may also depend on the task topics, the type of multimodal representation or the type or number of activities involved in the task may also play a role here. Aspects of the task texts will also be taken into account when discussing any differences between tasks. Thus, analysing the dialogue content should give insight into the workings of different types of multimodal representation tasks.

With respect to the latter, this study explores the correlation between talking about domain specific concepts and phenomena and learning outcomes for students working with multimodal representations. This relation should be a positive one, where more frequent talk about task content, including the use of concepts, should be accompanied by higher learning outcomes. Recent studies found a positive relation between talking about task content and learning outcomes (Fall, Webb, & Chudowsky, 2000; Van Boxtel, 2000).

3.2 Method[2]

3.2.1 Participants

The participants in this study (*N* = 143; aged 12 to 14) were pupils from six different first-year classes in three different pre-vocational secondary schools (VMBO), with one history teacher for each school. The majority of Dutch pupils in secondary school (some 60%) attend this school type. The language proficiency of these pupils is relatively low. History as a school subject is part of the compulsory curriculum for only the first two years for these pupils, so there is little time for developing a chronological frame of reference.

3.2.2 Experimental tasks

Working in pairs, participants carried out a series of tasks on the Early Middle Ages during three consecutive history lessons. The period of 500 to 1000 AD in Western European history – the Early Middle Ages – was selected for several reasons. First, it includes all the different types of phenomena that are dealt within the history curriculum (i.e., the events, structures and themes of an era). Also, the period is representative for the difficulty level of, and the types of relations in other periods in the history curriculum. In addition, the existing curriculum for this period includes different types of developments and social structures that are closely related. It marks a turning point in Western European history with the decline of the Roman Empire and the subsequent development of manorialism, and these two developments can only be understood in relation to each other. A number of important socioeconomic, political and religious changes took place during the Early Middle Ages. The period includes some very abstract concepts – such as ‘manorialism’ and ‘empire’ – and life during this period was in many ways very different from the lives of the participants in this study. On top of that, there is very little original visual material available from the Early Middle Ages that can help pupils in shaping a notion of this period. The specific task content was chosen on the basis of the report of the Committee of History and Social Sciences (2001) that proposes 10 eras with their specific aspects for the history curriculum in Dutch secondary schools. In this proposal, the Early Middle Ages are called “the time of monks and knights” (p. 20), and its specific aspects are the spread of Christianity, manorialism, and the rise and spread of Islam.

The experimental tasks were based on prior experiences with a pilot study. The pilot study was done with 22 participants to determine how much time was needed, and to try out earlier versions of the tests and the tasks. In short, there were three different experimental conditions: (1) Text, (2) Visual, and (3) Timeline. The current chapter deals only with analyses of the collaborative learning processes in the Text and Timeline conditions, because previous analysis of the learning outcomes showed no

[2] The data used for the analyses described in this chapter were drawn from the same experiment as the one described in Chapter 2. Hence, there is considerable overlap between the Method sections of Chapter 2 and Chapter 3.

significant differences between the Visual condition and both the Text and Timeline conditions, whereas there was a significant difference between the Text and Timeline conditions (see Chapter 2).

Pupils in all conditions worked in dyads and started out by reading the same accompanying texts. These were two to three pages in length for each task (of which there were four, see further), including appropriate illustrations. These illustrations were not actively used in the construction task that followed, but were provided to improve basic understanding of the text for all conditions. Important concepts were printed in a bold typeface, and participants were encouraged to use these concepts in their answers. The texts were available to the students throughout the rest of the experiment.

The participants in the Text condition were given tasks to produce just that: Text. Considering the learner level, the textual tasks did not consist solely of open questions, but also included fill-in-the-blanks and ordering sentences, in the same way that their history workbooks do. The tasks in the Timeline condition involved text, diagrams and pictures, and the separate task products had to be integrated in a timeline. There were four different tasks, each with its own specific content and matching representation form. The tasks used in the Timeline condition were designed according to Mayer's (2003) principles for multimedia learning. According to these principles students learn more when words and pictures are combined (multimedia effect), when extraneous material is excluded (coherence), and when words are placed near a corresponding picture (spatial contiguity). The representations for the four tasks reflected the differences in the types of task content, and can be categorised according to the taxonomy by Lohse et al. (1994) as: (1) process diagram (decline of the Roman Empire), (2) network chart (effects of the fall of the Roman Empire), (3) structure diagram (manorialism), and (4) cartogram (spread of Christianity and Islam). Figures A.1, A.2, A.3, and A.4 in Appendix A show examples of completed answer sheets for all four tasks in the Timeline condition.

All pictures used for the answer sheets were made specifically for the experiments by a professional illustrator. The picture for the third task was full-colour, while the other tasks were black-and-white. The task products in the Timeline condition were linked to each other in a large timeline the size of two sheets of flip chart paper (approximately 60x140 cm). The tasks in the Text condition covered the same content, but only textual answers were required, and no pictures were provided on the answer and instruction sheets. An overview of the tasks is shown in Table 3.1.

The tasks were concluded with a summary question, such as: "Finish the following sentence with two causes and use your answers above and the text: The Roman Empire disappeared because … and because … ". The summary question was accidentally left out in Task 4 of the Text condition.

An inventory was made of differences between characteristics of the four tasks in terms of length, possible complexity, and number of activities. This overview is shown in Table 3.2.

Table 3.1. Contents and activities of the collaborative tasks in the condition with textual representations (Text) and the condition with multimodal representations integrated in a timeline (Timeline).

	Activities in condition	
Tasks: Knowledge type and contents	**Text group activities**	**Timeline group activities**
Task 1 Development: The decline of the Roman Empire	Fill-in-the-blank Ordering sentences Summary question	Completing a storyboard: - Ordering pictures - Adding concepts and captions Summary question Adding text and colour to timeline
Task 2 Causal relations: The effects of the fall of the Western Roman Empire	Fill-in-the-blank Ordering sentences Summary question	Completing a causal network: - Choosing pictures - Adding concepts and captions Adding arrows to timeline Summary question
Task 3 Structure: Manorialism	Answering questions Summary question	Describing a historical image: - Answering questions linked to details in the picture Summary question Adding arrows, text and colour to timeline
Task 4 Development: Religion: spread of Christianity and Islam	Fill-in-the-blank	Completing two cartograms: - Fill-in-the-blank in tags linked to specific points on the maps - Colouring the spread of religions in the maps Summary question Adding text and colour to timeline

Table 3.2. Characteristics of the four texts accompanying the four tasks.

	Task 1	**Task 2**	**Task 3**	**Task 4**
Maximum no. of concepts	12	17	11	17
No. of activities	25	22	15	40
No. of words	403	293	333	614
No. of sentences	41	30	30	57
No. of paragraphs	20	10	10	19

3.2.3 Tests

All participants were given the same individual knowledge test three times: a pre-test, a post-test, and a retention test. A retention test was included to investigate long-term effects. The test consisted of three parts (a total of 27 items; see Appendix B) that were administered separately to prevent questions giving away each other's answers. To ensure coverage of the learning goals, the test was based on a test matrix, which in turn was based on a matrix of the learning goals for each task. The test started with a free association 'spider' on the Early Middle Ages in which pupils were asked to write down everything they know about the period in a concept map (see Appendix B Part A). The second part consisted of 8 open items, including both textual and multimodal questions (e.g., with pictures, maps, or timelines; see Appendix B Part B). The third and final part consisted of 18 multiple-choice items (e.g., multiple response questions, choosing which came first in time; see Appendix B Part C). All items dealt with historical concepts, phenomena and relations. In addition, 11 of these items included chronological aspects: 3 open items, and 8 multiple-choice items.

The three parts were scored on the basis of the task text content. Participants received one point for each association spider answer that appeared on an exhaustive list of concepts taken from the task text. No interrater reliability (Cohen's kappa) was calculated, as the percentage of agreement on 30 randomly chosen spider tests was 100%. The interrater reliability between two raters for the open items on 74 randomly chosen tests from different classes (28% of the total number of scored tests) was .89 (Cohen's kappa). Interrater reliability was not calculated for the multiple-choice items.

Cronbach's alpha was used to determine test reliability. One open item was excluded from further analyses, because it had zero variance in the pre-test, and little variance in the post-test and retention test. This item asked for the Arab Empire, while most participants seemed unable to disconnect the concept empire from the Romans. The pre-test shows low homogeneity (Cronbach's alpha = .31). However, this does not seem to show unreliability, but rather that the topic was completely unknown to the participants: Prior knowledge was low to non-existent, and thus the participants had to resort to guessing on the pre-test. Thus, we can reliably speak of a learning situation. Cronbach's alphas for the post-test (.58) and retention test (.66) were not high, but were acceptable. The participants did not receive feedback on their tests.

3.2.4 Setting and procedure

An experiment was conducted with dyads working in one of three conditions, two of which are compared and discussed in the study: textual representations (Text); and multimodal representations integrated in a timeline (Timeline). Scores for the multiple-choice section of the pre-test were used to divide the participants into three groups with low, intermediate or high pre-test scores. Participants from the group with the lowest scores and participants from the group with the highest scores were teamed with participants from the group with intermediate scores, resulting in teams that were composed mainly of low and intermediate level pupils, intermediate pupils only, or intermediate and high level pupils. Thus, no strongly asymmetrical dyads – combining low and high level pupils – were formed, but only dyads with contiguous ability ranges and intermediate dyads, because these combinations have been shown to result in an optimum balance between symmetry to enhance communication and relations and asymmetry to keep the dialogue going (Saleh, Lazonder, & De Jong, 2005). Due to absence of pupils during the first task lesson, some preassembled dyads had to be reshuffled. As a result, there is one low-level dyad (Timeline) and one high-level dyad (Text) in the final sample used for the process analyses. All but two dyads (Text) consisted of either only boys or only girls, because research has shown that mixed-gender pairs may result in inequality interaction and cooperation more than same-gender pairs (Wilkinson & Fung, 2002).

Dyads were randomly assigned to the conditions within each school, so that each condition was present in every school. In four out of six classes, the conditions were put in different classrooms with low/intermediate and high/intermediate scoring dyads evenly distributed over the conditions within a class.

Due to a number of causes a large portion of the original sample had to be discarded: (32% Text, and 37% Timeline). Participants or dyads were discarded if they missed the pre-test or the post-test, or if one or more tasks were unfinished. Participants were included in the final sample if they missed the retention test, or if their dyad partner was absent during one lesson. The high proportion of discarded dyads was mainly due to the lack of concentration in participants who were participating in the Muslim religious festival of Ramadan during the period that the experiment took place: Few of the dyads managed to finish all tasks. As a result, there is less diversity in the ethnic background of the entire sample (about 9% with a foreign background) than in the general population of Dutch pre-vocational secondary education (24%). There might also be a bias towards better attending, academically more able students. However, there is no reason to assume that this effect might have been unevenly spread across conditions. The final sample includes 54 participants (Text = 30; Timeline = 24).

The pre-tests were administered four or seven days before the start of the experiment. The total time taken by participants varied from 30 to 40 minutes. Participants were given brief instructions about the research study and the tasks.

The experiment was started on the first history lesson after administering the pre-test. The classes in two of the three schools were divided over two classrooms by condition. One of the researchers monitored one classroom, and an assistant monitored the other, while the teacher alternated between the classrooms. Apart from a short introduction before splitting up in conditions, there was no classroom instruction. The participants only received feedback on completion ("Is it finished? Yes or no?"), but not on content. Pupil discourse was recorded using small portable tape recorders placed on the participants' desks. When a dyad had completed one task, they were given the next. The participants had three lessons (approximately 150 minutes) to finish all four tasks. The post-test was administered one to five days after rounding off the experiment. The retention test was administered 33 to 49 days after the post-test ($M = 37$). Between the post-test and the retention test, regular classes were taught about the Early Middle Ages, but not about the specific topic of the experiment.

3.2.5 Instruments and analyses for student dialogues

Twenty dyads (40 participants) were selected from the final sample for the discourse analyses: 10 from the Text condition, and 10 from the Timeline condition. The dyads were randomly selected from all dyads from the final sample for which a full set of recordings was available.

The analyses of differences between the two conditions focused on a single task (Task 2). This task – about the effects of the fall of the Roman Empire – was chosen because the participants had had one task to get used to working together and to get accustomed to the type of task and setting. In addition to a comparison between the conditions Text versus Timeline for Task 2, the dialogues for all four tasks in the Timeline condition were coded and analysed to explore the differences in dialogues between the four multimodal representation types.

For each of the 20 dyads selected, the dialogues were transcribed, coded and analysed with utterances as the unit of analysis. All tasks consisted of two phases: an initial reading phase and a production phase. The initial reading phase was the same for all conditions. Dyads were given the text and instructed to read it out to each other – a strategy taken from reciprocal teaching (Slater & Horstman, 2002). After reading the text, the dyads were given the task assignment, which was the starting point for the second phase. The dialogue of the first phase was excluded from the analyses.

Table 3.3. Examples for the categories used for coding the student dialogues.

Code	**Examples**
Procedural utterances	"It's your turn now."
	"Do you have sticky tape?"
Social utterances	"Are you going to the party on Friday?"
	"You know who called last night?"
Content utterances by subtype:	
Core content	"There was little trade"
	"When the Romans left, bridges collapsed"
Auxiliary content	"Trade"
	"Viking boats are cool"
Read out content	"What happened here?"
Content: Reference to visual element	"And here is the fighting"
	"There is a fire in the background"
	"These men are walking around with a cow"
Content: Integration text-visual	"I am sure this is the Roman army"
	"In this picture I see the two emperors of the Western and Northern Roman Empire"
	"But this is the Early Middle Ages and Antiquity is this"
	"The bloke with the axe is a Viking"
Read out procedural	"Look at the five pictures"

The protocols were analysed in several steps, taking the utterance as the unit of analysis. First, the utterances were coded for their basic topic: Content, Procedural, Social or Other utterances. Table 3.3 shows some examples of utterances.

The focus of the analyses was on the part of the discourse that dealt with domain-specific content. Content utterances included utterances about historical phenomena and relations, about pictures, or about the answers to be given on the answer sheet. Procedural utterances were made for regulating the collaboration or the behaviour of the partner, or refer to physical characteristics of the task materials or to the spelling of the answer. Social utterances included all utterances by the dyad partners that were irrelevant to the task. The category Other included utterances by others (e.g., classmates, the teacher), and unintelligible utterances. The procedural, social and other utterances were not investigated further, but served as a valuable context for interpreting the content utterances. The next step was to indicate which utterances were passages read out from the text, the instructions, or the answer sheet (coded as Read out procedural or Read out content, depending on their topic). As the Read out content utterances did not originate from the participants, they were not coded further within the Content category. Again, these utterances provided a context for interpreting the Content utterances.

Next, the Content utterances were coded for different types of content: core content and auxiliary content (see Table 3.3 for examples). Content utterances were coded as Core when participants made a statement about a historical phenomenon or an utterance in which they related phenomena to each other. An exhaustive list of phenomena and relations (historical propositions) based on the task text was compiled for each task by two experimenters – one of whom has a degree in History – to ensure consistent coding of this category. Content utterances were coded as Auxiliary when they referred to the content of the text, answer sheet or pictures, but did not include an explicit description or reasoning about the content.

In addition to coding the dialogues utterance-by-utterance, the total number of concepts used within the whole set of content utterances was tabulated, as well as the total number of different concepts. One utterance could contain more than one concept, as shown in some of the examples in Table 3.3: “In this picture I see the two *emperors* of the Western and Northern *Roman Empire*”, and “But this is the *Early Middle Ages* and *Antiquity* is this”. Conjugations of the same stem were counted as one item (e.g., divided Empire-division, protection-protected).

Some extra coding was done for the Timeline condition as input for the comparison between the four multimodal representation types. Within Content, utterances referring to visual elements on the answer sheets of the Timeline condition (e.g., a picture, an arrow, the timeline itself) were coded as well, since it was assumed that these make it easier to talk about historical concepts and phenomena. Next, the Content utterances were coded for integration, which occurred when information from the text was verbalised and explicitly related to the multimodal representations or a part thereof (e.g., diagrams, pictures, the timeline itself) on the answer sheets.

Interrater reliabilities between two coders were calculated for four randomly chosen dialogue protocols (two from each condition, totalling 1060 utterances). These are all acceptable, and are shown in Table 3.4.

Table 3.4. Interrater reliabilities for dialogue coding.

Coding level	**% agreement**	**Cohen’s kappa**
Topic (Social, Procedural, Content, Read out)	82.80	.74
Content (Core, Auxiliary)	86.40	.76
Reference to visual elements	90.70	.76
Integration text-visual	93.46	.70
Concept count	82.40	.78

3.3 Results

3.3.1 Student dialogues: Textual vs. multimodal representations

Task 2 student dialogues in the Text condition were compared to those in the Timeline condition. In the Text condition the mean duration of the analysed dialogues was 19.43 minutes (SD = 8.90) and in the Timeline condition 19.42 (SD = 5.79). The durations did not differ significantly between conditions ($t(17) = .004$, $p = .997$ (two-tailed)). Hence, all results in this section report frequencies of utterance categories.

Differences between Text and Timeline in Task 2 dialogue utterances

Table 3.5 shows the means and standard deviations for the dialogue frequencies in both conditions. On the whole, the Timeline group talked significantly more than the Text group, as shown by an independent samples t-test for the total number of utterances ($t(15.60) = -2.92$, $p = .00$ (one-tailed), mean difference = 75.60). This general difference between the two conditions seems to be attributable to all three main topics – procedural, social, and content utterances – though only procedural utterances ($t(18) = -3.63$, $p = .00$ (one-tailed), mean difference = 36.10) and social utterances ($t(14.18) = -1.92$, $p = .04$ (one-tailed), mean difference = 23.10) show significantly higher frequencies for Timeline condition than for the Text condition. There is no significant difference for the number of content utterances ($t(18) = -1.51$, $p = .07$ (one-tailed)).

Table 3.5. Means and standard deviations for the frequencies of the dialogue variables in Task 2: Comparison between conditions.

		Mean		**SD**	
		Text	**Timeline**	**Text**	**Timeline**
Total number of utterances	**	85.00	160.60	45.05	68.22
Procedural utterances	**	18.80	54.90	13.36	28.44
Social utterances	*	17.20	40.30	18.64	33.12
Content utterances		49.00	65.40	22.16	26.23
Core content		17.20	16.70	9.77	10.22
Auxiliary content	*	25.70	43.80	11.35	24.02
Read out content		6.10	4.90	3.78	2.81
Read out procedural	**	1.00	3.00	1.70	1.70

Text: N = 10 dyads. Timeline: N = 10 dyads.
*: Timeline > Text at $p < .05$; **: Timeline > Text at $p < .01$.
Core content utterances + auxiliary content utterances + read out content = content utterances. Read out procedural utterances are a subset of procedural utterances.

Table 3.6. Excerpt from a dialogue on Task 2 in the Timeline condition.

Speaker	Utterance	Topic	Subtopic
1	"Here, that's finished, Wesley, well done"	Procedural	
2	"Now you do another sentence"	Procedural	
1	"Oh, let me think"	Procedural	
2	"No, write"	Procedural	
1	"What does that say?"	Content	Auxiliary
2	"Just that the Vikings travelled around"	Content	Core
1	"No, the bottom one, that village"	Content	Auxiliary
2	"That's because"	Content	Auxiliary
1	"Um, OK but it needs something after it"	Content	Auxiliary
1	"The Vikings plundered"	Content	Core
1	"With fifteen exclamation marks after it"	Social	
2	"Oh yeah"	Social	
1	"That makes the sentence a bit longer, that's a question mark idiot"	Social	
2	"Like this"	Social	

Taking a closer look at the content utterances shows that the two conditions have elicited roughly the same number of core content utterances, and similar amounts of content were read out from the task text and task sheet. Yet, even though there was no significant difference between the two conditions for the number of content utterances, the Timeline condition shows a significantly higher frequency for auxiliary content utterances ($t(12.83) = -2.15$, $p = .03$ (one-tailed), mean difference = 18.10). This suggests that perhaps the multimodal representations do encourage content talk, but not to name things or refer to them very specifically. Table 3.6 shows a typical example of pupils discussing what text should be written on the task sheet to describe a picture of Viking raids in Task 2 in the Timeline condition. The columns show the alternation between the two pupils in the dyad (Speaker), a translation of the utterances, the main topic (procedural, social utterances or content), and the subtopic (read out, core content, or auxiliary content). A clear example of an auxiliary utterance in this excerpt is "No, the bottom one, the village", in which the pupil refers to a picture without being explicit. The auxiliary utterances in this excerpt, such as 'What does that say' and 'That's because', seem to lead up to more explicit statements (i.e., core content utterances) about historical phenomena described in the task text. The auxiliary utterances often include deictic expressions, such as 'that' and 'it'.

The Timeline group also read out significantly more procedural information from the task sheet and instructions ($t(18) = -2.63$, $p = .01$ (one-tailed), mean difference = 2.00). This is not surprising, as the instructions for the Text condition were more concise.

Concept use in the two conditions
Analysis of the use of historical concepts used within the content-related utterances (core and auxiliary) showed no differences between the Text and Timeline conditions in the total number of concepts used, nor for the number of different concepts. Closer inspection of the data did not reveal differences between conditions in the choice of specific concepts. Table 3.7 shows the results for the concept use in the dialogue.

Table 3.7. Means and standard deviations for the concept scores of the dialogues in Task 2: Comparison between conditions.

	Mean		**SD**	
	Text	**Timeline**	**Text**	**Timeline**
Total number of concepts	27.70	24.00	16.57	10.74
Total number of different concepts	9.00	7.30	2.26	2.36

3.3.2 Student dialogues: Differences between multimodal tasks

In addition to the comparison between the two conditions (Text and Timeline), differences were analysed in dialogue content between the four different multimodal tasks – process diagram (Task 1), network chart (Task 2), structure diagram (Task 3), and cartogram (Task 4) – within the Timeline condition (see Table 3.1 and Table 3.2). In addition to the categories used to compare the Task 2 dialogues of the Text and Timeline conditions, two categories are introduced here that are specific for the Timeline condition: references to visual elements on the task sheets, and integration of the information provided in the texts given with the visual elements on the task sheets.

Percentages were used to describe and compare the data in this section, because the duration of a dialogue in minutes varied greatly from task to task (e.g., 10 to 30 minutes), and because the tasks are compared within one experimental condition (Timeline). The means and standard deviations of the dialogue percentages are shown in Table 3.8 for each of the four tasks.

The different dialogue score percentages were analysed in a number of multivariate analyses of variance with task (Task 1 vs. Task 2 vs. Task 3 vs. Task 4) as a within-subjects factor. The sphericity assumption was not met for all variables, in which case the Huyn-Feldt correction was applied. Post-hoc comparisons were performed using the Bonferroni adjustment for multiple comparisons for those variables that showed significant main effects. Differences between the four tasks are indicated in Table 3.8 by plus sign ('+') and a minus sign ('-').

Table 3.8. Means and standard deviations for the total number of utterances and for the percentages of the dialogue variables: Comparison between the four tasks in the Timeline condition (N = 10).

		Task 1		Task 2		Task 3		Task 4	
		Storyboard		**Causal network**		**Image**		**Cartogram**	
		Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
Total number of utterances	a	296.60 +	134.68	160.60 -	68.22	99.90 -	30.85	139.00	71.68
Procedural utterances %	a	24.04	5.19	35.96	12.85	32.84	8.62	38.21	23.07
Social utterances %		17.73	14.09	21.58	14.81	13.43	9.81	12.14	13.26
Content utterances %		58.22 +	16.11	42.46 -	14.03	53.73	8.59	49.65	22.36
Core content %		18.63 +	10.53	12.52	10.08	10.27 -	9.62	10.44	10.31
Auxiliary content %		35.37	8.66	26.16	10.73	35.02	11.78	32.47	14.75
Read out content %		4.23	2.54	3.78 -	3.27	8.44 +	5.62	6.73	5.56
Reference to visual element %		19.22 +	7.13	12.34	8.39	6.83 -	3.39	9.82 -	7.84
Integration text-visual %		9.02 +	5.30	4.79	5.58	2.38 -	2.54	2.42 -	3.08
Read out procedural %		2.42	1.29	2.01	1.58	2.38	2.42	2.87	3.06

Categories: Core content + auxiliary content + read out content = content utterances. References to visual elements are a subset of core and auxiliary content together, and integration text-visual is a subset of references to visual elements. Read out procedural utterances are a subset of procedural utterances.

All percentages were calculated from the total number of utterances by dyad before calculating the means for each task.

a: Huyn-Feldt-correction for sphericity was applied in the multivariate analysis of variance.

+ and – indicate differences between tasks: Means with a + sign are significantly higher than results with a - sign in the same row. Means with the same sign do not differ significantly from each other. Results without a + or - do not differ significantly from the other tasks on that dialogue variable.

The first question is whether the total number of utterances differs between the four tasks. This was analysed in a multivariate analysis of variance with task (Task 1 vs. Task 2 vs. Task 3 vs. Task 4) as a within-subjects factor. Since the sphericity assumption was not met, the Huyn-Feldt correction was applied. The main effect of task was significant ($F(1.99, 17.94) = 12.64$, $p = .00$, $\eta_p^2 = .58$). Post-hoc comparisons showed that the total number of utterances was significantly higher in Task 1 where the participants made a process diagram than in Task 3 where they made a structure diagram (mean difference = 196.70, $p = .02$), and it was also higher in Task 1 than in Task 2 where a network chart was made (mean difference = 136.00, $p = .00$). Although Task 4 had an even lower total number of utterances ($M = 139.00$) than Task 2 ($M = 160.60$), Task 4 which involved making two cartograms does not differ significantly from Task 1, due to an outlier dyad in Task 4 (mean difference = 157.60; $p = .06$). Still, the general tendency is that Task 1 elicits more utterances in general than each of the other three tasks.

The next step in the dialogue analyses was to investigate whether there are any differences between the percentages of the three main dialogue topics (procedural, social, and content utterances) in the four tasks. The variables were analysed in a multivariate analysis of variance with task (Task 1 vs. Task 2 vs. Task 3 vs. Task 4) as a within-subjects factor. Again, the sphericity assumption was not met for the percentage of procedural utterances, and the Huyn-Feldt correction was applied.

Task 1 elicits a lower proportion of procedural utterances than Task 2 (mean difference = 11.92, $p = .04$), although the overall effect for procedural utterances was not significant ($F(2.45, 22.01) = 2.36$, $p = .11$, $\eta_p^2 = .21$). It is interesting to note that Task 1 – the storyboard task with the process diagram about the decline of the Roman Empire – has the lowest percentage of procedural utterances (though this is not a significant difference) and the highest percentage of content utterances. This means that the high frequency of total utterances for this task is not due to a large proportion of extra procedural talk. In general, the percentage of procedural utterances varies between 24% and 38% for the different tasks. The overall effect for read out procedural utterances was not significant. This subcategory only made up 2.01% to 2.87% of the total number of utterances.

Social talk is the smallest category throughout, with percentages varying between 12% and 22%. There are no significant differences between the tasks for the proportion of social utterances.

The content utterances make up the largest part of the dialogue, with the proportion of content utterances varying between 42% and 58% for all four tasks. There was a significant main effect for the percentage of concept utterances ($F(3, 27) = 4.18$, $p = .02.$, $\eta_p^2 = .32.$). Post-hoc comparisons showed that Task 1 – just like the results for the total frequency of all utterances – has a significantly higher proportion of content utterances than Task 2, with content utterances in Task 1 making up 58% of the dialogue, leaving behind Task 2 with its 42% (mean difference = 15.76, $p = .03$).

However, this pattern of Task 1 eliciting relatively more content utterances than Task 2 is not reflected in significant differences between the percentages of the content subcategories: core, auxiliary or read out content utterances. Again, the variables were analysed in a multivariate analysis of variance with task (Task 1 vs. Task 2 vs. Task 3 vs. Task 4) as a within-subjects factor. The univariate effect of the variable Core content is significant ($F(3, 27) = 7.68$, $p = .00$, $\eta_p^2 = .46$). Here, Task 1 shows a higher percentage of core content utterances than Task 3 with almost 19% of the total number of utterances (Task 1) against just over 10% (Task 3). The univariate effect of the variable Read out content was also significant ($F(3, 27) = 6.59$, $p = .00$, $\eta_p^2 = .42$). Task 3 shows a higher percentage of read out content utterances than Task 2 with 8.44% against just under 4%. There is no significant difference between the tasks for the proportion of auxiliary content utterances, which is the largest of the content category percentages for all four tasks.

As previously stated, two dialogue utterance categories were added to the analysis of the Timeline condition dialogues to determine the extent to which participants used the visual elements (i.e., the schemas, maps, pictures and the timeline itself) in the four multimodal tasks. Reference to visual elements of the task sheets varied from 6% of the total number of utterances in Task 3 to just over 19% in Task 1. The analysis of variance with task (Task 1 vs. Task 2 vs. Task 3 vs. Task 4) as a within-subjects factor showed that the effect of task was significant ($F(3, 27) = 8.14$, $p = .00$, $\eta_p^2 = .48$). Here, Task 1 surpasses both Task 3 (mean difference = 12.39, $p = .00$) and Task 4 (mean difference = 9.40, $p = .03$).

On the whole, there is little integration of information from the task text with visual elements of the task sheet. The percentages vary from about 2% of the total number of utterances in Task 3 and Task 4 to just over 9% in Task 1. The analysis of variance with task (Task 1 vs. Task 2 vs. Task 3 vs. Task 4) as a within-subjects factor showed that the effect of task was significant ($F(3, 27) = 11.93$, $p = .00$, $\eta_p^2 = .57$). Again, Task 1 scores significantly higher than Task 3 (mean difference = 6.65, $p = .01$) and Task 4 (mean difference = 6.60, $p = .00$) on the percentage of utterances where visual references are integrated with information from the text. There is no significant difference between Task 1 and Task 2, which is not surprising, as both of these tasks required that students discuss what they saw in the pictures before choosing or ordering them, whilst this type of integrative instruction – focusing so strongly on discussing and comparing pictures – was not given for Task 3, which did not have different pictures to discuss and compare, whilst Task 4 contained a very different type of image – maps instead of concrete pictures – and these were only discussed near the end of the task whilst colouring the maps. However, if instructions to discuss pictures in detail were the distinguishing factor between Tasks 1 and 2 versus Tasks 3 and 4, why did Task 2 not show a significant difference here with Task 3 and Task 4? Perhaps the added value of Task 1 was that after discussing the pictures, Task 1 required students to order these pictures, whilst Task 2 required them merely to choose four out of eight. Table 3.9 shows an excerpt of pupils discussing the pictures included in Task 1 for the Timeline condition, where pictures had to be discussed and put in the correct chronological order. The columns show the alternation between the two pupils in the dyad (Speaker), a translation of the utterances, the main topic (procedure, social talk or content), and the subtopic (read out, core content, or auxiliary content). The final column indicates whether the utterances refer to visual elements on the task sheet, or even involve integrating such visual elements with information from the task text. In this excerpt pupils first try to describe the pictures, and to connect the pictures with the information from the text. In the second part of the excerpt, the pupils discuss the chronological order of the phenomena represented by the pictures.

Concept use in the four Timeline tasks

One of the research questions dealt with the use of historical concepts in the task dialogues. The means and standard deviations for the concept scores in the dialogues of the four different tasks are shown in Table 3.10. Since the maximum number of different concepts – as used in the texts accompanying each of the tasks – was different for each task (Task 1 = 12; Task 2 = 17; Task 3 = 11; Task 4 = 17) the total number of different concepts was corrected for this maximum by calculating and analysing the proportion of different concepts to the maximum possible number of different concepts. The use of concepts in the different tasks was also analysed in analyses of variance with task (Task 1 vs. Task 2 vs. Task 3 vs. Task 4) as a within-subjects factor. Post-hoc comparisons were performed using the Bonferroni adjustment for multiple comparisons for those variables that showed significant main effects.

The main effect of task was significant for the total number of concepts ($F(3, 27) = 14.41$, $p = .00$, $\eta_p^2 = .62$). Post-hoc comparisons showed that Task 1 elicited more historical concepts than Task 2 (mean difference = 32.90, $p = .00$) and Task 3 (mean difference = 33.20, $p = .00$). Since Task 1 elicited a higher percentage of content utterances than Task 2 in general, and a higher percentage of core content utterances than Task 3 (see Table 3.8), this is not a surprising result.

The total number of different concepts (see Table 3.2) did not differ significantly between the four tasks. It seems unlikely that this is attributable to a ceiling effect, as the means – at least for Task 2 and Task 4 – are not very close to the maximum possible scores.

The main effect of task was significant for the proportion of different concepts to the maximum possible number of different concepts – which is a correction for the total number of different concepts ($F(3, 27) = 8.07$, $p = .00$, $\eta_p^2 = .47$). Post-hoc comparisons showed that this proportion was higher for Task 1 than for Task 2 (mean difference = .35, $p = .00$), and this seems to indicate that Task 1 was more successful than Task 2 in eliciting talk about a wider variety of concepts within the task topic.

Table 3.9. Excerpt from a dialogue on Task 1 in the Timeline condition.

Speaker	Utterance	Topic	Subtopic	Visual/Integration
1	"Task 1. The Roman Empire disappeared, disappears. In this task you will make a storyboard about the departure of the Romans from Western Europe. 1: Look at the five pictures. Discuss what you see in the pictures."	Procedural	Read out	-
2	"Well, um, here you can see some kind of country or something and a um"	Content	Core	Reference to visual element
1	"Emperor? I think a Roman simply"	Content	Core	Integration text-visual
2	"Yes, I think he is upper class, because you can see that thing, that is Caesar"	Content	Core	Reference to visual element
1	"Yes"	Content	Auxiliary	
2	"He has a thing on his head, a wreath on his head"	Content	Auxiliary	Reference to visual element
1	"Look here you can see a um migration of peoples or something, kind, um, because they are going to, walking like this"	Content	Core	Integration text-visual
1	"Well, I don't know, they are walking somewhere, all soldiers"	Content	Auxiliary	Integration text-visual
1	"I think a soldier who is sitting on some kind of stool with a sword"	Content	Core	Reference to visual element
1	"I think that is the soldier who has left"	Content	Core	Reference to visual element
1	"Look, here he is again"	Content	Core	Reference to visual element
2	"Oh wait, I think maybe I know, I think this is um"	Content	Auxiliary	-
1	"Oh wait, look"	Content	Auxiliary	-
2	"He comes after this one and then they also have a country already or something"	Content	Core	Reference to visual element
1	"No, this comes together"	Content	Auxiliary	-
2	"Yes"	Content	Auxiliary	-
1	"Just look at this here, in this they walk like that all together that way, because first, and then this comes in, and then you have this"	Content	Core	Reference to visual element
2	"First in the entire country and there are battles"	Content	Core	-
1	"Yes, you can see that here"	Content	Auxiliary	Reference to visual element
2	"And then he is sitting here all alone with such a small country and so that has disappeared"	Content	Core	Reference to visual element
1	"Yes"	Content	Auxiliary	-
2	"I think it is something like this"	Content	Auxiliary	-

Table 3.10. Means and standard deviations for the concept scores of the dialogues: Comparison between the four tasks in the Timeline condition (N = 10).

	Task 1		**Task 2**		**Task 3**		**Task 4**	
	Storyboard		**Causal network**		**Image**		**Cartogram**	
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
Total number of concepts	56.90+	19.26	24.00-	10.74	23.70-	10.34	39.00	27.19
Total number of different concepts	9.40	1.43	7.30	2.36	7.40	2.17	9.00	3.94
Proportion different concepts to maximum # different concepts	.78+	.12	.43-	.14	.67	.20	.53	.23

+ and – indicate differences between tasks: Means with a + sign are significantly higher than means with a - sign in the same row. Means with the same sign do not differ significantly from each other. Means without a + or - do not differ significantly from the other tasks on that dialogue variable.

Conclusion

Table 3.11 gives a summary of the differences between the student dialogues in the four Timeline tasks.

Table 3.11. Overview of differences in student dialogues between four different multimodal tasks.

Variable	Significant differences
Total number of utterances	Task 1 > Task 2; Task 1 > Task 3
Content utterances %	Task 1 > Task 2
Core content %	Task 1 > Task 3
Read out content %	Task 3 > Task 2
Reference to visual elements %	Task 1 > Task 3; Task 1 > Task 4
Integration text-visual %	Task 1 > Task 3; Task 1 > Task 4
Total number of concepts	Task 1 > Task 2; Task 1 > Task 3
Proportion different concepts to maximum # different concepts	Task 1 > Task 2

On the whole, Task 1 elicits more utterances and also relatively more concepts in those utterances than the other tasks do – in particular Task 2 and Task 3. These differences might be attributable to characteristics of the tasks or the task texts as described in Table 3.1 and Table 3.2 in the Method section of this chapter. Table 3.2 showed that Task 1 and Task 4 both had slightly longer texts than Task 2 and Task 3, whilst Task 3 involved fewer activities – such as ordering, choosing, or describing – than the other three tasks. However, an exploratory analysis showed no significant correlations between the total number of utterances per task and any of the quantitative task characteristics shown in Table 3.2.

The qualitative differences between the tasks (see Table 3.1) show that one of the features that distinguishes Task 1 from the other tasks seems to be that it deals with a process diagram where participants had to determine the chronological order of a set of pictures of events, and then describe the pictures by writing a logical story. Unlike the other tasks, Task 1 involves the activity of ordering pieces of information (i.e., images depicting events on a storyboard). At the same time, it does not seem to involve many more activities (such as ‘discuss’, ‘write’, ‘choose’, or ‘colour’) than Task 2 and Task 4.

3.3.3 Correlating overall Timeline content dialogue with post-test scores

Comparisons were made for the Timeline condition between the dialogue variables and results on the post-test as a whole by calculating Pearson correlations. To this end, for each dyad the utterance frequencies for each of the four Timeline tasks were added to a grand total frequency. These totals are shown in Table 3.12. The frequencies for each dyad were assigned to the individual dyad members to enable correlating with individual learning outcomes. This was deemed more appropriate than calculating the individual learning outcomes means for each dyad, and then assigning these means to the dyads, because the dialogues were witnessed by each of the dyad partners, whilst the tests really were made individually.

Table 3.12. Means and standard deviations for the total number of utterances, the frequencies of the dialogue variables, and the concept scores in all four tasks together in the Timeline condition (*N* = 20 participants), and correlations with post-test scores.

	All Timeline tasks		**Correlations**
	Mean	*SD*	*Pearson r*
Total number of utterances	696.10	209.08	.02
Procedural utterances	205.00	80.90	-.19
Social utterances	141.70	104.33	-.16
Content utterances	349.40	84.10	.43
Core content	87.80	44.75	.36
Auxiliary content	226.00	64.79	.23
Read out content	35.60	20.27	.26
Reference to visual element	89.50	27.72	.48 *
Integration text-visual	34.60	24.53	.44
Read out procedural	15.60	6.75	.12
Total number of concepts	143.60	55.70	.45 *
Total number of different concepts	33.10	5.74	.31

Core content + auxiliary content + read out content = content utterances. References to visual elements are a subset of core and auxiliary content together, and integration text-visual is a subset of references to visual elements. Read out procedural utterances are a subset of procedural utterances.
* significant correlations with post-test scores at $p < .05$.

Significant positive correlations were found between the post-test result ($M = 18.65$, $SD = 5.07$, $N = 20$) and the number of utterances referring to visual elements on the task sheets ($r = .48$, $p = .03$), and between the post-test result and the total number of concepts used ($r = .45$, $p = .05$). This confirms the premise that more discussion of domain-specific concepts and active use of visual elements whilst discussing content are positively related to learning outcomes.

3.4 Discussion

The effects of multimodal representation tasks in a timeline (Timeline) and of textual tasks (Text) in a collaborative setting were examined. It was assumed that working with multimodal representations in a timeline would result in different dialogues and learning outcomes than working with textual representations would. In addition, the study explored whether different types of multimodal tasks resulted in differences in the dialogues.

The first research question investigated the effects of the collaborative completion and construction of integrated multimodal representations versus textual representations on the (domain-specific) content of the student dialogue. The results of the dialogue analysis of one out of four tasks showed that the multimodal group (Timeline) talked more in general, and produced more procedural utterances in particular than the textual group (Text). This is not surprising since the multimodal timeline task required more organisation and coordination compared to the textual task. Whilst the textual version of the task involved fill-in-the-blank sentences, ordering sentences, and answering a summary question, the multimodal timeline version of the task – completing a causal network chart – required choosing relevant pictures and discarding irrelevant ones, adding concepts and captions to the pictures chosen, relating visual and textual information, linking the elements to the previous task on the timeline, as well as answering the summary question. Van Boxtel, Van der Linden, and Kanselaar (1997) found a comparable difference between tasks. They compared a concept mapping task on electricity concepts with a poster task in which students had to use the same concepts to explain the working of an electric flashlight. In their study, the concept mapping task elicited significantly more talk about the concepts than the poster task did. Their explanation – that the poster task elicited more writing and drawing activities – is in line with the suggestion by Bennett and Dunne (1991) that tasks that require more physical activities are more likely to elicit less abstract talk: “for tasks which combine both action with abstract demands, talk related to action continues to dominate” (p. 113). Still, in the study presented here, the tasks with multiple representations asked for more physical activities, such as selecting and gluing pictures, which resulted in more procedural utterances, but not in fewer content utterances – about abstract historical phenomena and the relations between them – in comparison with the less physical tasks with only textual representations.

The Timeline group also talked more about topics not related to the task (social utterances). The nature of the multimodal timeline task – which involved more physical activity than its textual counterpart – might have made working in the classroom more turbulent, making it harder for pupils to concentrate on the task, thus eliciting more distraction and the tendency for more social utterances. This did not, however, detract from content related discussion. Also, pupil concentration may have been influenced by the fact that for many dyads this particular task was spread out over two non-consecutive lessons, and by the fact that these pupils are not used to working in groups for an entire lesson – although these circumstances were the same for the pupils in the textual condition. Restarting their work could require a certain amount of coordination (i.e., catching up on what has intervened) which could be achieved through social talk. The value of social talk in group work has been confirmed by Chen, Lee, Chu, Wang, and Jiang (2005), although it has to be noted that their research was done in a different cultural setting (Asian as opposed to Western European).

On the whole, the results do not strongly support the idea that the extra step of visualisation in the multimodal condition provokes and supports more extensive discussion of core content than in the textual condition. Co-constructing a multimodal representation in a timeline did not seem to invite more core content utterances, nor did it elicit the use of more or more different concepts. However, the results do indicate that auxiliary talk about the content – i.e., talking about content without making clear statements – was encouraged by the multimodal timeline task. It seems that learners often describe pictures in very general terms in auxiliary utterances – including non-specific referents such as ‘this’ and ‘that’ – because there is a tangible common point of reference, so they can afford to be less explicit. From that point of view, the multimodal representations have a deictic role (i.e., pointing to a picture or schema or to a part of it; see Suthers, Girardeau, & Hundhausen, 2003) that does not occur in the textual condition. At the same time, it might also be difficult for learners – considering their educational level – at this level to use abstract concepts to describe concrete pictures, and as a result the learners keep using everyday instead of domain-specific language. One might conclude that while the breadth of the dialogue did not increase, the depth did (Munneke, Andriessen, Kanselaar, & Kirschner, 2007).

The second and more exploratory research question asked whether different multimodal tasks – process diagram, network chart, structure diagram, and cartogram – have different effects on the content of the student dialogue. The results showed that Task 1, which involved completing a storyboard or process diagram, elicited a higher proportion of utterances about the content of the task, in particular about the core content, and more historical concepts were used in this discussion. This task also elicited the most references to visual elements of the task sheet (i.e., pictures, schematic elements and the overall timeline itself).

Although it is tempting to attribute the differences found between the tasks solely to the differences in multimodal representation types, other characteristics of the four tasks might also play a role. The findings are discussed in relation to different task characteristics in the following paragraphs.

One characteristic of Task 1 that seems to distinguish it from the other three is that it required the very concrete physical activity of ordering pictures to make the story work, which can be viewed as parallel to the cognitive activity of determining chronology. Although the Task 2 also involved discussing the contents of a set of pictures, this network chart task did not involve ordering the pictures. The results seem to indicate that Task 1 was more successful than Task 2 in eliciting utterances about a wider variety of concepts within the task topic. In hindsight, this raises the question whether a comparison between the text and Timeline conditions – as was done for Task 2 – would have led to very different results if Task 1 had been chosen for that purpose. This is a legitimate question, but it cannot be answered at this time.

Task 3 involved describing an image by answering questions about elements of the picture. Little physical activity was required here, and the activities required were also required in the other tasks (writing down answers, drawing arrows, and colouring the timeline). Structuring the task with questions might also have made Task 3 too simple while more open instructions might have led to different results. However, considering the educational level of the participants in this study (pre-vocational secondary education), the questions were placed next to the picture because a more open approach – 'describe what you see in the picture and use the text to do so' – was considered too difficult for them and was expected not to be able to elicit the appropriate concepts. Perhaps this caused the task to be too simple instead, and did not sufficiently encourage the participants to use the picture in finding the correct answers because the questions gave too many clues. This is in line with Bodemer, Ploetzner, Bruchmüller, and Häcker (2005) who found that integration was more beneficial for more complex tasks or topics.

Task 4 – completing two cartograms – did require extra physical activity. Two maps had to be coloured in addition to fill-in-the-blanks and colouring the timeline. Observations in the classroom and of the map questions in the post-test and retention test suggest that map reading is not a strong point of learners at this level. This might explain why the cartogram task, in spite of involving the physical activity of colouring, did not show more content utterances of any kind than Task 2 and Task 3. In addition, the first three tasks contained more concrete pictures that perhaps give easier leads for discussion than a map does as a more abstract visualisation,.

The different topic types of the tasks might also have played a role. Unlike the other tasks, Task 1 had a relatively clear narrative structure that had to be reconstructed by using the pictures. Task 2 dealt with several simultaneous effects of a certain event – and relations of cause and effect are often a stumbling block for learners at this educational level (Carretero, Asensio, & Pozo, 1991; Carretero, López-Manjón, & Jacott, 1997). Task 3 involved the visualisation of a rather complicated system (i.e., manorialism). Task 4 dealt with two maps showing two interacting developments (the spread of Christianity and the spread of Islam), which may have been hard to grasp for the participants, in part due to their lack of map reading skills.

Finally, the study explored whether the learning outcomes might be related to the content of the student dialogue. The results showed that referring to visual elements on the task sheets and using more topic-specific concepts were related to higher learning outcomes. This supports the idea that higher learning outcomes of learners working with integrated multimodal representations are mediated by discourse about domain-specific topics as well as about visualisations.

The findings of Bodemer, Ploetzner, Bruchmüller, and Häcker (2005) suggest that integration of visual and textual information is an important mechanism that explains why learners learn from working with multimodal materials. However, the finding that there was no significant correlation between the post-test results and the number of utterances in which textual and visual information were integrated does not support the idea that this integration should also be verbalised explicitly to result in higher learning outcomes. Perhaps integration is still important, whilst it does not necessarily need to be verbalised.

Future research in this area should be done on integrated multimodal representations where the tasks can be ordered randomly. Due to the nature of the domain, this may be harder to achieve in History, but other non-science domains – such as geography or social studies – may well offer suitable topics. Finally, there are still many unanswered questions with regard to the processes involved in acquiring content knowledge through multimodal representation. For example, why does referring to visual elements of the task seem to be related positively with learning outcomes, whilst integrating these visual elements with information from the topic text does not? Moreover, if referring to visual elements and talking about core concepts is so important, then future research should also focus on how we can encourage learners to perform these activities. In addition, future research could focus on the role of deictic properties of multimodal representations in face-to-face collaborative learning.

4. Concrete and abstract visualisations in history learning tasks[3]

History learning requires that students understand historical phenomena, abstract concepts and the relations between them. Students have problems grasping, using and relating complex historical developments and structures. A study was conducted to determine whether tasks with abstract and/or concrete visualisations can facilitate the history learning process. First-year pupils ($N = 104$) in pre-vocational secondary school worked in randomly assigned pairs. After reading a text, the pairs were given a learning task in one of four conditions: Textual, Concrete visualised, Abstract visualised, and Combined. It was hypothesised that the conditions with visualisations would perform better on a post-test and retention test. Post-test and retention test results showed no significant differences. There were some significant differences on the evaluation questionnaire. Combining text and different types of visualisations in learning tasks does not necessarily enhance history learning. Possible explanations given are the characteristics of the text that was used to construct multimodal representations, the ecological setting, the semiotics of the domain of history – that are not defined clearly – and the difficulty of unequivocally visualising historical concepts. Though no significant increase in learning was found, motivation and perceived competence were higher in the condition with concrete visualisations.

4.1 Introduction

Though modern history schoolbooks are often characterised by an impressive number of pictures, little research has been done into the workings and value of visual support within the domain of history. Distinctive features of history as a school subject are historical events, phenomena, the relationships between them, and the concepts that describe them. Students have problems grasping, using, and relating complex historical developments and structures (Carretero, Asensio, & Pozo, 1991). Visualisations in other domains have been shown to have added value for learning compared to purely textual representations. Peeck (1993), for example, discusses several studies that show that presenting appropriate pictures alongside text increases understanding and memorization, and Mayer's (2001) Cognitive Theory of Multimedia Learning argues that visualisations can be powerful learning tools. These positive effects of visualisations are often explained by the Dual Coding Theory (Paivio, 1991), which assumes that information is processed through one of two channels – verbal or visual – and predicts that adding pictures to text will benefit learning in most cases, as pictures can be processed both verbally and visually, resulting in more elaborate encoding and the availability of more retrieval cues to the learner. Both Peeck and Mayer, however, focus mainly on the domain of the natural sciences. The humanities and social sciences

[3] Prangsma, M.E., Van Boxtel, C.A.M., Kanselaar, G., & Kirschner, P.A. (2006). *Concrete and abstract visualisations in history learning tasks*. Manuscript submitted for publication.

– such as history and geography – remain underrepresented in the corpus of research on learning with visual and multimodal representations.

The type of information represented can strongly determine a representation's suitability for achieving its goal. O'Donnell, Dansereau, and Hall (2002) state that more research is needed on the match and mismatch between knowledge maps (e.g., causal or hierarchical schemas) and the macrostructure of the information they represent. It is possible that this can be extended to other types of visualisations. Schnotz and Bannert (2003) found that, for the success of visualisations of time-differences in geography, it is essential that the representation type used fits the knowledge type asked for. For example, in their study on different types of world maps showing time zones they found that it is easier to calculate time differences with a carpet diagram than with a circle diagram. Likewise, Butcher (2006) found that visual representations seem most successful when they are designed to support the specific cognitive processes needed for deep understanding; that is, when there is a match between representation type and knowledge type. Different domains have different needs. A timeline, for instance, will not often be found in geography schoolbooks, but it is very appropriate for history, because it visualises the temporal relations necessary for building a coherent representation of the past.

Visual representation types have been classified in different ways. One common dimension is the classification of visualisations as abstract vs. concrete (e.g., Lohse, Biolsi, Walker, & Rueter, 1994). Concrete and abstract visualisations each have advantages and disadvantages. Concrete visualisations have a strong resemblance to objects in the real world, for example, photographs and realistic drawings. They may also be easier to interpret than abstract visualisations, as they require little understanding of abstract visual conventions. On the other hand, they seem less appropriate for visualising structural and relational information. Concrete visualisations such as drawings also differ in the extent to which they show a realistic and detailed image of a phenomenon.

Original historical visualisations, such as photographs or paintings, can give a clear image of a historical phenomenon. For instance, building styles, specific tools, religious objects, or period costumes can help shape this image. However, the educational value of such concrete visualisations is often limited by problems of source reliability (after all, these visualisations are the creator's interpretation of objects or events), redundancy of information (e.g., decorative elements), the use of period-specific symbols (e.g., a dog in a medieval picture represents loyalty) and students' lack of experience with the general visual language of a particular period (e.g., frescos, icons or romantic paintings) (Husbands, 1996; Sauer, 2003). Realistic drawings made especially for educational purposes might be a good alternative, as redundant information can be left out and complex information can be simplified. Within the domain of the natural sciences, such drawings combined with text seem to be beneficial for learners. Butcher (2006), in a more recent study, investigated students learning about the heart and circulatory system using either a simplified drawing that highlights structural relations, or a more detailed and realistic visualisation. She found that simplified drawings best support factual learning and information integration.

Abstract visualisations show information units in a way that does not resemble tangible objects, but rather focus on certain aspects of the information, often containing visual items whose meaning is based on convention, such as arrows in a flow-chart, or colours used to show altitudes on geographic maps. Some examples of abstract visualisations are causal diagrams – that focus on the causal relations between the components of the information unit – and flow charts – where each item has a conventional meaning (‘file’, ‘defer’, etc.). O’Donnell et al. (2002) describe several advantages of more schematic representations, which they call knowledge maps. These maps can focus attention on the macrostructure of a body of information. Using knowledge maps resulted in higher recall of main issues in comparison with using text, and the maps seem to be especially supportive for students with weak verbal skills. Other studies on the use and construction of concept maps have also shown that abstract visualisations can support learning, both of concepts and of relations (Fischer, Bruhn, Gräsel, & Mandl, 2002; Robinson, Robinson, & Katayama, 1999; Van Boxtel, Van der Linden, Roelofs, & Erkens, 2002; Van Drie, Van Boxtel, Erkens, & Kanselaar, 2005). Understanding causal relations plays an important part in history learning, but these relations are also very complex: There are usually multiple causes for a single event, their importance or presence is often not immediately obvious, and in addition to their direct effect the causes also influence each other. A causal schema that visualises these relations might help create a clearer overview for the learner (Barnes, 2002).

Given the importance of combining the ‘building blocks’ of historical knowledge (i.e., phenomena, concepts, and relations) and the affordances and limitations of concrete and abstract visualisations that ‘fit’ these types of knowledge, the idea occurs that a combination of concrete and abstract visualisations of these different elements (phenomena, concepts, and relations) can support the acquisition of historical knowledge. It is thought that historical phenomena can be better understood if one can form an image of it *and* relate it to other phenomena (Carretero, Jacott, Limón, López-Manjón, & León, 1994; Husbands, 1996; Leinhardt, 1993). Fasulo, Girardet, and Pontecorvo (1998) argue that a picture can only show a snapshot of a series of events. To make such an image meaningful, it needs to be framed by a temporal plot and its context of antecedents and consequences. For example, ‘manorialism’ in the Early Middle Ages can only be fully understood in the context of the fall of the Western Roman Empire. Students need to understand how each phenomenon is related to other phenomena, both chronologically (i.e., temporal relations) and in terms of cause and effect (i.e., causal relations) (Masterman & Rogers, 2002). Although several studies have shed light on the effects of either schematic (abstract) or pictorial (concrete) visualisations on learning, little is known about the effects of combining representations highlighting different aspects of the topic content.

A previous study focused on differences between textual and multimodal tasks (see Chapter 2). The visualisations that students built in that study contained both pictures and schemas at the same time. The pictures that had to be added by the students each visualised a single historical phenomenon (e.g., trade by barter) and the pictures and text together were incorporated in causal schemas and timelines. That earlier study showed that integrating multimodal representations in a timeline led to significantly better learning outcomes than working on a textual task, but only in the short run. Since that study dealt with different combinations of abstract and concrete visualisations, we suspected that differences in abstractness and concreteness of the visualisations might have played a role. However, this could not be distilled from the data gathered. Therefore the study described here was designed to pull apart the three modes of representation – text, concrete pictures and abstract schemas – to try to find out what the effect was of each type separately as well as in combination.

4.1.1 Problem definition

This study was set up to research the differences in learning effects of different types of visualisations. It deals with differences between learning with textual tasks, with concrete (i.e., realistic) pictures added to the textual tasks, with text in an abstract causal map showing the relations between historical phenomena in the task, and with a combination of text and pictures in a causal map. The main question addressed by this research is: Does combining text and different types of visualisations – abstract and concrete – in history tasks enhance learning of historical phenomena and concepts?

4.2 Method

4.2.1 Participants

The participants in this study (*N*=104) were pupils from six different first-year classes in pre-vocational secondary school with three different teachers in two different schools (pupils aged 12 to 14). The majority of Dutch pupils in secondary school (approximately 60%) attend this type of school.

4.2.2 Experimental tasks

Working in pairs during one history lesson (approximately 45 minutes), participants carried out a task on the Early Middle Ages – 500 to 1000 AD in Western European history – and specifically on the effects of the fall of the Roman Empire. Each student pair was provided with a 328-word text (one page A4). To make sure that all participants had read the text before starting on the assignment, the pairs were instructed to read the text out loud to each other. The text was the same for all four conditions, and did not include visualisations. After reading the text, the pairs were given a task in one of four conditions: Textual, Concrete, Abstract, and Combined. The tasks were designed according to Mayer’s (2001) principles for multimedia learning, focusing specifically on the idea that students learn more when words and pictures are combined (i.e., multimedia effect), when extraneous material is excluded (i.e., coherence), and when words are placed near a corresponding picture (i.e., spatial contiguity). Participants were asked to insert appropriate concepts, and thus finish sentences about events, phenomena, and relationships.

In short, the focus of this study is on photorealistic drawings (i.e., concrete visualisations) of historical phenomena and concepts, and on causal maps (i.e., abstract visualisations) that represent relations between phenomena. The tasks in the Textual condition contained just that: Text in the shape of 12 fill-in-the-blank sentences about the main issues in the text. The Abstract condition contained the same fill-in-the-blank sentences, but in the form of a causal schema. The Concrete condition included 8 pictures to illustrate the 12 fill-in-the-blank sentences. Reading from left to right, the sentences in this condition were in the same order as those in the text condition. The tasks in the Combined condition combined all three elements: the fill-in-the-blank sentences and pictures were integrated in the causal schema. In both conditions with pictures, each picture represented all or part of the concepts that had to be filled in by the students. For example, there were pictures for *Vikings*, for *trade by barter*, and for *agriculture*. The black-and-white drawings used for the answer sheets were produced specifically for the experiments by a professional schoolbook illustrator and showed simplified versions of historical concepts. Only textual answers were required for all conditions. The task sheets are shown in Figures C.1 to C.4 in Appendix C.

All participants worked in pairs for two reasons. This was done to encourage active processing through discussion (Erkens, Jaspers, Prangsma, & Kanselaar, 2005; Roschelle, 1992). Also, the tasks were based on the tasks in the previous study discussed in the Introduction, which dealt with dyads because there it enabled us to study the learning process through the student dialogue.

4.2.3 Preparatory assignment

To give all participants the same starting point for the experimental task in the study, and to give them the required background knowledge about the topic, a preparatory assignment was given. Participants – having just rounded off a chapter on Antiquity – were asked to draw pictures to illustrate a text on the Fall of the Roman Empire. The text was divided into three sections. The first section was about the situation in the second century: a large Roman empire divided in provinces, governed by an emperor and with a strong army defending the borders with fortresses and soldiers. The second section included information about the weakening of the empire (due to incompetent emperors, the division of the empire in an eastern and a western part, invasions and overthrowing of the last of the Roman emperors). The third and final section concluded with the situation as it existed in approximately the year 500 AD, when the Western Empire was gone and only the Eastern Empire continued. The participants were then asked to make three drawings – one for each section of the text – that together would give an accurate representation of the Fall of the Roman Empire. The task was piloted in two history classes.

4.2.4 Tests

Participants completed the same individual test three times (pre-test, post-test, and retention test) which required them to indicate whether given statements were right or wrong. Each test consisted of the same 28 multiple-choice items, shown in Appendix D. Since the instructional text was only 328 words in length, it was difficult to construct parallel tests with a sufficient number of items. Together, the 28 items covered the full range of phenomena, relations and concepts in the text and the assignment. To avoid test-effect as much as possible, the order of the questions was reversed for the post-test. In addition, the post-test was preceded by a free recall test, in which participants were asked to write down everything they could remember from the text and the task for the period 500 to 1000 AD. They were encouraged to write full sentences, and when giving loose concepts at least try to explain them.

Prior knowledge was low, so pupils had to resort to guessing on the pre-test, resulting in low homogeneity (Cronbach alpha = .53). Cronbach’s alpha for the post-test, though not high (.68), was acceptable. Cronbach’s alpha for the retention test was .73.

The free recall test was analysed with respect to the number of different historical concepts used and the number of correct propositions stated. First, all were divided into segments. Then the concepts – using a list based on the text and task – were underlined, and the number of different concepts was counted for each participant. Then the segments were coded as proposition or non-proposition. A proposition was defined as a statement about a historical phenomenon, situation, or concept. Finally, the propositions were coded as either correct or incorrect. Interrater reliabilities were calculated for each step for a random selection of 15 post-tests. The interrater reliability (Cohen's kappa) for concept coding was very good (.86), for distinguishing propositions it was good (.76), and for correctness of propositions it was also good (.77). Examples of propositions from the free recall test are shown in Table 4.1.

Table 4.1. Examples of pupil propositions from the free recall test.

Correct propositions	**Incorrect propositions**
Without the Romans it was not safe in the city.	Western Europe was conquered first. *
The Vikings started plundering.	The Roman army fell apart. *
Roads and bridges were not looked after anymore.	People lived from trade. @
Serfs gave part of their crop to the lord in exchange for protection.	Serfs worked for the lord to get part of the crop.

* This happened before 500 AD.

@ Quite the opposite: people resorted to agriculture.

4.2.5 Evaluation questionnaire

We also included questions to evaluate the participants' perception of the task difficulty and enjoyability and of their perception of their own achievement, based on research by others on the effects of these factors on motivation and learning. Perceived competence, for example, has been shown to be an important factor that can influence motivation and success in learning (Bandura, 1993; Boggiano, Main, & Katz, 1988). Concrete visualisations such as the pictures used in our study may motivate students to focus attention on important parts of the text and the accompanying task. The pictures may also reduce the difficulty level of the task, because the pictures represent part of the concepts that have to be filled in by the students. Salomon and Leigh (1984), on the other hand, found that a representational format that is perceived as being easier results in lower mental effort, leading to less deep processing. The questionnaire was given to the participants after completion of the task. This questionnaire consisted of three questions each on a 4-point scale, targeting perceived difficulty level ("Did you find the task easy or difficult?"), enjoyability ("How did you like the task?"), and perceived learning gains ("How much did you learn from the task?").

4.2.6 Setting and procedures

The teachers assigned the pupils to three levels (low, intermediate, and high), based on their history grades on their report cards. These levels were used to distribute pupils over conditions within each class, and then divide them into dyads with contiguous ability ranges (low + intermediate and intermediate + high) and intermediate dyads, because these combinations have been shown to result in an optimum balance between symmetry to enhance communication and relations and asymmetry to keep the dialogue going (Saleh, Lazonder, & De Jong, 2005). Low-intermediate and high-intermediate dyads were evenly distributed over the conditions within each class. Participants with missing data (e.g., due to missed tests) were not included in the final sample. Table 4.2 shows the final distribution of pupils from different teachers over conditions.

Table 4.2. Final sample size: Distribution of participants over conditions and teachers.

	Condition				
	Textual	**Abstract**	**Concrete**	**Combined**	**Total per teacher**
Teacher A	10	7	8	12	37
Teacher B	10	5	9	7	31
Teacher C	10	13	6	7	36
Total per condition	30	25	23	26	104

For each teacher, two classes were included in the sample.

The experiment began in the first lesson after completion of the regular lessons on Antiquity. To ensure that all participants had a similar starting point for the period before 500 AD, the preparatory assignment on the fall of the Roman Empire was administered before the pre-test. The pre-tests were administered one to six days before the start of the experiment. The evaluation questionnaire was given to the participants directly after completion of the task. The post-test was administered directly after the questionnaire: After the free recall test was collected, the multiple choice part of the post-test test was handed out. The retention test was administered about 6 weeks after the experiment. The participants did not receive feedback on their tasks or tests during the entire period of the study. Between the post-test and the retention test, regular classes were taught about the Early Middle Ages, but not about the specific topic of the experiment.

4.2.7 Hypotheses

Significantly higher post-test scores were expected for the Abstract and Concrete conditions than for the Textual condition, because the concrete and abstract visualisations are expected to stimulate more elaborate encoding and can function as anchors for remembering the information. We expected the Combined condition to have significantly higher scores than the other three conditions, because this condition supports the formation of a clear image of historical phenomena, it makes causal relations more salient and it provides the most anchors. Furthermore, we expected that students in both conditions with concrete visualisations would perceive the task as easier and more enjoyable than students in the other conditions.

4.3 Results

There were no significant differences between conditions on the pre-test score (F(3, 100) = .39, p = .76, η_p^2 = .01). Since the data for the post-test were not normally distributed, a Kruskall-Wallis test was used. The analysis showed that there were no significant differences between any of the conditions on the post-test score (χ^2(3) = 2.48, p = .48), nor between conditions on the retention test score (F(3, 100) = .10, p = .96, η_p^2 = .00). Additional analyses showed that the scores increased between the pre-test and post-test for all conditions, as well as between the pre-test and the retention test, meaning that performance of all conditions improved. The difference between the post-test and retention test scores did not show a significant decline in learning results for any of the conditions. The descriptive results are shown in Table 4.3.

Table 4.3. Means and standard deviations for the pre-test, post-test and retention test scores.

		Pre-test		Post-test		Retention test	
Condition	*N*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Textual	30	17.50	3.45	22.73	3.35	23.97	3.18
Abstract	25	16.96	3.59	23.72	3.71	23.48	3.27
Concrete	23	16.57	3.91	22.78	3.00	23.74	3.53
Combined	26	16.58	3.94	23.88	3.34	23.65	3.21
Total	104	16.93	3.68	22.03	3.35	23.72	3.24

The free recall tests were analysed for use of historical concepts and propositions. An ANOVA showed that the four conditions did not differ significantly in the number of different concepts ($F(3, 100) = 1.56$, $p = .20$, $\eta_p^2 = .04$), nor in the number of correct propositions ($F(3, 100) = 2.29$, $p = .08$, $\eta_p^2 = .06$). Table 4.4 shows the mean number of different concepts and the mean number of correct propositions for each condition.

Table 4.4. Means and standard deviations for the number of different concepts and for the number of correct propositions in the free recall test.

		Concepts		Correct propositions	
Condition	*N*	*M*	*SD*	*M*	*SD*
Textual	30	6.67	2.89	5.10	3.39
Abstract	25	6.96	3.19	4.32	2.91
Concrete	23	5.35	2.39	3.74	2.34
Combined	26	6.65	2.64	5.85	3.21

A Kolmogorov-Smirnov test showed that the data of the evaluation questionnaire were not distributed normally. Table 4.5 shows the results of the questionnaire for each condition. On average, the participants in all four conditions agreed that the task was quite enjoyable ($M = 2.11$, $SD = .52$), as a Kruskall-Wallis test showed no significant differences ($\chi^2(3) = 5.75$, $p = .12$). However, the conditions did differ in their evaluation of the difficulty level ($\chi^2(3) = 14.22$, $p = .00$) and their estimation of how much they had learnt ($\chi^2(3) = 8.19$, $p = .04$). A series of Mann-Whitney tests with Bonferroni correction showed that the condition with concrete visualisations rated their task as significantly easier than the Textual condition ($U = 187.00$, $p = .002$) and the condition with abstract visualisations ($U = 148.50$, $p = .001$). At the same time, when judging how much they had learnt, the Concrete condition thought they had learnt more than the Textual condition ($U = 214.00$, $p = .007$) thought they had.

Table 4.5. Means and standard deviations for the three questions in the evaluation questionnaire.

	Enjoyability		Difficulty		Learning gains	
Condition	*M*	*SD*	*M*	*SD*	*M*	*SD*
Textual	2.23	.43	2.00*	.53	2.37*	.67
Abstract	2.08	.57	2.00*	.41	2.16	.55
Concrete	1.91	.53	1.50*	.51	1.86*	.56
Combined	2.16	.55	1.77	.59	2.27	.60

* $p < .05$.

4.4 Discussion

Returning to the research question, the results lead to the conclusion that tasks combining text and different types of visualisations do not necessarily enhance history learning more than textual tasks. The participants' perception that learning is easier and more fruitful with pictures was not confirmed by the test results. A number of factors may have influenced this outcome.

First of all, participants in the Abstract, Concrete and Combined conditions were not explicitly stimulated to use the schema and/or pictures provided or to talk about them. All four conditions received the same carefully written one-page text. It is possible that reading this short text in itself was enough to perform well on the tests, and that this reduced the added value of the tasks – even though the pictures and schemas were designed to support the learning process. Bodemer, Ploetzner, Bruchmüller, and Häcker (2005) found that integration of multimodal representations was effective only for complex information. Possibly, the text used in our research was not difficult enough to require the participants to make full use of the visualisations, the fill-in-the-blanks task provided enough opportunity for processing the information, or the text effectively explained the complexity of the information. The texts contained concrete examples and concepts. High-imagery texts are judged more interesting and result in higher comprehension scores than low-imagery texts (Clark & Paivio, 1991). Also, the test did not contain higher-order questions that asked for knowledge application. It is possible that a higher-order assessment would have given different results.

Another factor is reflected in the saying "Old habits die hard". It is possible that pupils are fixed in their habits, in their approach to doing history and dealing with history tasks. Perhaps pupils typically focus on the textual information since their teachers primarily test them on the content of texts, and – related to the previous point – do not use visual information to their full advantage unless explicitly told to do so – and also how to do it. Such explicit instruction and practice with using visual information was not included in the research design. De Westelinck, Valcke, De Craene, and Kirschner (2005) also suggest that learners may find it hard to learn with multimodal representations when they have limited understanding of the visualizations used. On the other hand, in studies in other domains simply adding pictures to text (without any constructive activity to process the information or relate text and pictures) appeared to be beneficial to learning (Peeck, 1993).

Further explanations for the findings can be placed under three themes: the setting, the nature of the domain, and the semiotics of the visualisations. All three were different in this study than in most studies that did confirm the superiority of multimodal tasks. First, the research was done in a classroom setting, where the tasks were incorporated in the normal history curriculum. This ecologically valid setting may have influenced the results through interfering circumstances, such as classroom dynamics, and attention span differences. While one task in one lesson might be too little to result in differences between conditions it is a type of task that is very common in history education. Such tasks are relatively brief (approximately 20 minutes). The same cannot be said for the tasks and visualisations found, for example, in Mayer's experiments, which often consist of very short animated clips. While this study was not intended as a replication study of Mayer's work it does raise the possibility that there may be a few problems with the validity of some of the assumptions of CTML in a classroom setting.

Another possible explanation for the findings of this study is the nature of phenomena in the subject domain. Previous studies on learning with visualisations by other researchers were mainly done in the domain of science. The nature of phenomena dealt with in science is often very different from the nature of historical phenomena. Many historical phenomena – such as manorialism and serfdom – are not easily unequivocally represented in concrete or even combined visualisations. Whereas in science often one picture can replace a thousand words, in history a thousand pictures is often not enough, and there is a need for more verbal information to interpret a picture or schema. In history learning, just a picture or schema with no or very little text can serve as an anchor, but it does not usually serve as a suitable replacement for just textual information.

Linked to this is the idea that the grammar – or semiotics – of visualisations in the domain of history has no unambiguous principles or rules as there are in science. A domain like science, even though it deals with complex phenomena like waves, has some basic governing principles – such as frequency and wavelength – that remain the same, whether the wave runs through water, air or solid matter. De Westelinck et al. (2005) found limitations of Mayer's theory of multimedia learning (Mayer, 2001) in a study within the domain of educational psychology in longer learning periods and with more abstract and semiotically ambiguous concepts, as the study discussed here did for history. Whilst the science domain offers a structured visual grammar and tangible topics, the semiotics of the domain of history are less clearly defined. Even seemingly simple symbols like arrows can be interpreted in different ways, for example as indicating dynamic relations such as causality, or merely as temporal relations, or as static relations describing the structure of phenomena (O'Donnell et al., 2002). In addition, even the basic governing principles such as cause and effect can be quite unpredictable, and ask for more than just arrows to describe, as a single event can trigger an avalanche of different possible consequences. Even when the context is known, for example the Early Middle Ages, pictures are also often open to multiple interpretations. This also makes it hard for students to learn to understand the form of representation – the importance of which is underlined by Ainsworth (2006): There are no fixed formats in history, for example for visualizing time, war or manorialism.

A bright spot in the results deals with the appreciation of the visualisations. Students in the visualisation condition ('Concrete') appreciated the visualisations, as shown by their rating of the materials as being easier and their feeling that they had learnt more from the task. Such positive apprehension of the materials should not be underestimated. The goal of educational innovation is not only to make learning more *efficient* so that learners learn the same amount of material in a shorter time span, and/or make learning more *effective* so that learners learn more in the same time span, but also to make learning more *enjoyable* such that the affective learning experience is pleasing and learners will want to learn (Kirschner, 2004). Educational research tends to focus only on determining how specific tools, environments, or student characteristics affect either the effectiveness and/or efficiency of learning, and not on how these factors may affect enjoyment. Although concrete visualisations were not found to have an effect on performance in the research reported here, they still had very important positive consequences for the way students worked on and appreciated the tasks. Including concrete visualisations led to a more enjoyable experience for the students. When designing learning materials, concrete visualisations should therefore be taken into account.

Future research should focus on gaining insight into the use and interpretation of different types of visualisations in the humanities and social sciences, for example in domains such as history or geography. In particular, more qualitative studies should shed light on the way pupils understand and interpret different kinds of visualisations of historical phenomena and relations and the conditions under which such visualisations can enhance history learning.

5. General discussion

We cannot – in the 21st century – imagine a school textbook not containing many different types of illustrations. Modern schoolbooks are packed with pictures, tables, graphs, and diagrams – in addition to texts. Visual or multimodal representations are seen as being instrumental in encouraging deep learning. When explaining my research to non-educationalists the most common response is: "Of course pictures make learning easier and better!" However, this opinion was not always predominant, and pictures were not always taken for granted as an ingredient of learning materials. The question, which has been researched since the 1960s, is: Do these pictures, tables, graphs, et cetera really do what we think they do and why? The main topic of this thesis is multimodal representation used in learning. The aim was to determine whether using multimodal representations in history learning tasks makes a difference for the learning outcomes achieved and/or the learning process carried out. Using representations is defined in this thesis as co-constructing representations in small groups.

This chapter is outlined as follows. First, the four research questions mentioned in the introduction to this thesis are answered, followed by an answer to the main research question. Next, methodological issues of the studies described in this thesis are considered. Next, theoretical implications of the findings are discussed. Finally, suggestions for future research are made.

5.1 Answering the research questions

What are the effects of the general type of co-constructed representation – textual, multimodal, or integrated multimodal – on the acquisition of a chronological-conceptual frame of reference?

Based on the literature on learning with multimodal representations (e.g., Ainsworth, 1999; Paivio, 1991; Mayer, 2001) the collaborative construction and completion of multimodal representations should contribute more to history learning than the collaborative completion and construction of textual representations. In addition, multimodal representations may also facilitate the acquisition of a chronological-conceptual frame of reference, which is important in history learning (e.g., Leinhardt, Stainton, & Virji, 1994; Carretero, Jacott, Limón, López-Manjón, & León, 1994). Consequently, integration of multimodal representations in a timeline should strengthen such an effect. The individual learning results presented in Chapter 2 showed that the collaborative completion and construction of multimodal representations did not contribute more to learning than the collaborative completion and construction of textual representations: Students who constructed and completed multimodal representations did not outperform students who constructed textual representations, although the means did point in the direction of the hypotheses. However, the results also showed that working on multimodal representations integrated in a timeline did lead to higher learning short-term outcomes than co-constructing textual representations, although this significant difference disappeared in the long run. These outcomes raised the next question, investigated in Chapter 3:

What are the effects of the collaborative completion and construction of integrated multimodal representations versus textual representations on the content of the student dialogue?

Chapter 2 showed that there were some differences in learning outcomes between the textual and timeline conditions. This raised the question – answered in Chapter 3 – whether the content quality of the learning process might mediate this in the timeline condition. Student dialogues were recorded during data collection for Study 1. The utterances for Task 2 in the Text and Timeline conditions were coded as procedural, social or content utterances. Content utterances were analysed further for level (core vs. auxiliary) and concept use.

The results showed that there were some differences between the two conditions. The multimodal timeline representation task elicited more utterances in general. This was reflected in all main topics (procedural utterances, social utterances and content utterances), although only the former two showed a significant difference. A closer analysis of the content utterances showed that the multimodal timeline elicited more auxiliary content utterances, i.e., it elicited more content utterances, but did this in less direct terms; by referring to content without using full sentences with historical concepts.

What is the effect of the type of multimodal representation – process diagram, network chart, structure diagram, or cartogram – on the content of the student dialogue?

The preliminary exploratory analyses of pupil discourse discussed in Chapter 2 suggested substantial differences in dialogue quality between tasks within the timeline condition, even though all four tasks were designed to elicit productive interaction. This outcome was further explored in Chapter 3. Student dialogues were recorded during data collection for Study 1. The utterances for all four tasks in the Timeline condition were coded as procedural, social or content utterances. Content utterances were analysed further for level (core vs. auxiliary), reference to visual elements, integration of modes, and concept use.

The results showed that there were indeed some interesting differences between the four multimodal representation task types. Task 1 – involving determining the chronological order of pictures and writing the story they tell underneath the pictures – elicited more utterances in general, more content utterances in particular, and more references to visual elements on the task sheet as well as integration of visual elements of the task sheet with information from the task text. In addition, this task elicited more historical concepts and a wider range of the concepts available from the task was used. Possible explanations might be found in differences between Task 1 and the other three tasks. Task 2 – in the shape of a causal network – involved only selecting pictures, not ordering them, and students were not asked to distinguish causes from effects. Task 3 – a structure diagram – contained a single picture, so that discussion of selecting or ordering pictures did not apply to this task, and questions focused mainly on the task text. This task was perhaps too easy and the picture not really necessary to complete the task successfully. Finally, the two cartograms in Task 4 did not contain very concrete visual elements, and pupils seemed to have problems understanding the maps.

The references to visual elements on the task sheet as well as the number of historical concepts used correlated positively with learning outcomes on the post-test, indicating that it is worth encouraging learners to do both of these things.

Further reflection on the studies described in Chapters 2 and 3 led to the idea that perhaps one of the distinguishing factors between different types of multimodal representations is the inclusion, exclusion or combination of abstract (schematic) and concrete (depictive) representations. This led to the following question, investigated in Chapter 4:

What are the effects of combining text and different types of visualisations – abstract and concrete – in collaborative history tasks on learning historical phenomena and concepts?

In a follow-up experiment that dealt with differences in learning results between different types of multimodal representations – abstract, concrete or a combination of abstract and concrete – the short term and long term learning outcomes for these three multimodal conditions were compared to those for a textual condition. As learning about a historical phenomenon requires understanding both abstract and concrete aspects of this phenomenon, concepts, and relations, a combination of abstract and concrete visualisations should have the greatest potential for enhancing learning. Although working on abstract, concrete or a combination of abstract and concrete representations did not result in significant differences in learning outcomes, students who worked with concrete visualisations rated their task as easier and thought that they had learnt more. It seems that combining text and different types of visualisations in learning tasks does not necessarily enhance history learning. Possible explanations are a lack of active use of the visualisations, elements of the real life setting (e.g., ability of pupils to concentrate, absence of whole-class teaching), the way representations are presented and relationships defined (i.e., the semiotics of the representations used) in the domain of history, and the difficulty of multi-interpretable representation of historical phenomena and concepts – that calls for a clear contextual frame of reference.

The central issue in this research is how multimodal representations can support verbalisation and co-construction of meanings and relations, and thus contribute to the attainment of domain specific conceptual knowledge. The studies in this thesis look at learning tasks with multimodal representations, and at integrating different representations in a timeline, and contrast these with learning tasks with textual representations. In other words: How can learning *activities* with *different* multimodal representations *in relation to each other* contribute to meaningful learning? The central question addressed in this thesis was specified as follows:

> *How does making and connecting different types of multimodal representations affect the collaborative learning process and the acquisition of a chronological-conceptual frame of reference in 12 to 14-year-olds enrolled in pre-vocational education?*

Several conclusions can be drawn from these results. First, learning with multimodal representations without integration in a timeline does not result in higher learning outcomes than learning with textual representations – neither in the short term, nor in the long run. Second, learning with multimodal representations integrated in a timeline resulted in higher learning outcomes than learning with textual representations – although the difference disappears in the long run. Third, compared to learning with textual representations, learning with multimodal tasks integrated in a timeline does affect the collaborative learning process by eliciting more interaction in general and more auxiliary content utterances in particular. Also, discussing content by referring to visual elements and using more historical concepts are positively related to short-term learning outcomes. Finally, learning with either concrete or abstract multimodal representations, or with a combination of concrete and abstract multimodal representations does not result in significantly higher learning outcomes than learning with textual representations.

5.2 Methodological issues

5.2.1 Participants

All studies in this thesis involved 12 to 14-year-old participants in the first year of pre-vocational secondary education (VMBO). A majority of Dutch pupils in secondary school (some 60%) attend this level. In general, the language proficiency of pupils attending such schools is lower than that of pupils in general secondary education, in part due to the fact that a considerable proportion of these pupils (about 24%) are from migrant families and do not have Dutch as a mother tongue.

However, because of their relatively low language proficiency, visualisations might be particularly suitable for these learners. This also means that there may be a differential effect for the use of visualisations. Previous research in the field of representations for educational purposes has focused primarily on the undergraduate and college preparatory levels. The existing lack of diversity in the choice of participants was, of course, also an important reason to choose a different target population for the research reported in this thesis. Rather than saying the findings reported here do not necessarily generalise to other learner types, we can now say that perhaps the findings of previous research do – at least to some extent – generalise to lower levels.

In addition, the practical nature of the tasks used in this research seems suitable for this learner level, as pre-vocational secondary education is practice-oriented to suit its pupils' needs. The evaluation questionnaire described in Chapter 4 showed that pupils enjoyed working on the multimodal tasks, and that they felt they had learnt from the tasks. At the same time, the pupils seemed to have some specific problems whilst working on the multimodal representations that might be related to their skills and experience. First, some of the participants in the studies described in Chapters 2 and 3 made some unexpected mistakes in colouring the maps in the post-test, for example by colouring the North Sea as Islamic. This suggests that some of the participants lacked basic map reading skills. It was beyond the scope of this research to investigate this further. However, several studies have detected difficulties in working with and interpreting maps (Postigo & Pozo, 1998; Scevak & Moore, 1998). Further research on cartogram tasks may shed light on the question whether our target group – 12 to 14-year-olds in pre-vocational secondary education – can recognise and interpret maps of Europe. Second, the tasks might have been more effective if they had also been reviewed in a whole-class discussion with the teacher, because these pupils are not used to working independently for three consecutive lessons. Of course, this would probably also have decreased differences between conditions.

There was considerable attrition in the experiment described in Chapter 2 and Chapter 3. Although the experiment started out with 143 pupils, the final sample includes only 85, an attrition rate of 40%. The high proportion of discarded dyads was mainly due to loss of concentration among pupils who were participating in the Muslim religious festival Ramadan during the period the experiment took place. Many of these pupils did not manage to finish all tasks, and several different pupils were absent from each lesson. Due to the random assignment of students to the conditions, no effect of this attrition on the results is expected.

5.2.2 Research design

The experiments described in this thesis were all carried out in ecologically valid settings. Choosing for ecological validity has its pros and cons. The experiments were closer to school learning situations than much previous research on multimodal representations has been, but choosing for an ecologically valid setting also sets some limitations on control and ethics. First, working with pupils and schools in an authentic school setting implies that some things will go wrong, for example because people are absent, rooms are not available, or pupils start walking around. Ethically, it is not possible in a setting like this one to have a proper control group that only does the pre-tests, post-tests and retention tests or one that does not include essential learning materials. The pupils in that condition would have to skip part of the curriculum or deal with the topic much later, at an illogical point in the curriculum.

In addition, the experimenters encountered some technical and practical problems with recording the pupil dialogues and organising sufficient space. First, out of the 85 pupils left in the final sample of the first experiment, only 26 dyads (52 pupils) delivered complete and usable recordings of all four tasks. Pupil discourse was recorded using small portable tape recorders placed on the pupils' desks. Problems encountered with these recorders concerned mainly technical malfunctioning and improper use of the machines by pupils.

The second practical problem was organising sufficient space. The original idea for the experiment described in Chapter 2 and Chapter 3 was to take dyads out of the classroom one by one (two pupils at a time, or four at a time if a research assistant was available) and give them three consecutive hours to finish the tasks. This would have allowed for video recordings instead of only audio tape recordings and thus for the capturing and analysis of full interaction. It would also have been possible to administer the post-test immediately after the tasks for each dyad, instead of one to several days after the task lessons. However, with three consecutive lessons for each dyad (out of eight for a normal school day), about 24 pupils in each class, two classes per school, one research assistant and two video cameras it would have taken six full days per school. Unfortunately, it was not logistically possible for the schools to set aside space for such an extended period of time.

No time was taken to prepare the participants for learning with multimodal representations, whilst all participants were experienced users of textual representations. The pilot study used for the studies described in Chapters 2 and 3 did not give reason to include a training session. This may have given the pupils in the textual conditions an advantage over those in the visual, timeline, concrete, abstract, and combined conditions. In hindsight, some pupils did seem to have problems with the cartogram task (see Chapter 3), and it might have been a good idea to build up the multimodal representations instead of introducing all elements at once, in particular in the experiment described in Chapter 4, where the abstract and concrete elements of the representations were presented rather than assembled by the pupils. Building up the representations might be especially important in the light of Nadolski, Kirschner, and

Van Merriënboer (2005) who suggest that there might be different levels of construction involved in complex learning tasks. This idea might be particularly helpful in improving the cartogram task by introducing steps to help pupils understand the maps.

The task texts used for the studies reported on in Chapters 2 and 3 had different numerical and topic characteristics, and these differences were not evened out between the four tasks (Chapter 3), in part because the different topics required different texts. Thus, it is possible to ascribe differences found between tasks to differences between the accompanying texts. Recently research has been done on suitable text characteristics focusing specifically on the same type of learners as in this thesis (pre-vocational secondary education, VMBO). Research by Land, Sanders and Van den Bergh (2005) shows that formal texts and narrative texts can have very different effects: Whilst the former improves comprehension, the latter increases enjoyment, and thus motivation. Although the texts used for the four tasks were carefully written to ensure their internal coherence, the texts were not aligned to each other regarding a number of text characteristics, such as the number of words. Possibly, the textual condition performed as well as the other two conditions because the texts by themselves were powerful enough to result in sufficient understanding of the topic. Still, the design of the first study might have been strengthened by removing the pictures from the task texts, although these texts and accompanying pictures were the same for all conditions. This might have enlarged the differences between the textual and multimodal conditions.

In addition, it should also be noted that one condition was missing from the study in Chapter 2: textual representations integrated in a timeline. The conclusion that the timeline combined with multimodal representations is the catalyst for integration of textual and visual information must be drawn cautiously: Possibly, differences between the conditions are attributable to the timeline per se, and not to the combination of timeline and multimodal representations.

In the experimental study with abstract and concrete visualisations (Chapter 4) no differences between conditions were found. However, the test in that experiment only measured factual recognition and recall. Van Meter and Garner (2005) showed that benefits of tasks in which students construct visual representations (in their case student-generated drawings) are revealed especially in higher-order assessments, such as the test items used in the first experiment (Chapter 2).

5.2.3 Domain and topics

The tasks in the studies in Chapters 2 and 3 were presented in a fixed order for all dyads in all conditions. It may seem a little odd to start a section on domain and topics by saying that the tasks were not ordered randomly, as in a solid research design they should be. However, this thesis deals with history learning, and chronology is an inevitable ingredient of any topic in the domain of history – probably even its main characteristic. As was noted on several occasions in this thesis, pupils have problems developing a consistent chronological frame of reference. Consequently, the chronological order of the topics used for the different tasks described in this thesis was kept intact.

The next question is: Would the results have been different if different topics had been chosen for the task content? This question can be looked at from two perspectives, one comparing different conditions, and another comparing different tasks. For the first point of view, it is important to realise that a topic can have a certain level of ‘imaginability’ or concreteness, but also that different texts about the same topic can have different levels of ‘imaginability’ or concreteness. For example, an explanation of the fall of the Roman Empire can be quite abstract, whereas an explanation of the retreat of the Roman administrators from Western Europe which is a subtopic of the former can be more easily visualised as actual people leaving a part of the world that still exists on today’s maps. Also, Vikings, for example, seem to tickle many children’s imagination, and once they know what they look like (the Vikings are part of the primary school history curriculum in The Netherlands) they will likely not need extra pictures to make their imagination run wild. Following this line of thought, a study comparing textual and multimodal tasks about abstract topics should lead to larger differences than a study comparing textual and multimodal tasks about concrete topics. However, the topics in the studies discussed in this thesis all incorporated both abstract and concrete aspects: not the abstractness or concreteness of the topics was varied, but the abstractness or concreteness of the representation tasks (see Chapter 4).

The second point of view takes us back to Chapter 3, where different types of multimodal representations were compared. There, the problem was that the four tasks also had different topics, and different types of topics (e.g., developments, structure, and causal relations). It might have been interesting to vary only one of these characteristics at a time across four tasks. The four tasks would then have had a) the same representation type and topic type, but a different topic (e.g., four developments, each shown in a process diagram), b) the same representation type and topic, but a different topic type, or c) the same topic and topic type, but a different representation type. However, although it is possible to represent a topic from the viewpoint of different topic types (situation b), for example by showing the *development* of manorialism, and the *structure* of manorialism, the contents covered would not be the same, and no single representation type would be suitable for representing both topic types. Similarly, using different representations to represent the same topic types can easily result in ill-matched topic-representation combinations. In other words, it is not always possible to vary the factors of representation type and topic type, as they are often inherently linked to each other.

5.3 Theoretical implications

Although Dual Coding Theory (Paivio, 1991) and the Cognitive Theory of Multimedia Learning (Mayer, 2001) and other previous research on learning with multimodal representations offer valuable insights, they do have their limitations, most notably in (1) their generalisability from the technical or natural science domains to other domains, (2) their tendency to focus on presented representations rather than construction of representations by learners, and (3) their focus on individual settings and learning outcomes rather than on collaborative settings and learning processes. The results of the studies described in this thesis are discussed in the light of these aspects below.

5.3.1 Multimodal representations in different domains

Because of the specificity of the domain of history, generalisability of the results of this thesis to other domains or topics may be limited. However, this is also one of the main reasons why the domain of history was chosen: Most studies on learning with multimodal representations have been limited to the natural sciences. Consequently, the research presented in this thesis shows that perhaps learning with multimodal representations per se is not as beneficial as some of the research on science learning suggests, although there does seem to be a bright future for learning by integrating multiple multimodal representations.

The introduction to this thesis described several differences between the domains of history and science. It was noted that representations in the natural sciences mainly show ‘How things work’, suiting the needs of learning in that domain. At the same time, these representations also often show ‘What things look like’. History learning on the other hand requires building a chronological frame of reference. Knowing ‘What things look like’ certainly plays a role in forming such a framework, but what holds it together is knowledge of ‘What things mean’ and ‘Why things happened when they happened’. For example, after reading this thesis one may be able to interpret Figure 5.1b, but Figure 5.1a can be interpreted more easily.

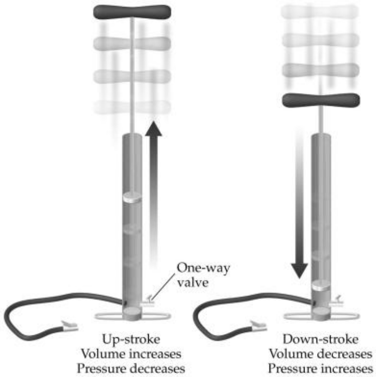


Figure 5.1a. Operation of a bicycle pump (from Tro, 2003).

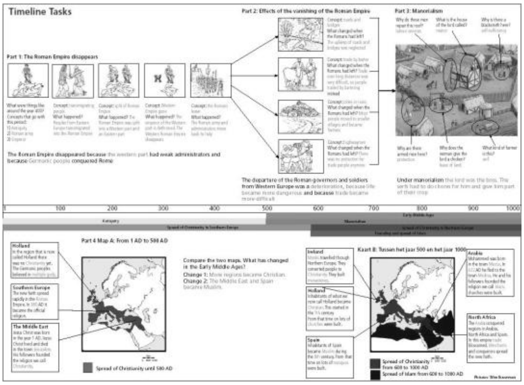


Figure 5.1b. Timeline of the transition between Antiquity and the Early Middle Ages.

On the other hand, the task in the first study that most closely resembled a ‘How things work’-approach (i.e., Task 1 with its process diagram) also seemed to elicit the most productive dialogue. Perhaps multimodal representations are most effective for topics dealing with step-by-step processes such as the fall of the Roman Empire in Task 1.

The scale of the represented phenomena might also play a role here, both in terms of length of time, and in terms of the size of objects or locations. Perhaps it is more difficult to understand a timeline covering several ages with multiple events and phenomena, than understanding how a bicycle pump fills a bicycle tire with air in a few seconds, because the latter can be felt, relived in its entirety, making it easier also to visualise. In addition, the scale of the overall objects is very different, as Figure 5.1a and Figure 5.1b show: A bicycle pump can be held in one hand, whilst a village or Empire cannot. Within the history domain this aspect of scale could be investigated further by comparing, for example, learning about European versus local issues.

As was noted earlier in Chapters 2 and 4, the grammar – or semiotics – of visualisations in the domain of history has no unambiguous principles or rules as there are in science. A domain like science, even though it deals with complex phenomena like waves, offers a structured visual grammar linked to basic governing principles – such as frequency and wavelength –that remain the same, whether the wave runs through water, air or solid matter. In history, however, even seemingly simple symbols like arrows can be interpreted in different ways, for example as indicating dynamic relations such as causality, or merely as temporal relations, or as static relations describing the structure of phenomena (O’Donnell et al., 2002). For example, the arrow shown in Figure 5.2, with a picture of a woman giving a chicken to a man, could be interpreted in several ways: Why does she give him a chicken? Did it escape? Is she paying him? Is he her husband, her boss, or a tax collector? Does she get anything in return? The truth in the context of the Early Middle Ages is that the woman is a serf, who gives part of her crop to the lord in exchange for protection. However, even when the context is known pictures and schemas are often still open to multiple interpretations. This makes it hard for students to learn to understand the form of representation – the importance of which is underlined by Ainsworth (2006): There are no fixed formats in history, for example for visualising time, war or manorialism.

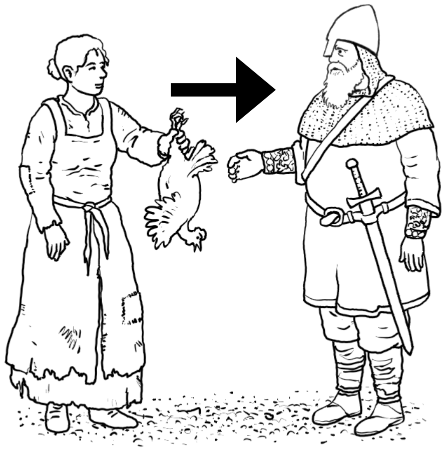


Figure 5.2. What does the arrow mean?

5.3.2 Activity continuum: presented vs. construction

As was noted in Chapter 1, some studies on collaborative learning with multimodal representations in small groups have shown that new domain knowledge can be built by articulating or representing ideas and by making connections in the shape of concrete, visual structures in concept maps (Van Boxtel, Van der Linden & Kanselaar, 2000; Roth & Roychoudhury, 1994). By comparing interaction processes and learning outcomes of learning activities with different types of representations we can gain insight into the possibilities, and even more so the limitations of each of the representational forms in (collaborative) history learning. The different types used in the studies in Chapters 2 and 3 were: making a timeline, completing a storyboard (process diagram), completing a causal map (network chart), describing a historical image (structure diagram), and completing a cartogram. A comparison of these different multimodal representation tasks in Chapter 3 showed that they do elicit different dialogue, and some of the characteristics of the four different tasks might well be linked to their place on the activity continuum. The task that seemed to encourage the most content utterances involved the activity of ordering or sequencing pictorial information, whilst less physical activity was involved in the tasks that elicited significantly fewer content utterances. Although no dialogue analysis was done for the study in Chapter 4, perhaps the absence of significant differences between conditions in that study is attributable to the lack of physical activities – thus minimizing differences between the conditions (i.e., textual, concrete, abstract, and combined). At the other end of the continuum, research by Van Boxtel and Van Rijn (2006) suggests that letting pupils draw their own pictures – in their study drawings of the fall of the Roman Empire – leads to better learning outcomes than presenting pictures to the pupils.

5.3.3 Facilitating collaborative learning with multimodal representations

At the start of this research project, it was thought that multimodal representations would have a facilitating role in collaborative learning, by providing a common referent that makes communication easier, and by providing a shared frame of reference. Dialogue analyses of the Timeline condition in Chapter 3 showed that some of the content utterances did indeed include references to visual elements of the multimodal representations. Thus, multimodal representations do seem to facilitate discussion of content. One of the differences found between the textual and integrated multimodal conditions that might explain the difference found in learning outcomes was, however, not an increase in explicit talking about the frame of reference, but rather an increase in implicit reference to content.

As was noted in Chapter 3, verbal deixis (e.g., through use of demonstrative pronouns such as “this”, “that one”) in the dialogue suggested that physical deixis (i.e., extra-linguistic behaviour such as pointing) may have played an important role in the communication between dyad partners. Pupils working individually would not point or refer to anything, simply because there would be no one to point things out for or to talk to. Still, the deictic properties of multimodal versus textual representations might be clarified to some extent by videotaping dialogues for collaborative groups, versus eye movement tracking for individuals. This would be an interesting topic for future research.

5.3.4 Integrating multiple multimodal representations

As was noted in the introductory chapter, acquiring a chronological frame of reference calls for visualisation and conceptualisation of specific periods in history and of relationships between aspects of these periods. Previous research has shown that students have trouble developing a consistent ‘chain of events and developments’, and as a result they cannot properly remember the chronological overview or reuse it, causing confusion of phenomena and concepts (Beck & McKeown, 1994; Leinhardt, Stainton & Virji, 1994). The results of the analysis of learning outcomes reported in Chapter 2 show that integrating multiple multimodal representations in a timeline does encourage the development of a chronological frame of reference. Dialogue analysis confirmed the idea that visualisation and conceptualisation are important, as there was a positive correlation between these aspects of the dialogue and learning outcomes for learners working with multimodal timeline tasks. However, integration did not lead to higher learning outcomes in the long run. Possibly, the chronological order of events is harder to remember in the long term than other types of relations, and a timeline difficult to reconstruct after a while so that it no longer functions as an anchor for remembering.

5.3.5 Artificial laboratory studies vs. ecologically valid field studies

The research reported on in this thesis was done in a classroom setting in pre-vocational secondary education, where the tasks were incorporated in the normal history curriculum. Pupils were graded for their work and worked for an entire lesson during each session. Learning with multimodal representations in the context of the Cognitive Theory of Multimedia Learning has been primarily tested in artificial, experimental, laboratory-like settings, where participants were – sometimes paid – volunteers, who did not receive grades for their learning outcomes, and where distraction by other participants or surroundings was eliminated as much as possible. In addition, the learning assignments were usually quite short, lasting from 30 seconds to 20 minutes. For determining the value of theories on learning with multimodal representations for use in education and in learning materials could benefit from an increase in the amount of ecologically valid research, as there is apparently a gap between traditional laboratory research and ecologically valid classroom-based research. As a result, it is not clear whether learning with multimodal representations has equivalent positive effects on learning outcomes in the context of the classroom where there is more interference from factors such as motivation, or class atmosphere. The studies in this thesis are an attempt to bridge this gap, and Table 5.1 gives an overview of some of the major differences between Mayer's research on which the Cognitive Theory of Multimedia Learning is based (Mayer, 2001), and the studies described in this thesis. This gap may be bridged further by replicating better controlled laboratory research in more ecologically valid settings, varying specific factors very carefully (e.g., learner characteristics, domain characteristics, use of collaboration, types of representations, integration of representations) one at a time. The research reported in this thesis also suggests that in realistic educational settings, additional requirements may play a role in the effectiveness of learning with multimodal representations. For example, multimodal tasks might be more effective when they are embedded in a series of lessons that include whole-class instruction and take into account the optimum attention span of the pupils.

Table 5.1. Comparison between Mayer's work and the studies in this thesis.

Aspects	Mayer	This thesis
setting	lab	school
domain	physics	history
age	18-20	12-14
education level	higher education psychology students	pre-vocational secondary education
location	US	NL
time	5 minutes	30 minutes
embedding	no embedding	in curriculum
group size	individual	collaborating dyads
activities	study the item	complete
integration	no integration of multiple representations	integration of multiple representations

5.4 Future research

It would be interesting to investigate how text can support correct interpretation of representations in different domains, as Ainsworth (2006) claims that text guides the interpretative process in picture comprehension. In history a picture without text will often result in a guessing game. Text is required there. Again, this might be very different for domains or topics that deal with 'How things work'. Future research could investigate what text characteristics are suitable for a task in which visual and textual elements need to be integrated.

Although no explicit attempt was made to falsify the Cognitive Theory of Multimedia Learning in this thesis, the results suggest that prudence is in order in generalising the principles that lie at the foundation of CTML to their use and effects in actual learning situations in home, study and school settings. Many other studies referring to Mayer's work – whether they are replication studies or indirect validations such as the studies in this thesis – have raised the same doubts. The modality effect, for example, could not be confirmed by a number of studies (De Westelinck, Valcke, De Craene, & Kirschner, 2005; Jeung & Chandler, 1997; Kalyuga, Chandler, & Sweller, 2000; Tabbers, 2002; Tabbers, 2006). It would be interesting to see how well CTML holds when one element is changed at a time, for example, the context, the domain, or learner type.

The research was done in pre-vocational schools with students who do not have much prior knowledge of the topics chosen for the tasks. The effects of using multimodal representations were often weaker than expected. In future research variation in student characteristics – such as prior knowledge, level of experience in the construction of multimodal representations in the domain of history, and collaboration skills – is necessary to investigate the prerequisites for using this type of material in pre-vocational education.

The results in this thesis suggest that it is worth taking a closer look at the role of deictic properties of multimodal representations in collaborative learning, both with respect to verbal deixis and physical deixis (Suthers, Girardeau, & Hundhausen, 2003). A similar approach is also suggested by Jewitt, Kress, Ogborn, and Tsatsarelis (2001), who make a plea for a multimodal approach to interaction research where not just the linguistic interaction is analysed, but also the visual and physical interaction.

Finally, there is a dearth of research on how different types of representations (such as the eleven categories defined in Lohse, Biolsi, Walker, and Rueter, 1994) work differentially by eliciting differences in learning processes. In addition, more research is needed on how these different types of representations work together. From a domain-specific point of view, this could lead to further research on understanding and remembering temporal relations through construction of timelines, including through integration of multiple (multimodal) representations.

References

Ainsworth, S. (1999). The functions of multiple representations. *Computers & Education, 33*, 131-152.

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction, 16*, 183-198.

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist, 28*, 117-148.

Barnes, S. (2002). Revealing the big picture: Patterns, shapes and images at key stage 3. *Teaching History*, (107), 6-12.

Beck, I. L., & McKeown, M. G. (1994). Outcomes of history instruction: Paste-up accounts. In M. Carretero & J. F. Voss (Eds.), *Cognitive and instructional processes in history and the social sciences* (pp. 237-256). Hillsdale, NJ: Lawrence Erlbaum.

Bennett, N., & Dunne, E. (1991). The nature and quality of talk in co-operative classroom groups. *Learning and Instruction, 1*, 103-118.

Bodemer, D., Ploetzner, R., Bruchmüller, K., & Häcker, S. (2005). Supporting learning with interactive multimedia through active integration of representations. *Instructional Science, 33*, 73-95.

Boggiano, A. K., Main, D. S., & Katz, P. A. (1988). Children's preference for challenge: The role of perceived competence and control. *Journal of Personality and Social Psychology, 54*, 134-141.

Butcher, K. R. (2006). Learning from text with diagrams: Promoting mental model development and inference generation. *Journal of Educational Psychology, 98*, 182-197.

Carney, R. N., & Levin, J. R. (2002). Pictorial illustrations *still* improve students' learning from text. *Educational Psychology Review, 14*, 5-26.

Carretero, M., Asensio, M., & Pozo, J. I. (1991). Cognitive development, historical time representation and causal explanations in adolescence. In M. Carretero, M. Pope, R. J. Simons, & J. I. Pozo (Eds.), *Learning and instruction: Vol. 3. European research in an international context* (pp. 27-48). Oxford: Pergamon.

Carretero, M., Jacott, L., Limón, M., López-Manjón, A., & León, J. A. (1994). Historical knowledge: Cognitive and instructional implications. In M. Carretero & J. F. Voss (Eds.), *Cognitive and instructional processes in history and the social sciences* (pp. 357-376). Hillsdale, NJ: Lawrence Erlbaum.

Carretero, M., López-Manjón, A., & Jacott, L. (1997). Explaining historical events. *International Journal of Educational Research, 27*, 245-253.

Chen, F.-C., Lee, Y.-W., Chu, H. C., Wang, H. R., & Jiang, H.-M. (2005). Effective discussions, social talks and learning: A paradox on learning in discussion forums. In T. Koschmann, D. D. Suthers, & T.-W. Chan (Eds.), *Computer supported collaborative learning 2005: The next 10 years! Proceedings of the international conference on computer supported collaborative learning 2005* (pp. 33-42). Mahwah, NJ: Lawrence Erlbaum.

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review, 3*, 149-210.

Committee of History and Social Studies. (2001). *Verleden, heden en toekomst* [Past, present and future]. Enschede, The Netherlands: SLO, Netherlands Institute for Curriculum Development.

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

Cox, R., & Brna, P. (1995). Supporting the use of external representations in problem solving: The need for flexible learning environments. *Journal of Artificial Intelligence in Education, 6*, 239-302.

Dawson, I. (2004). Time for chronology? Ideas for developing chronological understanding. *Teaching History*, (117), 14-24.

De Jong, T., Ainsworth, S., Dobson, M., Van der Hulst, A., Levonen, J., Reimann, P., Sime, J.-A., Van Someren, M. W., Spada, H., & Swaak, J. (1998). Acquiring knowledge in science and mathematics: The use of multiple representations in technology-based learning environments. In M. W. Van Someren, P. Reimann, H. P. A. Boshuizen, & T. De Jong (Eds.), *Learning with multiple representations* (pp. 9-40). Oxford: Pergamon.

De Westelinck, K., Valcke, M., De Craene, B., & Kirschner, P. (2005). Multimedia learning in social sciences: Limitations of external graphical representations. *Computers in Human Behavior, 21*, 555-573.

Erkens, G., Jaspers, J., Prangsma, M., & Kanselaar, G. (2005). Coordination processes in computer supported collaborative writing. *Computers in Human Behavior, 21*, 463-486.

Erkens, G., Kanselaar, G., Jaspers, J., & Schijf, H. (2001). Computerondersteund samenwerkend leren [Computer-supported collaborative learning]. In W. A. Wald & J. Van der Linden (Eds.), *Leren in perspectief [Viewpoints on learning]* (pp. 85-97). Apeldoorn, The Netherlands: Garant.

Fall, R., Webb, N. M., & Chudowsky, N. (2000). Group discussion and large-scale language arts assessment: Effects on students' comprehension. *American Educational Research Journal, 37*, 911-941.

Fasulo, A., Girardet, H., & Pontecorvo, C. (1998). Seeing the past: Learning history through group discussion of iconographic sources. In J. F. Voss & M. Carretero (Eds.), *International review of history education: Vol. 2. Learning and reasoning in history* (pp. 132-153). London: Woburn Press.

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

Friedman, W. J. (1982). Conventional time concepts and children's structuring of time. In W. J. Friedman (Ed.), *The developmental psychology of time* (pp. 171-208). New York: Academic Press.

Hacquebord, H. (2004). Taalproblemen en taalbehoeften in het voortgezet onderwijs: Leerlingen- en docentenvragenlijsten als instrumenten voor taalbeleid [Language problems and language needs in secondary education: Student and teacher questionnaires as instruments for language policy]. *Levende Talen, 5*(2), 17-28.

Hoodless, P. (1996). *Teaching of History Series: Vol. 69. Time and timelines in the primary school.* London: The Historical Association.

Horton, P. B., McConney, A. A., Gallo, M., Woods, A. L., Senn, G. J., & Hamelin, D. (1993). An investigation of the effectiveness of concept mapping as an instructional tool. *Science Education, 77*, 95-111.

Hunt, M. (2000). Teaching historical significance. In J. Arthur & R. Phillips (Eds.), *Issues in history teaching* (pp. 39-53). London: Routledge Falmer.

Husbands, C. (1996). *What is history teaching? Language, ideas and meaning in learning about the past*. Buckingham, UK: Open University Press.

Jeung, H.-J., & Chandler, P. (1997). The role of visual indicators in dual sensory mode instruction. *Educational Psychology, 17*, 329-343.

Jewitt, C., Kress, G., Ogborn, J., & Tsatsarelis, C. (2001). Exploring learning through visual, actional and linguistic communication: The multimodal environment of a science classroom. *Educational Review, 53*, 5-18.

Jonassen, D. H., Reeves, T. C., Hong, N., Harvey, D., & Peters, K. (1997). Concept mapping as cognitive learning and assessment tools. *Journal of Interactive Learning Research, 8,* 289-308.

Jones, B. F., Pierce, J., & Hunter, B. (1988, 1989). Teaching students to construct graphic representations. *Educational Leadership, 46*(4), 20-25.

Kalyuga, S., Chandler, P., & Sweller, J. (2000). Incorporating learner experience into the design of multimedia instruction. *Journal of Educational Psychology, 92*, 126.

Kirschner, P. A. (2002). Cognitive load theory: Implications of cognitive load theory on the design of learning. *Learning and Instruction, 12*, 1-10.

Kirschner, P. A. (2004). Design, development, and implementation of electronic learning environments for collaborative learning. *Educational Technology Research and Development, 52*(3), 39-46.

Land, J., Sanders, T., & Van den Bergh, H. (2005). *De ideale leertekst voor het vmbo: Welke kenmerken maken een leertekst het meest geschikt voor een leerling op het vmbo?* [The ideal learning text for pre-vocational secondary education: What characteristics make a text most suitable for pupils in pre-vocational secondary education?]. Utrecht, The Netherlands: Universiteit Utrecht & Stichting Lezen.

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 11*(1), 65-99.

Lee, P. (2005). Historical literacy: Theory and research. *International Journal of Historical Learning, Teaching and Research, 5*(1), 25-40.

Leinhardt, G. (1993). Weaving instructional explanations in history. *British Journal of Educational Psychology, 63*, 46-74.

Leinhardt, G. (1994). History: A time to be mindful. In G. Leinhardt, I. L. Beck, & C. Stainton (Eds.), *Teaching and learning in history* (pp. 209-256). Hillsdale, NJ: Lawrence Erlbaum.

Leinhardt, G., Stainton, C., & Virji, S. M. (1994). A sense of history. *Educational Psychologist, 29*, 79-88.

Lohse, G. L., Biolsi, K., Walker, N., & Rueter, H. H. (1994). A classification of visual representations. *Communications of the ACM, 37*(12), 36-49.

Masterman, E., & Rogers, Y. (2002). A framework for designing interactive multimedia to scaffold young children’s understanding of historical chronology. *Instructional Science, 30*, 221-241.

Mayer, R. E. (2001). *Multimedia learning.* Cambridge, UK: Cambridge University Press.

Mayer, R. E., & Chandler, P. (2001). When learning is just a click away: Does simple user interaction foster deeper understanding of multimedia messages? *Journal of Educational Psychology, 93*, 390-397.

Mayer, R. E. (2003). The promise of multimedia learning: using the same instructional design methods across different media. *Learning and Instruction, 13,* 125-139.

Mayer, R. E., & Sims, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a Dual-Coding Theory of Multimedia Learning. *Journal of Educational Psychology, 86,* 389-401.

Moreno, R., & Valdez, A. (2005). Cognitive load and learning effects of having students organize pictures and words in multimedia environments: The role of student interactivity and feedback. *Educational Technology Research & Development, 53*(3), 35-45.

Munneke, L., Andriessen, J., Kanselaar, G., & Kirschner, P. (2007). Supporting interactive argumentation: Influence of representational tools on discussing a wicked problem. *Computers in Human Behavior, 23*, 1072-1088.

Munneke, L., Van Amelsvoort, M., & Andriessen, J. (2003). The role of diagrams in collaborative argumentation-based learning. *International Journal of Educational Research, 39,* 113-131.

Nadolski, R. J., Kirschner, P. A., & Van Merriënboer, J. J. G. (2005). Optimizing the number of steps in learning tasks for complex skills. *British Journal of Educational Psychology, 75*, 223-237.

Novak, J. D. (1990). Concept maps and Vee diagrams: Two metacognitive tools to facilitate meaningful learning. *Instructional science, 19*, 29-52.

O'Donnell, A. M., Dansereau, D. F., & Hall, R. H. (2002). Knowledge maps as scaffolds for cognitive processing. *Educational Psychology Review, 14*, 71-86.

Paivio, A. (1991). *Images in mind: The evolution of a theory*. London: Harvester Wheatsheaf.

Peeck, J. (1993). Increasing picture effects in learning from illustrated text. *Learning and Instruction, 3,* 227-238.

Postigo, Y., & Pozo, J. I. (1998). The learning of a geographical map by experts and novices. *Educational Psychology, 18,* 65-80.

Reimann, P. (2003). Multimedia learning: Beyond modality. *Learning and Instruction, 13*, 245-252.

Robinson, D. H., Robinson, S. L., & Katayama, A. D. (1999). When words are represented in memory like pictures: Evidence for spatial encoding of study materials. *Contemporary Educational Psychology, 24*(1), 38-54.

Roschelle, J. (1992). Learning by collaborating: Convergent conceptual change. *Journal of the Learning Sciences, 2*, 235-276.

Roth, W.-M., & Roychoudhury, A. (1994). Science discourse through collaborative concept mapping: New perspectives for the teacher. *International Journal of Science Education, 6*, 437-455.

Saleh, M. A., Lazonder, A., & De Jong, T. (2005). Effects of within-class ability grouping on social interaction, achievement, and motivation. *Instructional Science, 33*, 105-119.

Salomon, G., & Leigh, T. (1984). Predispositions about learning from print and television. *Journal of Communication, 34*(2), 119-135.

Sauer, M. (2003). *Bilder im geschichtsunterricht: Typen, Interpretationsmethoden, Unterrichtsverfahren* [Images in history teaching: Types, interpretation methods, didactics] (2nd ed.). Seelze-Velber, Germany: Kallmeyersche Verlagsbuchhandlung.

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

Scevak, J. J., & Moore, P. J. (1998). Levels of processing effects on learning from texts with maps. *Educational Psychology, 18*, 133-155.

Schnotz, W. (1993). On the relation of dual coding and mental models in graphics comprehension. *Learning and Instruction, 3*, 247-249.

Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction, 13*, 141-156.

Shemilt, D. (2000). The caliph's coin: The currency of narrative frameworks in history teaching. In P. N. Stearns, P. Seixas, & S. Wineburg (Eds.), *Knowing, teaching, and learning history: National and international perspectives* (pp. 83-101). New York: New York University Press.

Slater, W. H., & Horstman, F. R. (2002). Teaching reading and writing to struggling middle school and high school students: The case for reciprocal teaching. *Preventing School Failure, 46*, 163-166.

Stern, E., Aprea, C., & Ebner, H. G. (2003). Improving cross-content transfer in text processing by means of active graphical representation. *Learning and Instruction, 13*, 191-203.

Stow, W., & Haydn, T. (2001). Issues in the teaching of chronology. In J. Arthur & R. Phillips (Eds.), *Issues in history teaching* (pp. 83-97). London: Routledge Falmer.

Suthers, D. D., & Hundhausen, C. D. (2001). *Learning by constructing collaborative representations: An empirical comparison of three alternatives.* Paper presented at the 1st European conference on Computer Supported Collaborative Learning, Maastricht, The Netherlands.

Suthers, D. D., & Hundhausen, C. D. (2003). An empirical study of the effects of representational guidance on collaborative learning. *Journal of the Learning Sciences, 12*, 183-219.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*, 257-285.

Sweller, J., & Chandler, P. (1991). Evidence for cognitive load theory. *Cognition and Instruction, 8*, 351-362.

Sweller, J., Van Merriënboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251-296.

Tabbers, H. K. (2002). *The modality of text in multimedia instructions: Refining the design guidelines.* Unpublished doctoral dissertation, Open University of The Netherlands, Heerlen, The Netherlands.

Tabbers, H. K. (2006). Where did the modality effect go? A failure to replicate one of Mayer's multimedia learning effects and why this still might make some sense. *Proceedings of the EARLI SIG Text and Graphics Comprehension*, Nottingham University, UK, 27-29.

Teasley, S. D., & Roschelle, J. (1993). Constructing a joint problem space: The computer as a tool for sharing knowledge. In S. P. Lajoie & S. J. Derry (Eds.), *Computers as cognitive tools* (pp. 229-257). Hillsdale, NJ: Lawrence Erlbaum.

Tro, N. J. (2003). Companion website for Introductory Chemistry. Retrieved February 15, 2007, from http://wps.prenhall.com/wps/media/objects/476/488316/ch11.html

Van Amelsvoort, M., Andriessen, J., & Kanselaar, G. (In press). Representational tools in computer-supported collaborative argumentation-based learning: How dyads work with constructed and inspected argumentative diagrams. *Journal of the Learning Sciences*.

Van Boxtel, C. A. M. (1994). *Historische begrippen: Een onderzoek naar de mentale representaties van leerlingen uit het voortgezet onderwijs* [Historical concepts: A study into the mental representations of pupils in secondary education]. Unpublished Master’s thesis, Utrecht University, Utrecht, The Netherlands.

Van Boxtel, C. A. M., & Van Rijn, M. (2006). Picturing colligatory concepts in history: Effects of student-generated versus presented drawings. *Proceedings of the EARLI SIG Text and Graphics Comprehension*, Nottingham University, UK, 5-7.

Van Boxtel, C. A. M., Van der Linden, J. L., Roelofs, E., & Erkens, G. (2002). Collaborative concept mapping: Provoking and supporting meaningful discourse. *Theory into Practice, 41*(1), 40-46.

Van Boxtel, C. A. M., Van der Linden, J., & Kanselaar, G. (1997). Collaborative construction of conceptual understanding: Interaction processes and learning outcomes emerging from a concept mapping and a poster task. *Journal of Interactive Learning Research, 8*, 341-361.

Van Boxtel, C. A. M., & Van Drie, J. (2004). Historical reasoning: A comparison of how experts and novices contextualise historical sources. *International Journal of Historical Learning, Teaching and Research, 4*(2), 89-97.

Van Boxtel, C. A. M., Van der Linden, J., & Kanselaar, G. (2000). Collaborative learning tasks and the elaboration of conceptual knowledge. *Learning and Instruction 10,* 311-330.

Van der Meij, J., & De Jong, T. (2003, August). *Learning with multiple representations: Supporting students' translation between representations in a simulation-based learning environment.* Paper presented at the 10th European Conference for Research on Learning and Instruction, Padova, Italy.

Van Drie, J. (2005). *Learning about the past with new technologies: Fostering historical reasoning in computer-supported collaborative learning.* Unpublished PhD thesis, Utrecht University, Utrecht, The Netherlands.

Van Drie, J., Van Boxtel, C., Erkens, G., & Kanselaar, G. (2005). Using representational tools to support historical reasoning in computer-supported collaborative learning. *Technology, Pedagogy and Education, 14*, 25-42.

Van Drie, J., Van Boxtel, C. A. M., Jaspers, J., & Kanselaar, G. (2005). Effects of representational guidance on domain specific reasoning in CSCL. *Computers in Human Behavior, 21*, 575-602.

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.

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

Vogt, R. (1999). Principals of biology II, Exam Number Two. Retrieved February 15, 2007, from http://www.biol.sc.edu/courses/bio102/answer2-99.html

Wiley, J., & Ash, I. K. (2005). Multimedia learning of history. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 375-391). New York: Cambridge University Press.

Wilkinson, I. A. G., & Fung, I. Y. Y. (2002). Small-group composition and peer effects. *International Journal of Educational Research, 37*, 425-447.

Zhang, J. (1997). The nature of external representations in problem solving. *Cognitive Science, 21*, 191-217.

Appendices

Appendix A: Model answers for the four tasks in the Visual and Timeline conditions (Chapters 2 and 3)

Text shown in italics had to be added by the students.

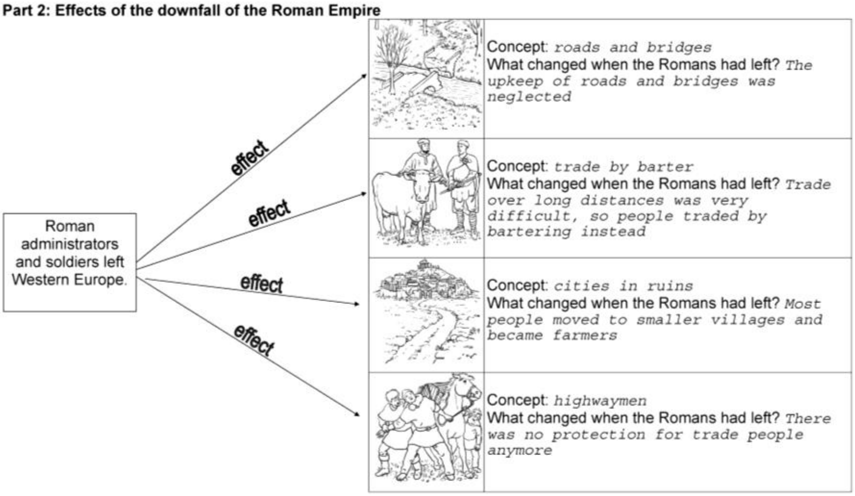
Part 1: The Roman Empire disappears

What were things like around the year 400? Concepts that go with this period: 1) *Antiquity* 2) *Roman army* 3) *Emperor*	Concept: *transmigrating people* What happened? *Peoples from Eastern Europe transmigrated into the Roman Empire*	Concept: *split of Roman Empire* What happened? *The Roman Empire was split into a Western part and an Eastern part*	Concept: *Western Empire gone* What happened? *The emperor of the Western part is dethroned. The Western Roman Empire disappears*	Concept: *the Romans leave* What happened? *The Roman army and administrators move back to Italy*

The Roman Empire disappeared because *the western part had weak administrators* and because *Germanic people conquered Rome*

Figure A.1. Model answers for Task 1 in the Visual and Timeline conditions: ordering pictures in a process diagram.

Part 2: Effects of the downfall of the Roman Empire



The departure of the Roman administrators and soldiers from Western Europe was a deterioration, because *life became more dangerous* and because *trade became more difficult*

Figure A.2. Model answers for Task 2 in the Visual and Timeline conditions: choosing pictures in a causal diagram.

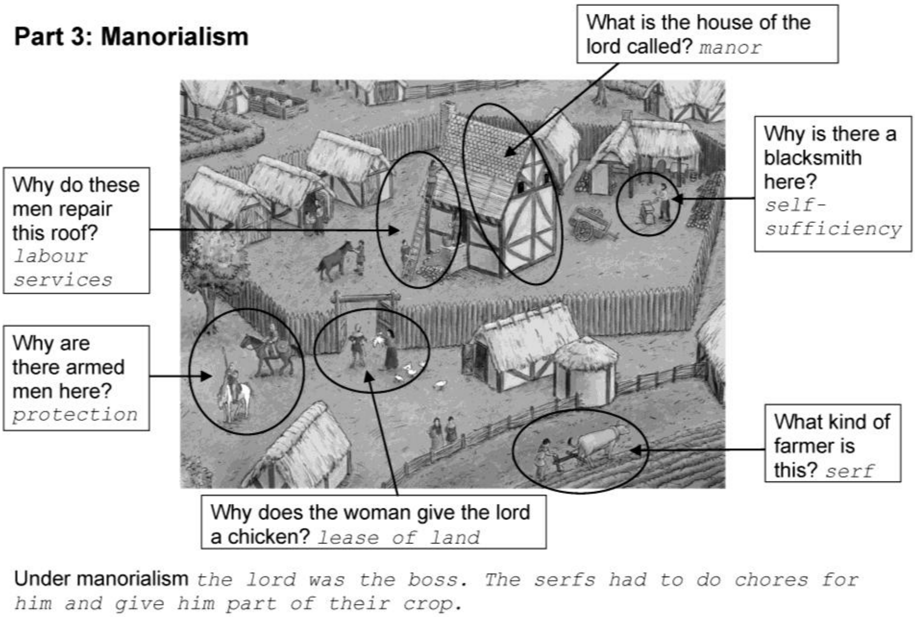


Figure A.3. Model answers for Task 3 in the Visual and Timeline conditions: labelling an image.

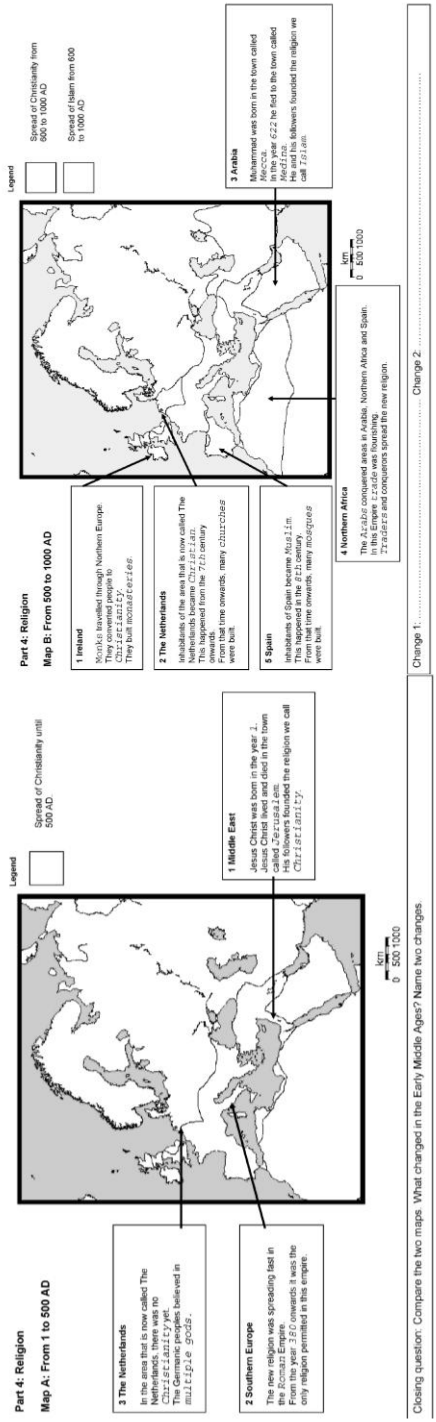
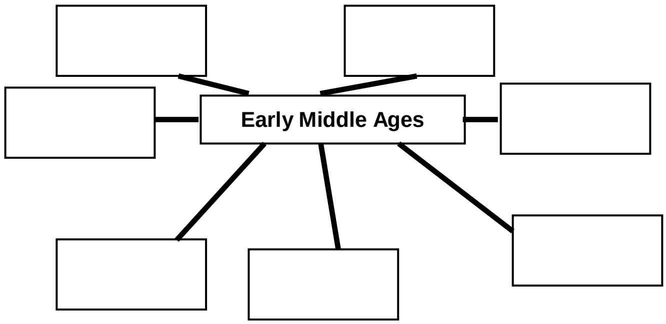


Figure A.4. Model answers for Task 4 in the Visual and Timeline conditions: completing two cartograms.

Appendix B: Test questions (Chapters 2 and 3)

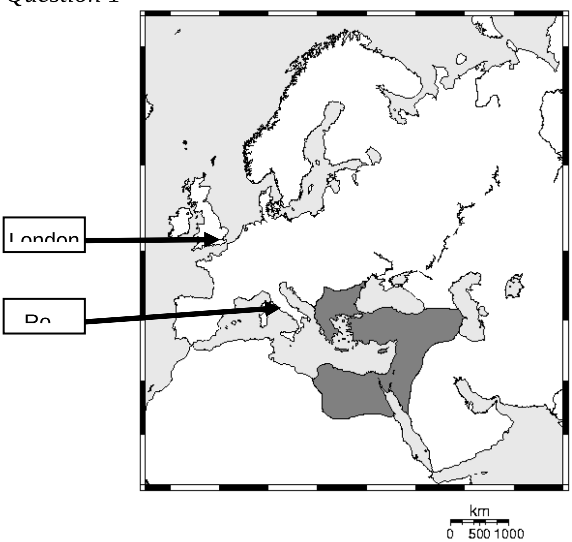
Part A: Free association spider

What do you know about the Early Middle Ages?
Write down as many things as possible in the boxes. You can also draw extra boxes.
You have 5 minutes to answer this question.



Part B: Open questions

Question 1



It is the year 500. John lives in London. He is looking for work. He used to be a merchant. He often travelled to Rome and bought fabrics and dried fruit there. He sold his goods on the London market. Now he has stopped with that long-distance trade.

What event in history is this related to?

Question 2

It is the year 600. Jan and his family have moved.
He used to live in a town. Now he is a farmer and lives in a village.
He and his wife and children live in a small cabin.
They work the land every day: ploughing, sowing, harvesting.
They also collect their chickens' eggs.
In the centre of the village lives a rich farmer.
Every week, Jan gives a chicken, eggs or part of his crop to that farmer.

2.a. Why did Jan give away those things? Jan did this because: …
2.b. What did he get in return? This is what he got in return: …

Question 3

It is the year 54. The monk Bonifatius travels through Europe. He came from Ireland, but now he is in Friesland. In the picture you can see Bonifatius with his **bible** in his hand. What was the monk Bonifatius doing in Northern-Europe?

Question 4

It is the year 900. Juan lives in Spain. He just moved.
He used to live in a village. Now Juan is a merchant and lives in a town.
He now lives with his wife and children in a beautiful house.
Juan and his wife both work in their shop every day.
They sell fabrics from Asia and Northern Africa.
They do a lot of business with foreign merchants.

4.a. In which empire did Juan and his wife live?
4.b. Juan and his family practiced the religion we call … .

Question 5

Write your answer to questions 5.a. and 5.b. on the timeline below.

5.a. Around what year did the Early Middle Ages begin? And when did they end? Colour the period in the timeline below. Write the name 'Early Middle Ages' above it.

5.b. Around what year did Islam arise?
Put a cross in the right place on the timeline. Write the name 'Islam' below it.

1 500 1000 1500 2000

Part C: Multiple choice questions

Question 1

The great Roman Empire disappeared. What caused this to happen?
Tick the **two** correct answers.

- ❑ The empire was increasingly badly governed.
- ❑ There was no longer any long-distance trade.
- ❑ Germanic peoples wandered into the empire.
- ❑ The Roman Empire became Christian.
- ❑ The Arabs marched into the empire.
- ❑ Roads and bridges were no longer cared for.

Question 2

When the Romans left Western Europe, trade changed.
What changed? Tick the **two** correct answers.

Trading became

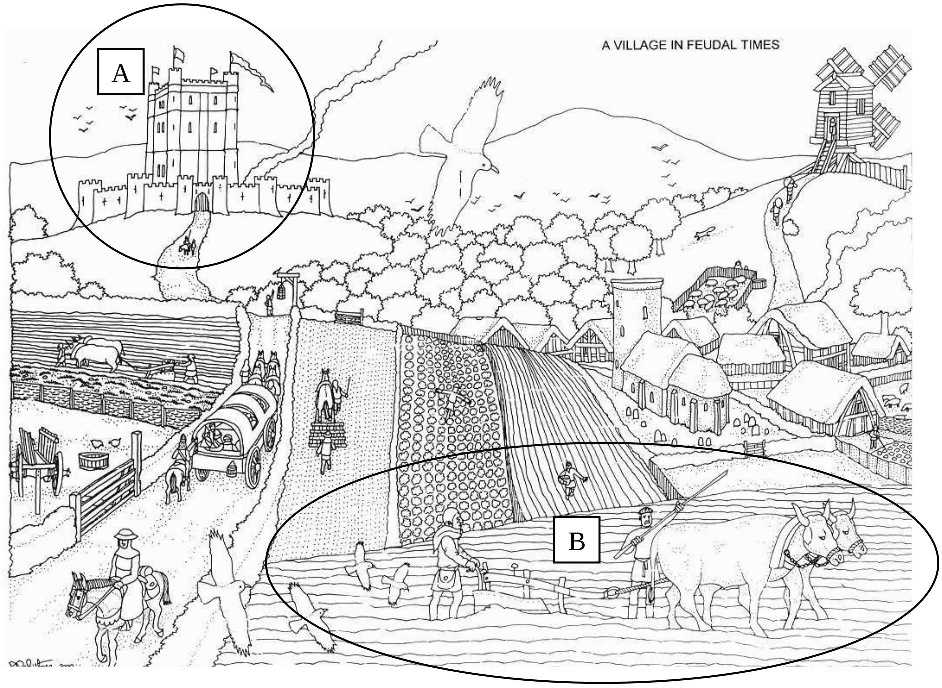
- ❑ harder, because upkeep on roads and bridges came to a halt.
- ❑ easier, because the Romans no longer patrolled the borders.
- ❑ more dangerous, because it was unsafe on the roads.
- ❑ safer, because there was protection from peoples travelling around the country.
- ❑ harder, because almost everyone converted to Christianity.
- ❑ easier, because it became more common to pay with coins.

Question 3

What goes with Roman times (Antiquity) and what goes with the Early Middle Ages? Circle the correct answer for each sentence.

	Circle the correct period:
There is an emperor who rules a large empire from Rome.	Romans / Early Middle Ages
There are lots of cities	Romans / Early Middle Ages
Vikings travels around looting.	Romans / Early Middle Ages
Upkeep on roads and bridges comes to a halt.	Romans / Early Middle Ages
There is a lot of long-distance trade.	Romans / Early Middle Ages

Question 4
Look at the picture below about the Early Middle Ages and answer the questions.



Circle **A**: Who lives in this building? There is only one correct answer.

- ❑ A lord
- ❑ A serf
- ❑ A Roman emperor
- ❑ A mayor

Circle **B**: These farmers did not work their own land. They have to work on the land owned by their boss. What do we call such farmers? There is only one correct answer.

- ❑ Labourers
- ❑ Serfs
- ❑ Slaves
- ❑ Poor farmers

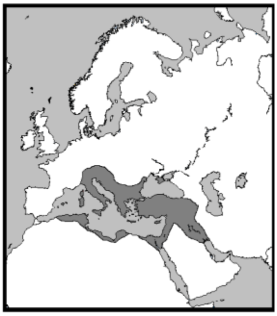
Question 5

Many buildings like this one were built in The Netherlands between 700 and 900 AD. Why did this happen during that period? There is one correct answer. Tick the correct answer.

- ❑ This is when the Romans were ruling The Netherlands.
- ❑ This is when most people in The Netherlands became Christian.
- ❑ This is when Christianity was founded.
- ❑ This is when people had lots of money to build churches.

Question 6

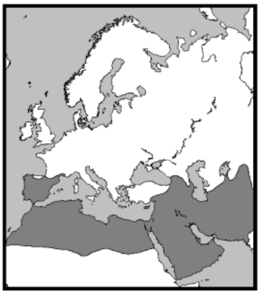
Below you see four maps. Which map represents Christian territory around the year 1000? Tick the box under the correct map.



Map A ☐

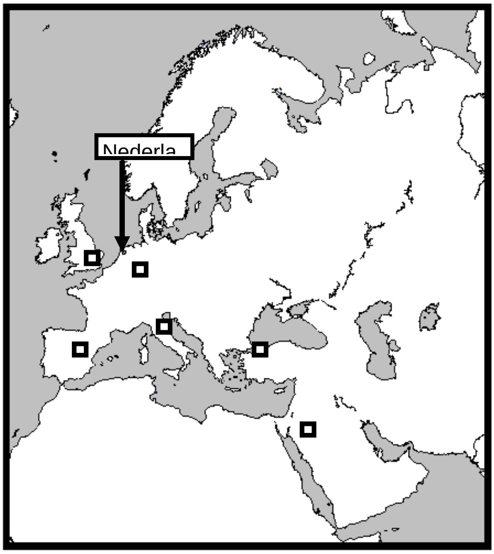
Map B ☐

Map C ☐



Map D ☐

Question 7
The map below shows six areas with a check box.
Tick the two areas that became Islamic during the Early Middle Ages.



Question 8
What came first? Each time, circle what came **first**.

What came first:	**Roman Empire**	or	**manorialism**
What came first:	**wandering of nations**	or	**rise of Islam**
What came first:	**Christianity in the Netherlands**	or	**decline of Roman Empire**

Appendix C: Task sheets for the Textual, Abstract, Concrete and Combined conditions (Chapter 4)

500 to 1000 AD - The Western Roman Empire disappears: What changed?

There are no more ___________ who organise everything.

There are no more ___________ to protect people.

Upkeep on ___________ stopped.

Therefore ___________ became more difficult.

After 500 AD most people lived from ___________

And they made everything they needed themselves.

There was also ___________

It became ___________

___________ traveled around looting.

Farmers looked for protection with a ___________

He employed ___________

The farmers gave ___________ to the lord in exchange for ___________

These farmers are called ___________

Figure C.1. Task sheet for the Textual condition.

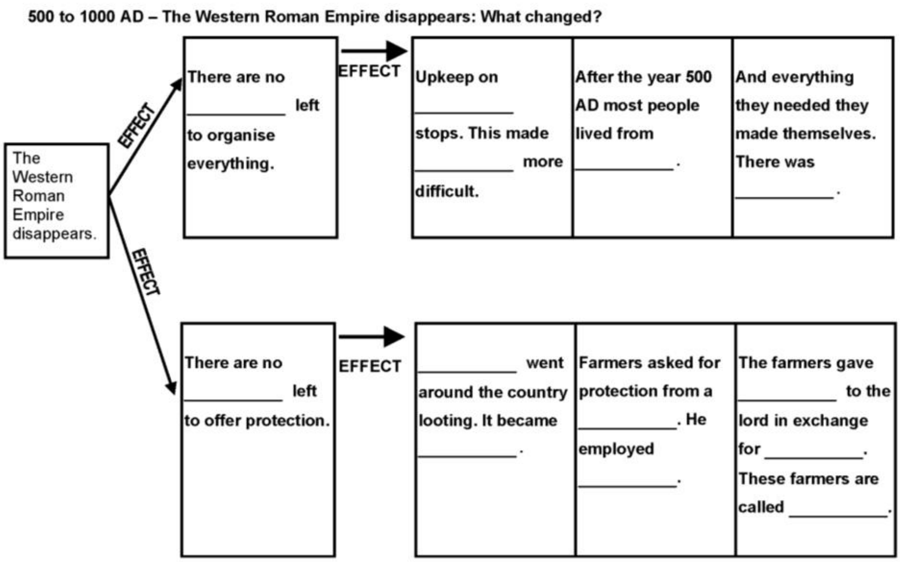


Figure C.2. Task sheet for the Abstract condition.

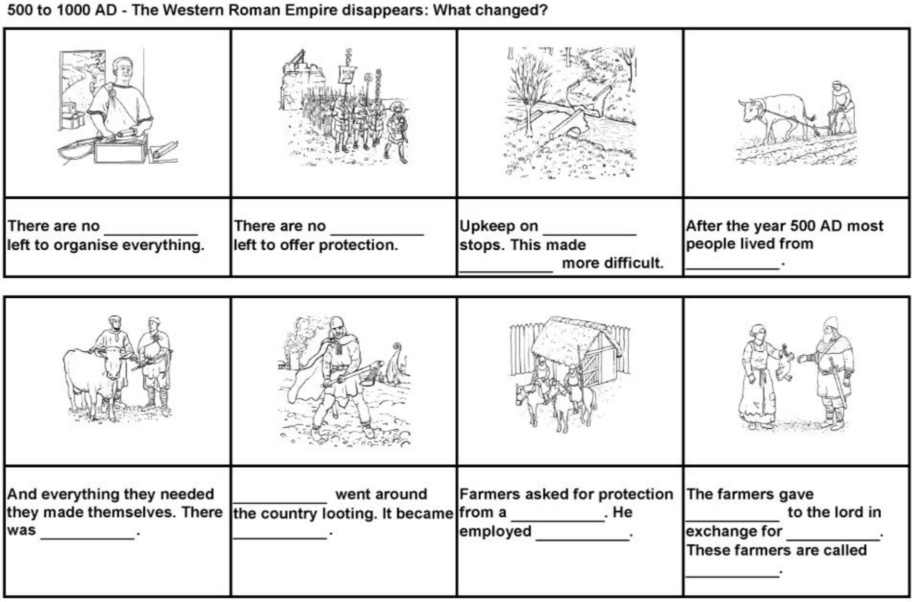


Figure C.3. Task sheet for the Concrete condition.

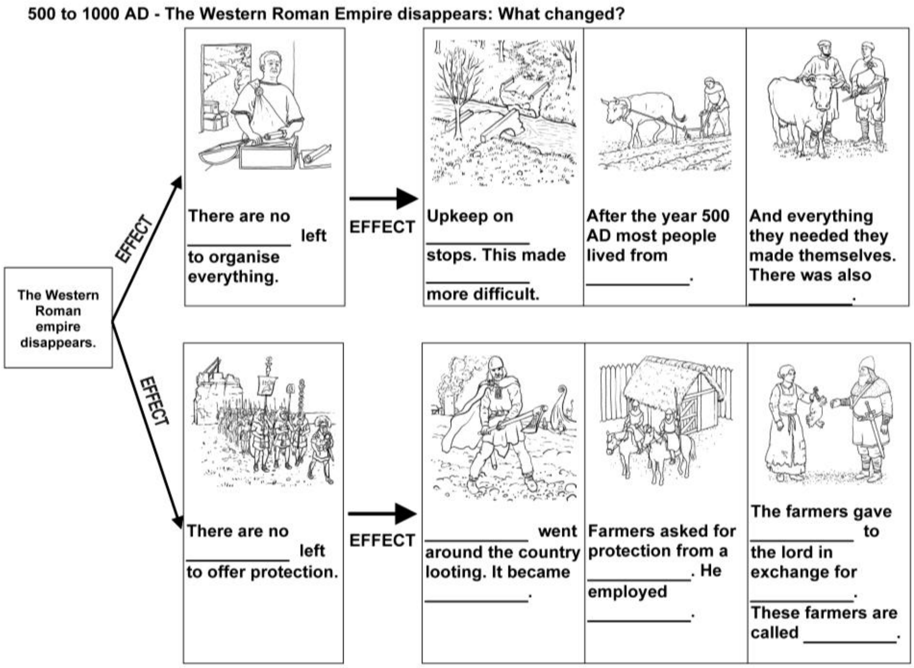


Figure C.4. Task sheet for the Combined condition.

Appendix D: Test questions (Chapter 4)

Part A

Around 500 AD the Western Roman Empire disappeared. This had far-reaching consequences. What changed when the Western Roman Empire disappeared?
Circle the correct answer (RIGHT or WRONG) for each position.

1	Viking raids were giving people trouble.	RIGHT / WRONG
2	The empire was gone, but the Roman era was not over yet.	RIGHT / WRONG
3	Lots of new roads and bridges were built.	RIGHT / WRONG
4	There were no Roman administrators left to take care of things.	RIGHT / WRONG
5	Long-distance trade became more difficult.	RIGHT / WRONG
6	Life became safer.	RIGHT / WRONG
7	More people had to live from agriculture.	RIGHT / WRONG
8	People started trading more.	RIGHT / WRONG
9	People were less protected against looting and attacks.	RIGHT / WRONG
10	The Roman era ended.	RIGHT / WRONG
11	Roman administrators stayed to govern the territory.	RIGHT / WRONG
12	Roman soldiers left.	RIGHT / WRONG
13	Vikings started ruling the area.	RIGHT / WRONG
14	Many roads became dilapidated.	RIGHT / WRONG

Part B

How did people live between 500 and 1000 AD, when the Western Roman Empire had disappeared?
Circle the correct answer (RIGHT or WRONG) for each position.

15	Serfs gave part of their crop to the lord.	RIGHT / WRONG
16	Farmers lived on the land of the lord.	RIGHT / WRONG
17	People paid with things instead of money.	RIGHT / WRONG
18	As it was less safe, farmers sought protection from a lord.	RIGHT / WRONG
19	Most people lived from trade.	RIGHT / WRONG
20	People made most of the things they needed themselves.	RIGHT / WRONG
21	Serfs gave part of their crop to a rich Roman.	RIGHT / WRONG
22	There was a lot of long/distance trade..	RIGHT / WRONG
23	As life was more dangerous, people moved into the towns.	RIGHT / WRONG
24	The serfs owned the land they worked on.	RIGHT / WRONG
25	Roman soldiers protected the farmers.	RIGHT / WRONG
26	Serfs gave part of their crop in exchange for protection.	RIGHT / WRONG
27	Almost everyone lived on agriculture.	RIGHT / WRONG
28	Roads and bridges were well-kept.	RIGHT / WRONG

Summary

We cannot – in the 21st century – imagine a school textbook not containing many different types of illustrations. Modern schoolbooks are packed with pictures, tables, graphs, and diagrams – in addition to texts. Visual or multimodal representations are seen as being more than just motivational or illustrative in nature, but rather also as being instrumental in encouraging and facilitating meaningful and/or deep learning. Multimodal representations are representations containing a combination of text, schemas, pictures, and/or animations. Most research on learning with multiple representations is based on Paivio's Dual Coding Theory and on Mayer's Cognitive Theory of Multimedia Learning (Paivio, 1991; Mayer, 2001). Dual Coding Theory assumes that information is processed through one of two separate but interrelated channels: the verbal channel or the visual channel. It predicts that adding pictures to text will benefit learning in most cases, as pictures are more likely to be processed both verbally and visually. This will result in more elaborate encoding, and the learner is provided with more retrieval cues (Paivio, 1991). Positive effects of using multimodal representations in learning materials have been found within the domain of science and technology and in the context of individual use of presented multiple representations (Mayer, 2001). However, different fields of knowledge may raise different possibilities for the use of multimodal representations (De Westelinck, Valcke, De Craene, & Kirschner, 2005). Also, theory seems to suggest that assembling and constructing multimodal representations – as opposed to simply presenting them – more strongly encourages articulation of ideas and content, discussion, and deep processing (Cox, 1999). Recent research on collaborative construction of representations has yielded positive results for both learning processes and learning outcomes (Suthers & Hundhausen, 2003; Van Drie, Van Boxtel, Jaspers, & Kanselaar, 2005; Van Boxtel, Van der Linden, & Kanselaar, 2000).

The studies described in this thesis deal with the potential and the effects of using multimodal representations for learning from the domain of history in co-construction tasks. The main topic of this thesis is multimodal representation used in acquiring a chronological frame of reference in history. The aim is to determine whether using multimodal representations in history learning tasks makes a difference for the learning outcomes achieved by students and/or the learning process that they carried out. The PhD thesis focuses on the question: How does making and connecting different types of multimodal representations affect the collaborative learning process and the acquisition of a chronological-conceptual frame of reference in 12 to 14-year olds enrolled in pre-vocational education?

In the school subject of history, students need to construct a chronological frame of reference that they can use to investigate, describe, explain and evaluate historical and contemporary phenomena. A chronological frame of reference is the knowledge base that is used when reasoning about the past. It consists of knowledge about: (1) historical phenomena, (2) temporal and causal relations between those phenomena, and (3) concepts for describing those phenomena and relations. A prerequisite for remembering and applying a chronological frame of reference is that students can visualise the historical phenomena that are part of the chronological overview to be acquired and that they can relate these phenomena to other phenomena and to concepts describing these phenomena and relations. Research by Beck and McKeown (1994) has shown that students have difficulty developing a coherent chain of events, and that the schemas students use are too general to offer ready slots to fit the

specific information that they might have gleaned from the learning materials. In addition, the specific information is too sparse to be useful in connecting it to more general information. Furthermore, students have particular difficulty forming a notion of complex historical developments and structures (Carretero, Asensio, & Pozo, 1991; Husbands, 1996). Making these developments, structures, temporal and causal relations visible through pictures and diagrams can render abstract phenomena and relations more explicit. The timeline is a representation type that is typical for history, and that functions as a tool to represent time and to give a chronological overview. However, a timeline with dates or periods and textual descriptions of historical phenomena only visualises the temporal relationships. It does not show the underlying causal relationships between the events and phenomena that it contains. Moreover, a timeline only gives limited support in visualising and grasping the character of an era. Other representations – such as historical pictures, animations, matrices comparing periods, numerical information and causal networks – might combine with a timeline to facilitate visualisation and construction of relationships, requiring explicit linking of the different representations. The historical phenomena being visualised need to be contextualised in time. Aspects of continuity and change cannot be recognised and causal relationships cannot be drawn until the temporal relationships are clear. In addition, causal relationships do not just require insight into temporal relations, but also insight into the types of phenomena to be explained. Connecting different types of information is thus crucial in history learning. Integrating different multimodal representations can support this, for example through a timeline consisting of elements such as pictures, verbal descriptions, and arrows indicating the causal relations. The studies in this thesis look at the possible added value of a method in which different types of representations are related to each other, namely of a timeline incorporating verbal, depictive and schematic representations.

The thesis discusses four research questions. It was hypothesised that the active completion and construction of multimodal representations and especially the integration of verbal and visual information in a timeline would help students to develop an overview of historical phenomena and relations.

All studies in this thesis worked with 12 to 14-year-old participants in the first year of prevocational secondary education (VMBO). A majority of Dutch students in secondary school (approximately 60%) attend this type of school. The language proficiency of these students is often lower than desired. History, as a school subject, is part of the compulsory curriculum for only the first two years for these students, so there is little time for developing a chronological frame of reference.

The topics of all learning materials used in the three studies reported on here were taken from the Early Middle Ages. This period – 500 to 1000 AD in (Western) European history – was selected for several reasons. First, the period includes a wide range of types of phenomena that are dealt with in the history curriculum (e.g., events, developments, structures, themes). In addition, the existing curriculum for this period includes different types of developments and social structures that are closely related. It marks a turning point in Western European history with the decline of the Roman Empire and the subsequent development of manorialism (Dutch: hofstelsel), and these two developments can only be understood in relation to each other. A number of important socioeconomic, political and religious changes took place during the Early Middle Ages. The period includes some very abstract concepts (e.g., manorialism, empire), and life during this period was in many ways very different from our students' lives today. On top of that, there is very little original visual material available from the

Early Middle Ages that can help students in shaping a notion of this period. The multimodal representations used in this thesis were constructed especially for educational purposes.

The first study – described in **Chapter 2** – answers the first research question: What are the effects of the general type of co-constructed representation – textual, multimodal, or integrated multimodal – on the acquisition of a chronological-conceptual frame of reference? Eighty-five students in pre-vocational secondary education worked in dyads on a series of four history tasks in one of three conditions: with textual representation tasks, with multimodal representation tasks, or with multimodal representation tasks integrated in a timeline. All students took a pre-test, post-test and retention test. The four tasks dealt with (1) the decline of the Roman Empire (process diagram), (2) the effects of the fall of the Roman Empire (network chart), (3) manorialism (structure diagram), and (4) the spread of Christianity and Islam (two cartograms).

Based on the literature on learning with multimodal representations and what is important for the acquisition of a chronological-conceptual frame of reference in history, the collaborative completion and construction of multimodal representations was expected to contribute more to learning than the collaborative completion and construction of textual representations. Moreover, integration of such multimodal representations in a timeline was expected to strengthen this effect, because the timeline also visualises temporal and causal relations between different phenomena.

Results show that working on multimodal representations integrated in a timeline leads to higher short-term results than co-constructing textual representations. The chapter also discusses an exploratory analysis of the dialogues of two dyads working in the condition with multimodal representations integrated in a timeline who showed a strong improvement in their test results between pre-test and post-test. This analysis indicates that the extent to which students verbally integrate textual and visual information differs for the four different multimodal timeline tasks.

The second study – described in **Chapter 3** – answers two research questions: (1) What are the effects of the collaborative completion and construction of integrated multimodal representations versus textual representations on the content of the student dialogue, and (2) are there any differences in domain-specific content of the student dialogue between collaborative completion and construction of different types of multimodal representations – structure diagram, network chart, process diagram, or cartogram?

The recorded dialogues of 20 dyads working on a series of four history tasks in one of two conditions were analysed: with textual representation tasks, or with multimodal representation tasks integrated in a timeline. The student utterances were coded as procedural utterances, social utterances or content utterances. Content utterances were further analysed for degree of specificity (core vs. auxiliary), reference to visual elements, integration of modes, and concept use. The comparison between conditions focused on the second of four tasks (that is, the network chart task dealing with the effects of the fall of the Roman Empire). Additional analyses were done on the 10 dyads in the Timeline condition to compare different multimodal tasks, and to correlate the dialogue outcomes and the post-test results.

The results show that working on integrated multimodal representations do – to some extent – lead to more discussion about domain content, as well as about procedural issues. Furthermore, a comparison between the dialogues for different integrated multimodal tasks showed that the task involving completion of a storyboard or process diagram elicited a higher proportion of content utterances, use of more historical concepts, and a higher proportion of content utterances with reference to visual elements than the other tasks. The results also show that discussing content with reference to schemas and/or pictures as well as using more domain-specific concepts correlate positively with learning outcomes.

The third study – described in **Chapter 4** – answers the fourth and final research question: What are the effects of combining text and different types of visualisations – abstract and concrete – in collaborative history tasks on learning historical phenomena and concepts?

First-year students (*N* = 104) worked in randomly assigned pairs. After reading a text, the pairs were given a learning task in one of four conditions: Textual, Concrete visualised, Abstract visualised, or Combined. The tasks in the Textual condition contained just fill-in-the-blank sentences. The Abstract condition contained the same fill-in-the-blank sentences, but in the form of a causal schema. The Concrete condition included eight pictures to illustrate the fill-in-the-blank sentences. The tasks in the Combined condition combined all three elements: the fill-in-the-blank sentences and pictures were integrated in the causal schema. It was hypothesised that the conditions with visualisations would perform better on a post-test and retention test. As learning history requires understanding both abstract and concrete phenomena, concepts, and relations, a combination of abstract and concrete visualisations should have the greatest potential for enhancing learning.

Post-test and retention test results revealed no significant differences between the conditions. There were some significant differences on the evaluation questionnaire. Combining text and different types of visualisations in learning tasks does not necessarily enhance history learning. Possible explanations given are the lack of activity with the multimodal representations, the ecological setting of the research, the semiotics of visualisations in the domain of history – that are not defined clearly – and the difficulty of unequivocally visualising historical concepts. Though no significant differences between conditions were found, the evaluation questionnaire showed that motivation and perceived competence were higher in the condition with concrete visualisations.

Several conclusions can be drawn from these results of the studies described in this thesis. First, learning with multimodal representations without integration in a timeline does not result in higher learning outcomes than learning with textual representations – neither in the short term, nor in the long run. Second, learning with multimodal representations integrated in a timeline resulted in higher learning outcomes than learning with textual representations – although the difference disappears in time. Third, compared to learning with textual representations, learning with multimodal tasks integrated in a timeline does affect the collaborative learning process by eliciting more dyadic interaction in general and more auxiliary content utterances in particular. Also, discussing content by referring to visual elements and using more historical concepts are positively related to short-term learning outcomes. Finally, learning with either concrete or abstract multimodal representations, or with a combination of concrete and abstract multimodal representations does not result in significantly higher learning outcomes than learning with textual representations alone.

Future research should focus on how different types of representations work differentially by eliciting differences in learning processes. In addition, more research is needed on how these different types of representations work together.

References

Beck, I. L., & McKeown, M. G. (1994). Outcomes of history instruction: Paste-up accounts. In M. Carretero & J. F. Voss (Eds.), *Cognitive and instructional processes in history and the social sciences* (pp. 237-256). Hillsdale, NJ: Lawrence Erlbaum.

Carretero, M., Asensio, M., & Pozo, J. I. (1991). Cognitive development, historical time representation and causal explanations in adolescence. In M. Carretero, M. Pope, R. J. Simons, & J. I. Pozo (Eds.), *Learning and instruction: Vol. 3. European research in an international context* (pp. 27-48). Oxford, UK: Pergamon.

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

De Westelinck, K., Valcke, M., De Craene, B., & Kirschner, P. (2005). Multimedia learning in social sciences: Limitations of external graphical representations. *Computers in Human Behavior, 21*, 555-573.

Husbands, C. (1996). *What is history teaching? Language, ideas and meaning in learning about the past*. Buckingham, UK: Open University Press.

Mayer, R. E. (2001). *Multimedia learning.* Cambridge, UK: Cambridge University Press.

Paivio, A. (1991). *Images in mind: The evolution of a theory*. London: Harvester Wheatsheaf.

Suthers, D. D., & Hundhausen, C. D. (2003). An empirical study of the effects of representational guidance on collaborative learning. *Journal of the Learning Sciences*, *12,* 183-219.

Van Boxtel, C. A. M., Van der Linden, J., & Kanselaar, G. (2000). Collaborative learning tasks and the elaboration of conceptual knowledge. Learning and Instruction 10, 311-330.

Van Drie, J., Van Boxtel, C. A. M., Jaspers, J., & Kanselaar, G. (2005). Effects of representational guidance on domain specific reasoning in CSCL. *Computers in Human Behavior, 21,* 575-602.

Samenvatting

We kunnen ons in de 21ste eeuw geen schoolboek voorstellen dat geen veelheid aan verschillende illustraties bevat. Moderne schoolboeken staan bol van de plaatjes, tabellen, grafieken en diagrammen – naast tekst. Visuele of multimodale representaties worden niet alleen gezien als motiverend of illustratief, maar ook als instrumenteel in het aanmoedigen en vergemakkelijken van betekenisvol en/of diep leren. Multimodale representaties zijn representaties die een combinatie bevatten van tekst, schema’s, afbeeldingen en/of animaties. Het meeste onderzoek naar leren met multipele representaties is gebaseerd op Paivio’s Dual Coding Theory en op Mayer’s Cognitive Theory of Multimedia Learning (Paivio, 1991; Mayer, 2001). Dual Coding Theory gaat ervan uit dat informatie wordt verwerkt via één van twee aparte, maar gerelateerde kanalen: het verbale of het visuele kanaal. De theorie voorspelt dat het toevoegen van afbeeldingen aan tekst in de meeste gevallen het leren verbetert, omdat de kans groter is dat afbeeldingen zowel verbaal als visueel worden verwerkt. Dit kan resulteren in een uitgebreidere codering en de lerende kan beschikken over meer stimuli om informatie uit het geheugen op te roepen retrieval cues (Paivio, 1991). Positieve effecten van het gebruik van multimodale representaties in leermiddelen zijn gevonden in het domein van de natuurwetenschappen en in de context van individueel gebruik van aangeboden multiple representaties (Mayer, 2001). Verschillende kennisgebieden kunnen echter verschillende mogelijkheden bieden voor het gebruik van multimodale representaties (De Westelinck, Valcke, De Craene, & Kirschner, 2005). Ook lijkt de theorie erop te wijzen dat het afmaken en construeren van multimodale representaties – in tegenstelling tot het simpelweg beschikbaar hebben – een sterker positief effect heeft op het expliciteren van ideeën en inhoud, op discussie en op diepe verwerking (Cox, 1999). Recent onderzoek op het gebied van samenwerkend construeren van representaties laat positieve resultaten zien voor zowel leerprocessen als leeruitkomsten (Suthers & Hundhausen, 2003; Van Drie, Van Boxtel, Jaspers, & Kanselaar, 2005; Van Boxtel, Van der Linden, & Kanselaar, 2000).

De studies die in dit proefschrift worden beschreven, hebben betrekking op de potentie en de effecten van het gebruik van multimodale representaties voor samenwerkend leren in het domein geschiedenis. Het hoofdonderwerp van dit proefschrift is multimodale representatie voor het verwerven van een chronologisch referentiekader bij het vak geschiedenis. Het doel is te bepalen of het gebruik van multimodale representaties bij geschiedenistaken een verschil kan maken voor de leeruitkomsten van leerlingen en/of voor het leerproces. Het proefschrift richt zich op de vraag: Hoe kan het maken en verbinden van verschillende typen multimodale representaties invloed uitoefenen op het proces van samenwerkend leren en op de verwerving van een chronologisch-conceptueel referentiekader door 12 tot 14-jarigen in het VMBO?

Bij het vak geschiedenis moeten leerlingen een chronologisch referentiekader construeren dat ze kunnen gebruiken om historische en hedendaagse verschijnselen te onderzoeken, te beschrijven, uit te leggen en te evalueren. Een chronologisch referentiekader is de kennisbasis die wordt gebruikt wanneer men nadenkt over het verleden. Het bestaat uit kennis over: (1) historische verschijnselen, (2) temporele en causale relaties tussen die verschijnselen, en (3) begrippen voor het beschrijven van die verschijnselen en relaties. Een voorwaarde voor het onthouden en toepassen van een chronologisch referentiekader is dat studenten de historische verschijnselen die onderdeel zijn van het te verwerven chronologisch overzicht kunnen visualiseren en dat ze deze verschijnselen kunnen relateren aan andere verschijnselen en aan begrippen die deze verschijnselen en relaties beschrijven. Onderzoek van Beck en McKeown (1994) heeft aangetoond dat leerlingen moeite hebben met het ontwikkelen van een coherente keten van gebeurtenissen en dat de schema's die leerlingen gebruiken te algemeen zijn om ze pasklare vakjes te kunnen bieden waarin de specifieke informatie past die ze wellicht uit de leermaterialen hebben gehaald. Bovendien is de specifieke informatie te beperkt om van nut te kunnen zijn in het verbinden aan meer algemene informatie. Ook hebben leerlingen vooral moeite met het vormen van begrip van complexe historische ontwikkelingen en structuren (Carretero, Asensio, & Pozo, 1991; Husbands, 1996). Door deze ontwikkelingen, structuren, temporele en causale relaties zichtbaar te maken via afbeeldingen en diagrammen kunnen abstracte verschijnselen en relaties explicieter gemaakt worden. De tijdbalk is een representatietype dat typisch is voor geschiedenis dat functioneert als een hulpmiddel voor het representeren van tijd en een chronologisch overzicht. Een tijdbalk met data of perioden en tekstuele beschrijvingen van historische verschijnselen visualiseert echter alleen de tijdsrelaties. Het toont niet de onderliggende causale relaties tussen de gebeurtenissen en verschijnselen. Bovendien ondersteunt een tijdbalk slechts in beperkte mate het visualiseren van een periode en het begrip van de aard ervan. Andere representaties – zoals historische afbeeldingen, animaties, vergelijkingsmatrices, cijfermatige informatie en causale netwerken – vergemakkelijken wellicht in combinatie met een tijdbalk de visualisatie en constructie van relaties, waarvoor expliciet verbinden van verschillende representaties noodzakelijk is. De historische verschijnselen die worden gevisualiseerd, moeten in de tijd worden geplaatst. Aspecten van continuïteit en verandering kunnen niet worden herkend en causale relaties kunnen niet worden gelegd totdat de tijdsrelaties helder zijn. Daarnaast is voor causale relaties niet alleen inzicht nodig in tijdsrelaties, maar ook in de typen verschijnselen die moeten worden verklaard. Het verbinden van verschillende informatietypen is dus van cruciaal belang bij geschiedenis leren. Het integreren van verschillende multimodale representaties kan dit ondersteunen, bijvoorbeeld door een tijdbalk bestaande uit elementen zoals afbeeldingen, verbale beschrijvingen, en pijlen die causale relaties aangeven. De studies in dit proefschrift bekijken de mogelijke toegevoegde waarde van een methode waarbij verschillende typen representaties met elkaar worden verbonden, namelijk een tijbalk met daarin verbale, beeldende en schematische representaties.

Dit proefschrift bespreekt vier onderzoeksvragen. De hypothese was dat het actief afmaken en construeren van multimodale representaties en met name het integreren van verbale en visuele informatie in een tijdbalk leerlingen zou helpen een overzicht te ontwikkelen van historische verschijnselen en relaties.

Alle studies in dit proefschrift hebben betrekking op 12 tot 14 jaar oude leerlingen in het eerste jaar van het voorbereidend middelbaar beroepsonderwijs (VMBO). Een meerderheid van de Nederlandse leerlingen in het voortgezet onderwijs (circa 60%) bezoekt dit schooltype. De taalvaardigheid van deze leerlingen is vaak lager dan gewenst. Geschiedenis als schoolvak is alleen in de eerste twee jaar een verplicht onderdeel van het curriculum voor deze leerlingen, zodat er weinig tijd is voor het ontwikkelen van een chronologisch referentiekader.

De onderwerpen van alle leermaterialen in de drie studies in dit proefschrift zijn afkomstig uit de Vroege Middeleeuwen. Deze periode – 500 tot 1000 na Christus in de (West-)Europese geschiedenis – werd gekozen om verschillende redenen. Ten eerste omvat deze periode een breed scala aan verschillende typen verschijnselen die worden onderscheiden binnen het geschiedeniscurriculum (bijvoorbeeld gebeurtenissen, ontwikkelingen, structuren, thema’s). Daarnaast bevat het bestaande curriculum voor deze periode verschillende typen ontwikkelingen en sociale structuren die nauw met elkaar zijn verbonden. De periode markeert een keerpunt in de West-Europese geschiedenis met het verval van het Romeinse Rijk en de daaropvolgende ontwikkeling van het hofstelsel, en deze twee ontwikkelingen kunnen alleen in relatie tot elkaar worden begrepen. Een aantal belangrijke sociaaleconomische, politieke en religieuze veranderingen vond plaats tijdens de Vroege Middeleeuwen. De periode omvat een aantal zeer abstracte begrippen (bijvoorbeeld hofstelsel, keizerrijk) en het leven tijdens deze periode verschilde in veel opzichten sterk van het moderne leven van onze leerlingen. Bovendien is er erg weinig origineel visueel materiaal bewaard gebleven uit de Vroege Middeleeuwen dat leerlingen kan helpen een beeld te vormen van deze periode. De multimodale representaties die zijn gebruikt voor dit proefschrift zijn alle speciaal ontworpen voor onderwijsdoelen.

De eerste studie – beschreven in **Hoofdstuk 2** – beantwoordt de eerste onderzoeksvraag: Wat zijn de effecten van het type geco-construeerde representatie in het algemeen – tekstueel, multimodaal, of geïntegreerd multimodaal – op de verwerving van een chronologisch-conceptueel referentiekader? Vijfentachtig leerlingen uit het VMBO werkten in tweetallen aan een serie van vier geschiedenistaken binnen één van drie condities: met tekstuele representatietaken, met multimodale representatietaken, of met multimodale representatietaken geïntegreerd in een tijdbalk. Alle leerlingen maakten een voortoets, een natoets en een retentietoets. De vier taken gingen over (1) de val van het Romeinse Rijk (procesdiagram), (2) de gevolgen van de val van het Romeinse Rijk (netwerkdiagram), (3) hofstelsel (structuurdiagram), en (4) de verspreiding van Christendom en Islam (twee kartogrammen).

Op basis van de literatuur over leren met multimodale representaties en wat belangrijk is voor het verwerven van een chronologisch-conceptueel referentiekader bij geschiedenis, werd verwacht dat het gezamenlijk afmaken en construeren van multimodale representaties meer zou bijdragen aan leren dan het gezamenlijk afmaken en construeren van tekstuele representaties. Bovendien werd verwacht dat integratie van dergelijke multimodale representaties in een tijdbalk dit effect zou versterken, omdat de tijdbalk ook de temporele en causale relaties tussen verschillende verschijnselen visualiseert.

De resultaten laten zien dat het werken met multimodale representaties geïntegreerd in een tijdbalk leidt tot betere resultaten op de korte termijn dan co-constructie van tekstuele representaties. Het hoofdstuk bespreekt ook een explorerende analyse van de dialogen van twee tweetallen die werkten in de conditie met multimodale representaties geïntegreerd in een tijdbalk en die een sterke verbetering in toetsresultaten lieten zien tussen voortoets en natoets. Deze analyse wijst erop dat de mate waarin leerlingen tekstuele en visuele informatie verbaal integreren verschillend is voor de vier verschillende multimodale tijdbalkopdrachten.

De tweede studie – beschreven in **Hoofdstuk 3** – beantwoordt twee onderzoeksvragen: (1) Wat zijn de effecten van het gezamenlijk afmaken en construeren van geïntegreerde multimodale representaties versus tekstuele representaties op de inhoud van de leerlingdialogen, en (2) zijn er verschillen in domeinspecifieke inhoud van de leerlingdialogen tussen het gezamenlijk afmaken en construeren van verschillende typen multimodale representaties – structuurdiagram, netwerkdiagram, procesdiagram, of kartogram?

De opgenomen dialogen werden geanalyseerd van 20 tweetallen die werkten aan een serie van vier geschiedenistaken in één van twee condities: met tekstuele representatietaken, of met multimodale representatietaken geïntegreerd in een tijdbalk. De leerlinguitspraken werden gecodeerd als procedurele uitspraken, sociale uitspraken of inhoudelijke uitspraken. Inhoudelijke uitspraken werden verder geanalyseerd voor de mate van specificiteit (kern vs. algemeen), verwijzing naar visuele elementen, integratie van modi, en gebruik van begrippen. De vergelijking tussen condities richtte zich op de tweede van vier taken (het netwerkdiagram over de gevolgen van het verdwijnen van het Romeinse Rijk). Aanvullende analyses werden gedaan voor de 10 tweetallen in de Tijdbalk-conditie om verschillende multimodale taken met elkaar te vergelijken en om de dialooguitkomsten te correleren met de natoetsscores.

De resultaten laten zien dat werken aan geïntegreerde multimodale representaties tot op zekere hoogte leidt tot meer discussie over domeininhoud en ook over procedurele zaken. Verder laat een vergelijking tussen de dialogen voor verschillende geïntegreerde multimodale taken zien dat de taak waarbij een stripverhaal of procesdiagram moest worden afgemaakt, een grotere proportie inhoudelijk uitspraken ontlokte dan de andere drie taken, meer historische begrippen, en een grotere proportie inhoudelijke uitspraken waarbij verwezen werd naar visuele elementen. De resultaten laten ook zien dat het bespreken van inhoud met verwijzing naar schema's en/of afbeeldingen en het gebruik van meer domeinspecifieke begrippen positief gecorreleerd is aan leeruitkomsten.

De derde studie – beschreven in **Hoofdstuk 4** – beantwoordt de vierde en laatste onderzoeksvraag: Wat zijn de effecten van het combineren van tekst en verschillende typen visualisaties – abstract en concreet – in samenwerkend geschiedenis leren op het leren van historische verschijnselen en begrippen?

Eerstejaars leerling ($N = 104$) werkten in random samengestelde tweetallen. Na het lezen van een tekst kregen de leerlingen een taak binnen één van vier condities: Tekstueel, Concreet visueel, Abstract visueel, of Gecombineerd. De taken in de Tekstuele conditie bevatten alleen invulzinnen. De Abstracte conditie bevatte dezelfde invulzinnen, maar in de vorm van een causaal schema. De Concrete conditie bevatte acht afbeeldingen ter illustratie van de invulzinnen. De taken in de Gecombineerde conditie combineerden alle drie de elementen: de invulzinnen en afbeeldingen waren geïntegreerd in het causaal schema. De hypothese was dat de condities met visualisaties beter zouden presteren op een natoets en een retentietoets. Voor geschiedenis leren is het noodzakelijk dat leerlingen begrip hebben van abstracte en concrete verschijnselen, begrippen en relaties, en daarom zou een combinatie van abstracte en concrete visualisaties het grootste potentieel moeten hebben voor het ondersteunen van leren.

De resultaten van de natoets en de retentietoets lieten geen significante verschillen zien tussen de condities. Er waren wel enkele significante verschillen voor de evaluatievragen. Het combineren van tekst en verschillende typen visualisaties in leertaken verbetert niet noodzakelijkerwijs geschiedenisleren. Mogelijke verklaringen die worden besproken zijn het gebrek aan activiteit met de multimodale representaties, de ecologische setting van het onderzoek, de semiotiek van de visualisaties in het domein geschiedenis – die niet duidelijk is gedefinieerd – en de moeilijkheid om historische begrippen eenduidig te visualiseren. Hoewel geen significante verschillen tussen condities werden gevonden, lieten de evaluatievragen zien dat de motivatie en de geschatte prestatie door de leerlingen hoger waren in de conditie met concrete visualisaties.

Uit de resultaten van de studies in dit proefschrift kunnen verschillende conclusies getrokken worden. Ten eerste resulteert leren met multimodale representaties zonder een geïntegreerde tijdbalk niet in betere leerresultaten dan leren met tekstuele weergave van geschiedenis – noch op korte termijn, noch op lange termijn. Ten tweede resulteerde leren met multimodale representatie geïntegreerd in een tijdbalk in hogere leeruitkomsten dan leren met tekstuele representaties – hoewel dit verschil op de lange termijn verdwijnt. Ten derde beïnvloedt leren met multimodale taken geïntegreerd in een tijdbalk het samenwerkend-leerproces meer dan leren met tekstuele representaties doordat meer interactie wordt ontlokt binnen het tweetal in het algemeen, en meer algemene inhoudelijke uitspraken in het bijzonder. Bovendien is er een positief verband tussen het bespreken van inhoud door te verwijzen naar visuele elementen en het gebruik van meer historische begrippen enerzijds en de leeruitkomsten op de korte termijn anderzijds. Ten slotte resulteert leren met concrete of abstracte multimodale representaties of met een combinatie van concrete en abstracte multimodale representaties niet in significant hogere leeruitkomsten dan leren met alleen tekstuele representaties.

Toekomstig onderzoek zou zich moeten richten op hoe verschillende typen representaties verschillend werken door het ontlokken van verschillende leerprocessen. Daarnaast is meer onderzoek noodzakelijk naar hoe deze verschillende typen representaties samenwerken.

Referenties

Beck, I. L., & McKeown, M. G. (1994). Outcomes of history instruction: Paste-up accounts. In M. Carretero & J. F. Voss (Eds.), *Cognitive and instructional processes in history and the social sciences* (pp. 237-256). Hillsdale, NJ: Lawrence Erlbaum.

Carretero, M., Asensio, M., & Pozo, J. I. (1991). Cognitive development, historical time representation and causal explanations in adolescence. In M. Carretero, M. Pope, R. J. Simons, & J. I. Pozo (Eds.), *Learning and instruction: Vol. 3. European research in an international context* (pp. 27-48). Oxford, UK: Pergamon.

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

De Westelinck, K., Valcke, M., De Craene, B., & Kirschner, P. (2005). Multimedia learning in social sciences: Limitations of external graphical representations. *Computers in Human Behavior, 21*, 555-573.

Husbands, C. (1996). *What is history teaching? Language, ideas and meaning in learning about the past*. Buckingham, UK: Open University Press.

Mayer, R. E. (2001). *Multimedia learning.* Cambridge, UK: Cambridge University Press.

Paivio, A. (1991). *Images in mind: The evolution of a theory*. London: Harvester Wheatsheaf.

Suthers, D. D., & Hundhausen, C. D. (2003). An empirical study of the effects of representational guidance on collaborative learning. *Journal of the Learning Sciences, 12*, 183-219.

Van Boxtel, C. A. M., Van der Linden, J., & Kanselaar, G. (2000). Collaborative learning tasks and the elaboration of conceptual knowledge. Learning and Instruction 10, 311-330.

Van Drie, J., Van Boxtel, C. A. M., Jaspers, J., & Kanselaar, G. (2005). Effects of representational guidance on domain specific reasoning in CSCL. *Computers in Human Behavior, 21,* 575-602.

List of publications

Accepted/submitted/in preparation

Prangsma, M. E., Van Boxtel, C. A. M., Kanselaar, G., & Kirschner, P. A. (2006). *Concrete and abstract visualisations in history learning tasks*. Manuscript submitted for publication.

Prangsma, M. E., Van Boxtel, C. A. M., Kanselaar, G., & Kirschner, P. A. (2007). *Multimodal representations in collaborative history tasks: A look at student dialogues*. Manuscript in preparation.

Erkens, G., Prangsma, M. E., Van Gisbergen, M., & Kanselaar, G. (2007). *Articulation of cognitive and metacognitive activities in computer-supported collaborative learning: The role of tools.* Manuscript in preparation.

Journal articles and book chapters

Prangsma, M. E., Van Boxtel, C. A. M., & Kanselaar, G. (in press). Developing a 'big picture': Effects of collaborative construction of multimodal representations in history. *Instructional Science*.

Erkens, G., Prangsma, M. E., & Jaspers, J. G. M. (2006). Planning and coordinating activities in collaborative learning. In A. M. O'Donnell, C. E. Hmelo-Silver, & G. Erkens (Eds.), *Collaborative learning, reasoning, and technology* (pp. 233-265). Mahwah, NJ: Lawrence Erlbaum.

Erkens, G., Jaspers, J. G. M., Prangsma, M. E., & Kanselaar, G. (2005). Coordination processes in computer supported collaborative writing. *Computers in Human Behavior, 21*(3), 463-486.

Erkens, G., Kanselaar, G., Prangsma, M. E., & Jaspers, J. G. M. (2003). Computer support for collaborative and argumentative writing. In E. De Corte, L. Verschaffel, N. Entwistle, & J. van Merriënboer (Eds.), *Powerful learning environments: Unravelling basic components and dimensions* (pp. 157-176). Amsterdam: Pergamon/Elsevier Science.

Kanselaar, G., Erkens, G., Andriessen, J. E. B., Prangsma, M. E., Veerman, A. L., & Jaspers, J. G. M. (2003). Designing argumentation tools for collaborative learning. In P. A. Kirschner, S. J. Buckingham Shum, & C. S. Carr (Eds.), *Visualizing argumentation: Software tools for collaborative and educational sense-making* (pp. 51-74). London: Springer-Verlag.

Erkens, G., Theil, J., Kanselaar, G., Prangsma, M. E., & Jaspers, J. G. M. (2002). 'W8ff:-)' Chattalk in een coöperatieve leeromgeving [Chat talk in a collaborative learning environment]. *Levende Talen Tijdschrift, 3*(2), 24-33.

Kanselaar, G., Andriessen, J. E. B., Erkens, G., Jaspers, J. G. M., Prangsma, M. E., & Veerman, A.L. (2002). Co-construction of knowledge in computer supported collaborative argumentation (CSCA). In P. Kirschner (Ed.), *Three worlds of CSCL: Can we support CSCL?* (pp. 93-130). Heerlen: Open Universiteit.

Erkens, G., Prangsma, M. E., Jaspers, J. G. M., & Kanselaar, G. (2002). *Computer support for collaborative and argumentative writing*. Final report NWO project. Utrecht: Department of Educational Sciences.

Erkens, G., Kanselaar, G., Prangsma, M. E., & Jaspers, J. G. M. (2002). Using tools and resources in computer supported collaborative writing. In G. Stahl (Ed.), *Computer support for collaborative learning: Foundations for a CSCL community* (pp. 389-399). Hillsdale, NJ: Lawrence Erlbaum.

Erkens, G., Jaspers, J. G. M., Tabachneck-Schijf, H. J. M., & Prangsma, M. E. (2001). Computer-supported collaboration in argumentative writing. In P. Dillenbourg, A. Eurlings, & K. Hakkarainen (Eds.), *European Perspectives on Computer-Supported Collaborative Learning (pp. 205-212). Proceedings of the First European Conference on Computer-Supported Collaborative Learning*. Maastricht: Maastricht University.

Conference presentations

Prangsma, M. E., Van Boxtel, C. A. M., Kanselaar, G., & Kirschner, P. A. (2007). *Concrete and abstract visualisations in history learning tasks.* Paper presented at the 2007 AERA Annual Meeting in Chicago, Illinois.

Prangsma, M. E., Van Boxtel, C. A. M., & Kanselaar, G. (2006, November). *Developing a 'big picture': Effects of collaborative construction of external representations in history.* Paper presented at the research meeting of the Interuniversity Centre for Educational Research (ICO), Amsterdam.

Prangsma, M. E., Van Boxtel, C. A. M., Kanselaar, G., & Kirschner, P. A. (2006). Concrete and abstract representations in history learning tasks. *Proceedings of the EARLI SIG Text and Graphics Comprehension*, Nottingham University, UK, 68-70.

Prangsma, M. E., Van Boxtel, C. A. M., Kanselaar, G., & Kirschner, P. A. (2006). Interactief geschiedenis leren met visuele representaties: Een blik op processen [Interactive history learning with visual representations: A look at processes]. *Proceedings of the 33rd OnderwijsResearchdagen* (pp. 99-100). Amsterdam: Vrije Universiteit Amsterdam.

Prangsma, M. E., Van Boxtel, C. A. M., & Kanselaar, G. (2005). Developing a ‘big picture': Collaborative construction of multi-modal representations in history [Abstract]. In C. P. Constantinou, D. Demetriou, A. Evagorou, M. Evagorou, A. Kofteros, M. Michael, C. Nicolaou, D. Papademetriou, & N. Papadouris (Eds.), *Integrating multiple perspectives on effective learning environments: Proceedings of the 11th European Conference for Research on Learning and Instruction* (p. 296). Nicosia, Cyprus: University of Cyprus.

Erkens, G., Jaspers, J. G. M., Prangsma, M. E., & Kanselaar, G. (2003, April). *Coordination processes in computer supported collaborative writing.* Paper presented at the invited VOR-ICT symposium of the 2003 AERA Annual Meeting, Chicago, Illinois.

Prangsma, M. E., Erkens, G., Jaspers, J. G. M., & Kanselaar, G. (2002). Computerondersteund samenwerken bij argumentatieve tekstproductie [Computer-supported collaborative writing]. In F. Daems, R. Rymenans, & G. Rogiest (Eds.), *Onderwijsonderzoek in Nederland en Vlaanderen*. *Proceedings of the 29th OnderwijsResearchdagen*. (pp. 533-535). Antwerpen: Universiteit Antwerpen.

Theil, J., Erkens, G., Prangsma, M. E., & Jaspers, J. G. M. (2002). "W8 ff :-)": 'Chattalk' in een coöperatieve leeromgeving ["W8 JAS :-)": Chat talk in a collaborative learning environment]. In F. Daems, R. Rymenans, & G. Rogiest (Eds.), *Onderwijsonderzoek in Nederland en Vlaanderen*. *Proceedings of the 29th OnderwijsResearchdagen* (pp. 543-545). Antwerpen: Universiteit Antwerpen.

Prangsma, M. E., Erkens, G., Jaspers, J. G. M., & Kanselaar, G. (2002, July). *Computer support of linearization in collaborative writing.* Paper presented at the 8th International EARLI SIG Writing Conference, Stafford, UK.

Erkens, G., Prangsma, M. E., & Jaspers, J. G. M. (2002, May). *Planning and coordinating activities in collaborative learning.* Paper presented at the RISE Conference on Collaborative Learning, Reasoning, and Technology, Newark, NJ.

Curriculum vitae

Maaike Prangsma was born on 22nd June 1974 in Apeldoorn, The Netherlands. She graduated in 1999 from Groningen University in English language and linguistics and joined Utrecht University in 2000 as a junior researcher. Between 2000 and 2002, Maaike worked on two projects investigating computer-supported collaborative learning, the second of which within the domain of history learning – also the subject of her PhD project. Maaike began her PhD in 2002, completing her manuscript precisely according to her planning in March 2007. Friends and family expected no less as Maaike also runs her own company specialising in time-management advice and coaching.

Acknowledgements

There is not enough room here to thank everyone who contributed to the completion of this dissertation, and I will undoubtedly have missed out a few names. If you cannot find your name here, please, accept my apologies and feel free to jot in your name in the appropriate category.

My husband

Matthew

Supervisors and coaches

Gellof	Carla	Paul
Jeanet	Marieke	Esther
Ike		Carla

Family and friends

Pap	Mam	Jord
Martine	Pieter-Anne	Sjoerd
Kees	Margreet	David
Anne-Marie	Jérôme	Lis
Johannes	Markus	Daniëlle
Opa		Tante Lies

Colleagues

Bert	Chiel	Chris	Clari	Co	Crina	Daniëlle
Dortie	Eline	Ellie	Elly	Erica	Frans	Frieda
Gijsbert	Hans	Hein	Hendrien	Jan	Jannet	Jantine
Jentine	Jeroen	Jessica	Joris	Jos	Karel	Katrien
Lennart	Liesbeth	Lisette	Maarten	Marcel	Marieke	Marije
Mirjam	Patrick	Richard	Ronny	Sandy	Sjaak	Theo
	Thoni	Tim	Wim	Wouter	Yvette	

Participants

Jaap	Jerry	Koen	Sheila	Thea

Pupils and board at Fioretti College	Pupils and board at Vader Rijn College	Pupils and board at Kalsbeek College	Pupils and board at Meridiaan College

Assistants

Erica	Johan	Hanna
Maaike	Liesbeth	Sophie
Corine	Joanneke	Karin

Lijn 296

Floor	Koen	Gonny
Annemieke	Lineke	Henk
George		Petra