CHB 611
6 November 2004; Disk Used
ARTICLE IN PRESS
No. of Pages 20, DTD = 5.0.1

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Computers in Human Behavior xxx (2004) xxx–xxx
Computers in Human Behavior
www.elsevier.com/locate/comphumbeh

Multimedia learning in social sciences: limitations of external graphical representations

Katrien De Westelinck [a,*], Martin Valcke [a,*]
Brigitte Decraene [b], Paul Kirschner [b]

[a] *Department of Education, Ghent University, H. Dunantlaan 2, 9000 Gent, Belgium*
[b] *Educational Technology Expertise Center, Open University of the Netherlands, Netherlands*

Abstract

In a series of six experimental studies, each consisting of three sub-studies, the central question was researched whether adding external graphical representations to printed or electronic learning materials improves retention and transfer scores. These studies research the degree of generalizability of Mayer's cognitive theory of multimedia learning (CTML) to the knowledge domain of the social sciences. The research hypotheses build on the assumption that this knowledge domain differs in the way instructional designers are able to develop adequate depictive external graphical representations. Earlier CTML-research was mostly carried out in the field of the natural sciences, where graphical representations are depictive in nature and/or where representations can be developed from existing or acquired iconic sign systems. The results indicate that alternative guidelines might need to be considered when learners study learning materials with external graphical representations that reflect low levels of repleteness and do not build on an iconic sign system previously mastered or acquired by the learners. The research results reveal that studying this type of representation does not result in higher test performance and does not result in lower levels of mental load.

* Corresponding authors. Tel.: +32 9 264 8675; fax: +32 9 264 8688.
E-mail addresses: katrien.dewestelinck@ugent.be (K.D. Westelinck), martin.valcke@ugent.be (M. Valcke).

0747-5632/$ - see front matter
doi:10.1016/j.chb.2004.10.030

1. Introduction

The cognitive theory of multimedia learning (CTML) posited by Mayer (2001a) presents a clear framework to direct instructional design of both printed and interactive multimedia materials. The power of CTML and these design guidelines is not only linked to a clear theoretical base, but also builds on the empirical evidence presented by Mayer, his colleagues, and other researchers. Consequently, instructional designers find the theory theoretical and practical appealing. But daily teaching experience of the authors of the present article, responsible for freshman courses in the knowledge domain of educational sciences, is not in line with CTML. Students appear to have difficulties in coping with graphical representations such as schemas, tables and graphs. And, as will be discussed in the next sections, recent research is not always able to replicate the positive findings that have been reported in earlier CTML-studies in other knowledge domains.

Through testing the CTML-design principles in another subject domain the question of extending or generalizing the cognitive theory of multimedia learning is raised. Printed and computer multimedia learning materials are used to test the original CTML-based research hypotheses, but do this in the context of alternative hypotheses that are put forward to explain results/expectations not completely in line with CTML design principles.

2. Basic assumptions and design guidelines of CTML

Mayer's theory of multimedia learning (2001a, 2003) is based on three central assumptions. The *dual channel assumption* states that two separate channels are used

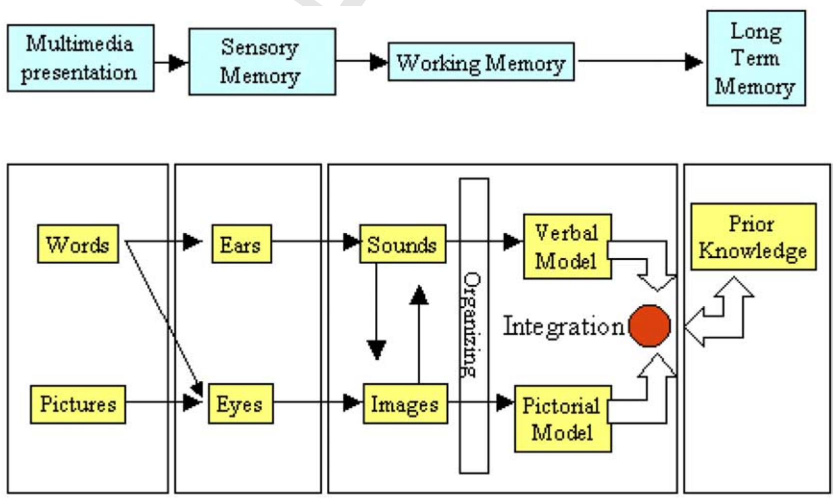

Fig. 1. The cognitive theory of multimedia learning (Mayer, 2001a,b, p. 44).

to process information (see Fig. 1). A first channel processes sounds in working memory, resulting in verbal models. A second channel is used to process images, resulting in pictorial models. The construction of both verbal and pictorial models can be influenced by prior knowledge retrieved from long term memory. Both are integrated into one coherent structure to be stored in long term memory. The second CTML-assumption focuses on the processing of all sensory input: the *active processing assumption*. This implies that the learner is actively engaged in processing information and makes an effort to construct coherent mental models. Typical cognitive processes involved in the latter are selecting, organizing and integrating. The third assumption is the *limited capacity assumption*. This implies that learners are limited in the amount of information they can process simultaneously along each channel.

These three theoretical assumptions are related to comparable notions in the literature. The dual channel assumption is also found in the working memory model of Baddeley (1992) and Chandler and Sweller (1991), the multiple channel communication model of Moore, Burton, and Myers (1996), the dual-coding theory of Paivio (1978, 1991), the sensory-semantic model of Nelson (1979) and the multiple-channel communication theory of Broadbent (1956), Shannon and Weaver (1949) and others. The second assumption about limited capacity is related to the "cognitive load theory" (CLT) of Sweller and colleagues (1988, 1989, 1994) who also tried to describe and explain the difficulties learners meet when dealing with complex knowledge domains. The active processing assumption is central to most cognitive theories and is, for example, explicitly mentioned by Wittrock (1989).

The practical relevance of CTML is evidenced by the definition of design principles for multimedia learning materials and is as such most clearly directed towards the instructional designer community (Reimann, 2003). The design guidelines (formulated as stated in the book Multimedia Learning by Mayer) are applicable to printed and interactive multimedia learning materials: (a) the *multimedia principle*: learners benefit more from words and pictures than from words alone, (b) the *temporal contiguity principle*: learners perform better when corresponding words and pictures are presented in close temporal proximity (e.g., simultaneously) instead of successively, (c) the *spatial contiguity principle*: learning is fostered when words and pictures are represented close to one another on a page or screen, (d) the *coherence principle*: learning performance is better when extraneous sounds, words, pictures are excluded, (e) the *modality principle*: learners learn more from animation enriched with audio (narration) than from animation enriched with printed text, (f) the *redundancy principle*: learners perform better when presented with animation and narration instead of animation and narration combined with printed text when the printed text matches the narration, and (g) the *individual differences principle*: all design principles have a stronger impact with low-prior knowledge learners and learners with high-spatial abilities (see Mayer, 2001a, 2001b, 2003 for an overview). Next there is the phenomenon called *expertise reversal*: what is optimal for low prior knowledge learners is suboptimal for experts and vice versa (Kalyuga, Ayres, Chandler, & Sweller, 2003) Mayer stresses the generic nature of these design guidelines (2001a, p. 193). He states that they can help to explain why instructional designers, such as Tufte (1983, 1990) stressed to enrich text with graphical representations such as tables, graphs, diagrams and charts. The research here questions the generic nature of the design guidelines by focusing on some problems related to the nature of external graphical representations in a particular knowledge domain.

3. Nature and impact of types of external graphical representations

There is a long tradition in theoretical and empirical research about external graphical representations in learning materials (see Anglin, Towers, & Levie, 1996 for an overview). This article focuses in particular on the CTML to study the theoretical and empirical impact of external graphical representations in learning materials. Although CTML-research has given a lot of proof that using the design principles developing learning materials result in higher performance on retention- and transfer tests recent CTML-related research presents inconsistent results about the impact on student performance. Goldman (2003), in a recent review of external representations studies, asks in this context for a *second generation* of research. She considers Mayer's work as *first generation* research focusing on generic principles to understand consistencies in the processing of verbal and visual information. The *second generation* should be helpful for understanding the affordances of external graphical representations in view of task demands, the active processing of learners, the support learners receive in processing the learning materials and low or high prior knowledge. The research presented here is a contribution to this second generation since it focuses on the affordances of external graphical representations in view of the active processing by learners in a specific domain. It especially questions whether learners are sufficiently acquainted with the basis of the symbol system as reflected in external graphical representations. The question is also related to the nature of knowledge domains.

Mayer differentiates between verbal and pictorial representations, noting that verbal representations require more mental effort to be processed by the learner. Pictorial representations are considered more original modes of knowledge representation. Mayer (2001a, p. 68) states that pictorial representations are more intuitive and closer to visual experience. Presenting both text and pictures invokes deep learning because the learner is required to develop both verbal and pictorial mental representations and connections between them.

Schnotz and Bannert (2003) elaborated on this theoretical distinction between verbal (descriptive) and pictorial (depictive) representations in an alternative way. In their view, *descriptive* representations such as text, formulae or logical expressions build on the use of symbols related to content via conventions. An important part of the symbol system is used to reflect relationships between the symbols (e.g., verbs and prepositions). Of importance for the present study is that such descriptive representations like printed text on paper or a screen can build on available and/or acquired *iconic sign systems*. Goodman (1976) notes that *depictive* representations such as pictures, graphics, or sculptures do not build on such iconic sign systems. Each type of depictive representation possesses inherent structural features that have very specific associations with the content represented. The example in Fig. 2 demonstrates how a learner has to interpret that the arrows to the left and right of the bus indicate the distance is to and from the two destinations. In other words, the learner has to know or learn and understand these associations between the structural features of the representation and the content represented. In this example an iconic sign system is available to understand a part of the representation (i.e., what 50 or 200 km means), but to understand the specific meaning of the arrows, most learners will have to rely on prior knowledge to assign the meaning "from" and "to". Most learners will interpret this part of the depictive representation analogously and also that the tower is the Eiffel tower in Paris. As to the meaning of the woman with the child, there is room for multiple interpretations (e.g., friend, girlfriend, wife, family, mother or grandmother). In this example, alternative representations of this part of the representation will not result in a lack of understanding of the overall content of the representation.

The fact that learners need to be acquainted with the iconic sign system used to develop an external graphical representation is the core of this study. Mayer, as explained earlier, would state that the depictive representation of the Eiffel tower and the woman with child are more intuitive and closer to visual experience than the descriptive representations of distances and directions. Learners are expected to process these depictive representations much faster than they would the descriptive ones. In other words, the learner builds a pictorial model with the correct visual-perceptual relationships. At question here is whether learners have sufficient and adequate prior knowledge to understand the depictive representations. The implication is that prior knowledge influences mastery of the iconic sign system at the base of the representations and that learners could have more difficulties and/or need more time to develop mental models when confronted with new or unknown iconic sign systems. There can, in other words, be a mismatch between the iconic sign system of a learner and the iconic sign system used in the representations, which can cause learners to experience more difficulties and/or need more time to develop mental models when confronted with new or unknown iconic sign systems. Goodman (1976) calls this a low level of *repleteness*, an index of the number of elements that are significant for the learner. Low repleteness implies a limited similarity to the realistic representation, which in turn implies a high cognitive load when confronted with such depictions and thus little space for learning processes. If this is the case, the benefits of adding graphical representations to achieve meaningful learning,

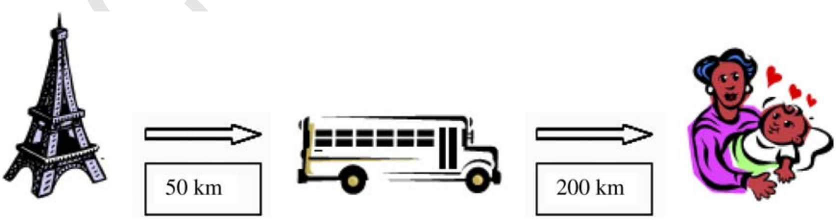

Fig. 2. Example of a depictive external representation.

which are typical for Mayer's studies, may not be found here. Stenning (1999) and Dobson (1999) qualify this via the variable *expressiveness*. Lower levels of expressiveness lead to more room for interpretations. Lowe (2003), for example, indicates that novices are easily captivated by the perceptually salient features of the displays and miss in this way the underlying principles and relationships. Stern, Aprea, and Ebner (2003) come to comparable conclusions finding that students who do not understand the fundamental concepts of graphs are prevented form noticing the key relationships in them. Also Lewalter (2003) points to the critical problem of students who do not succeed in identifying relevant information presented in external graphical representations. Consequently, Goldman (2003, p. 240) stresses the fact that representations "are only successful in improving learning from text to the degree that learners are able to interpret the cues". Mayer and Gallini (1990) indicate, for example, that learners might experience difficulties in identifying the relevant information presented in an illustration.

A review of the research literature from the perspective of iconic sign sytems reveals two important issues. First, there are *inconsistencies* in the way external graphical representations have been studied. Not all the studies make use of depictive external graphical representations (see Table 1). Mayer's original studies of (2001a, 2003) about lightning, pumps, and brakes are clear examples of depictive studies. But other studies, however, focus on more descriptive since they build on the use of symbols related to content by means of convention. These studies add external graphical representations such as flowcharts, formula editors, mathematical symbol sets, chemical formulas, and chemical reaction representations. This may be the source of inconsistencies in the findings of these studies about the CTML-guidelines. Second, most studies have been set up in the natural sciences. But, knowledge domains differ in their use of iconic sign systems. Recent CTML-studies set up in other knowledge domains can provide a significant extension of CTML.

The central hypothesis of the present research is that learners in the social sciences will experience difficulties with depictive external graphical representations as opposed to descriptive external graphical representations (e.g., text), due to interpretation difficulties of the iconic sign system used to develop these representations. Whereas the natural sciences can more easily build on intuitive (or acquired) consensual graphical representations, this is less apparent in the social sciences. These difficulties are expected to affect selection, processing and organizational processes of the learners. Due to less unequivocal (i.e., unambiguous) external graphical representations and the less known or unfamiliar iconic sign systems used, students are more likely to experience higher cognitive load. As a result of this increased cognitive load learners will develop less effective mental models and the deep-level learning predicted by Mayer, will hardly occur. Consequently, retention and/or transfer is expected to equivalent or lower than when the depictions are absent. If this is the case, then CTML design guidelines might be extended by taking the nature of the knowledge domain and/or the mastery of iconic sign system by learners into account.

Table 1
The knowledge domain and type of external graphical representations in CTML-research

Topic/knowledge domain	Study	External representation approach
Original studies of Mayer		
Pumps	Mayer and Anderson (1991)	Depictive with high repleteness: step-by-step drawings of a pump in different states
Brakes	Mayer and Anderson (1992)	Depictive with high repleteness: step-by-step drawings of brakes in different states
Lightning	Mayer et al. (1996)	Depictive with high repleteness: step-by-step drawings and animations
Generators	Mayer and Gallini (1990)	Depictive with high repleteness: step-by-step drawings of generators in different states
Lungs	Mayer and Sims (1994)	Depictive with high repleteness: step-by-step drawings of lungs indifferent states
Recent CTML-studies hard sciences		
Soldering	Kalyuga et al. (1999)	Depictive with high repleteness: videos of soldering workmen
Chemistry	Kozma (2003)	Descriptive (chemical formula) and depictive (set-up of chemical experiment) of process
Ecology	Roth and Bowen (1999)	Descriptive: Cartesian graphs representing cause-effects
Machines	Hegarty and Just (1993)	Depictive with high repleteness of machine functions
Vitamines & minerals	Seufert (2003)	Depictive with chemical set-up and chemical elements in the process
Meteorology	Lowe (2003)	Descriptive: meteorological maps in different states
Geographical time differences	Schnotz and Bannert (2003)[a]	Descriptive and depictive with low repleteness: carpet and circle diagrams
Recent CTML-studies in other field of sciences		
Training program for "experimental research"	Tabbers et al. (in press[a])	Depictive with low repleteness: diagrams
Introduction to instructional design	Tabbers et al. (in press[b])	Depictive with low repleteness: diagrams
Financial decision making	Stern et al. (2003)	Descriptive and depictive with low repleteness: mathematical graphs
First order logic	Dobson (1999)	Depictive (high and low repleteness: 3D-pictures versus diagrams) and descriptive
First order logic	Stenning (2003)	Descriptive (logical expressions) and depictive with low repleteness: logic tables
First order logic	Dobson (1995)	Depictive with low repleteness: Venn & Euler representations

4. Research

In a series of six separate experiments the basic tenets of CTML were tested as to their validity in the social sciences and how CTML might be extended.

5. Methods

5.1. Participants

In total 190 freshmen studying educational sciences at a Flemish university participated in this study. They represent the entire population of first-year students in the second semester 2002–2003. Participation was a formal part of the course "Instructional Sciences". Informed consent was obtained from all students prior to experimentation.

5.2. Procedure

The studies were set up during two sessions, organized during two consecutive weeks. Students were randomly assigned to the experimental conditions. The groups were formed by selecting the students as they appeared on the alphabetical tuition list. There were six experiments consisting of three sub-studies each focusing on a theme, related to the selected learning content (see Section 5.3). No students were assigned to the same condition in successive sessions. Each experimental condition was organized in a different room. Students, at the start of each session, received a study package consisting of: (a) a prior knowledge test, (b) a specific elaboration of the learning materials to be studied, and (c) a post-test of mastery of the complex knowledge elaborated in the learning materials (retention and transfer). After the second sub-study of each session, students were invited to indicate the cognitive load experienced during study. No time limit was set for studying the materials and/or completing the tests. The study package of students in computer conditions (i.e., to test the principles with dynamic representations) only consisted of pre-tests, cognitive load measures and the post-test for each sub theme in the session. Students in these conditions studied the multimedia materials in a computer room.

The answers to the retention and transfer questions were scored by three independent researchers not involved in the current study. The scoring was based on a scoring checklist that provided an optimal answer to each individual question. A score was given depending on the number of elements in a student's answer. To facilitate interpretation of the test scores, all scores were standardized, with a maximum score of 20 for each pre- and post-test.

5.3. Materials

The content of the learning materials was both complex and new to the students: an introduction to the learning styles literature (the learning content). Nine themes were outlined to be presented to the students: (a) the conceptual differentiation between behavior, mental activities, learning strategies and learning styles, (b) Curry's typology to differentiate between learning style as a personality trait, an information processing style or an instructional preference, (c) Dunn and Dunn's learning style approach, (d) Kolb's learning style approach, (e) Witkin's learning style model, and (f) Vermunt's learning style model. This learning content is complex and at a high difficulty level for freshman.

To guarantee the optimal design of the external graphical representations, Mayer's recommendations were taken into account (2001a, p. 191–193). He states that the graphical representations should have a potentially meaningful structure (a cause-effect relationship, interdependencies or hierarchies) and depict the different states of the complex structure. Building on these guidelines, the authors and a group of 20 fourth-year psychology students taking a course in instructional design, developed a series of possible graphical representations for each theme from which the authors selected and finalized the multimedia representations for each learning styles theme. Special care was taken when representing the structural relationships in the body of knowledge (such as dependent upon, consisting of, different from, follows from, affects, contains, et cetera). Fig. 3 depicts a page of printed learning materials with integrated external graphical representations about Kolb's learning style approach. It is clear from the example that the external graphical representations do not build on a formal and/or existing iconic sign system. Moreover, the approach is similar to the typical external graphical representations found in psychology and educational sciences textbooks.

For the design of the dynamic representations, computer animations were developed that were equivalent to those in the printed learning materials. The animations show, step by step, the build up of the representations incorporated in the printed materials. The students controlled the speed of the animations by clicking on the *continue* button on the screen.

5.4. Instruments

A pre-test and post-test were presented to the students which consisted of retention and transfer questions. Retention questions measure what students remember about a topic (e.g., What are the different operational approaches that Vermunt incorporates in his approach towards learning styles?). Transfer questions are related to problem solving. They test the deeper understanding of the content by having students explain phenomena that cannot immediately be retrieved from memory (e.g., What is the relationship between cognitive style and personality in Witkin's approach?). The analysis section reports the test results separately for each type of question, along with a total test score.

In the literature, measurement of cognitive load is mainly based on the learners' subjective report of their perceived *mental effort*. This results in a subjective cognitive load scale (Paas, van Merrienboer, & Adam, 1994) in which students note the amount of effort they experienced on a scale varying from 1 (very, very, very easy)

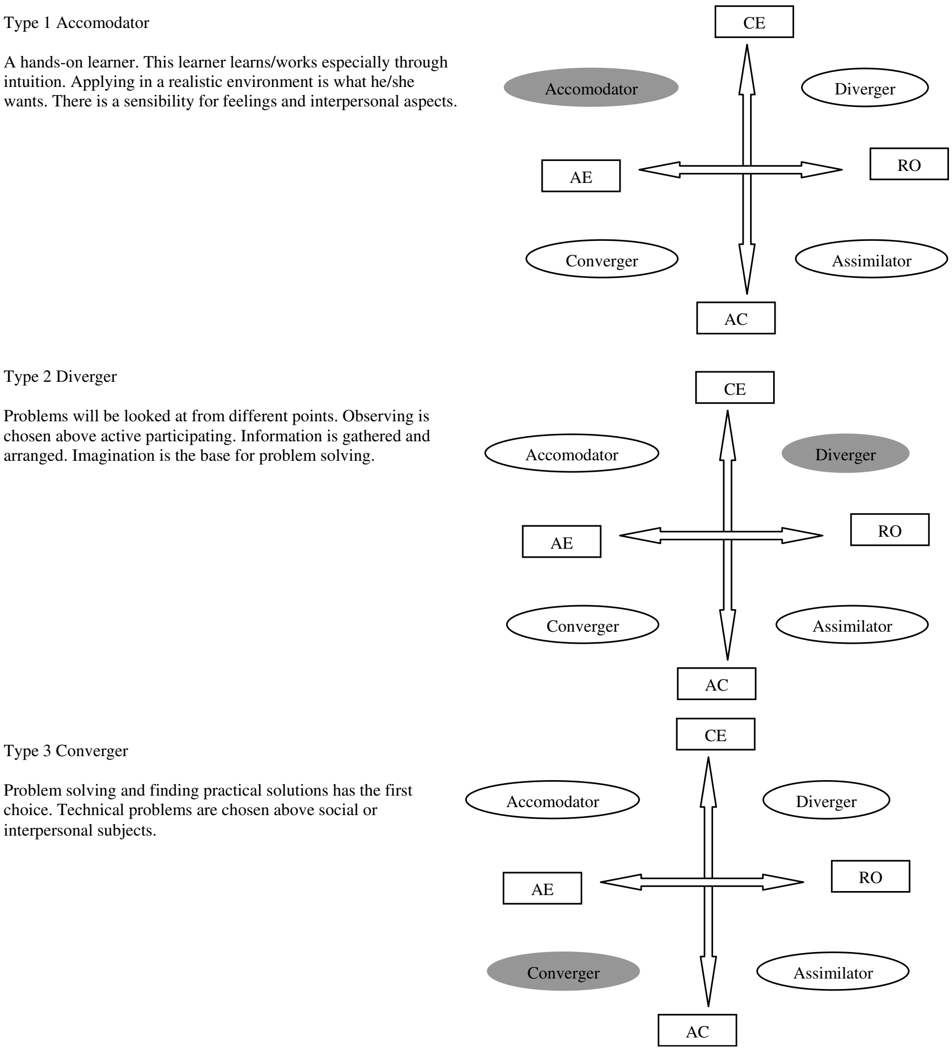

Fig. 3. Example of learning materials to test hypothesis three related to the coherence principle.

to 9 (very, very, very difficult). Application of this kind of scale results in high reliability measures (Cronbach's α) of .90–.82 (Paas, 1992; Paas et al., 1994).

5.5. Statistical analysis

All analyses are based on the comparison of mean test scores of students in the different conditions. Analysis of variance is applied after testing for homogeneity of variances. A significance level of $p < .01$ is used as the critical value. In case of sta-

Table 2
Mean scores and standard deviation for the retention, transfer and total scores in each experiment and for each sub study

Central hypothesis in the experiment		Multimedia principle				Spatial contiguity principle				Coherence principle			
		Text without representations		Text with external representations		Representations not integrated		Integrated representations		Summaries with representations		Expanded with illustrations	
Session 1		*M*[a]	SD	*M*[b]	SD	*M*[e]	SD	*M*[f]	SD	na[i]		na	
Sub 1	Retention	17.50	5.27	19.43	2.01	20.00	0.00	20.00	0.00				
	Transfer	9.03	5.18	10.14	6.36	17.79	4.12	14.43	5.25				
	Total	13.26	4.34	14.78	3.34	18.89	2.06	17.21	2.63				
Sub 2	Retention	19.26	2.65	18.66	2.70	19.41	2.52	19.23	2.15				
	Transfer	6.80	3.61	3.14	2.99	7.20	3.30	6.86	4.55				
	Total	12.14	2.68	9.70	2.33	12.44	2.32	12.16	2.97				
Sub 3	Retention	7.78	7.60	6.00	8.12	15.88	6.57	15.42	6.57				
	Transfer	11.48	3.42	11.04	4.56	11.95	6.09	12.57	5.55				
	Total	10.00	3.88	9.03	4.58	13.52	5.02	13.71	4.78				
	Mental load	5.17	2.03	–[j]	–[j]	4.61	2.19	3.97	2.05				
Session 2		*M*[c]	SD	*M*[d]	SD	na		na		*M*[g]	SD	*M*[h]	SD
Sub 4	Retention	14.97	4.38	14.40	4.70					13.83	2.97	11.69	4.62
	Transfer	13.99	6.23	12.13	7.38					12.79	5.60	11.39	4.99
	Total	14.52	4.24	13.35	5.38					13.36	3.58	11.56	3.45
Sub 5	Retention	7.20	2.59	6.04	2.18					7.25	10.12	4.44	2.37
	Transfer	6.66	4.51	6.13	3.68					6.97	4.18	7.13	4.15
	Total	6.97	2.32	6.07	1.59					7.13	6.55	5.51	2.04
Sub 6	Retention	19.80	1.00	18.00	4.56					16.76	6.46	16.28	6.91
	Transfer	12.26	6.85	11.46	8.05					10.15	8.44	7.44	7.99
	Total	16.57	3.09	15.20	3.47					13.92	5.42	12.49	5.58
	Mental load	5.56	1.32	5.00	1.95					5.95	1.87	4.86	1.74

(*continued on next page*)

Table 2 (*continued*)

Central hypothesis in the experiment		Modality principle				Redundancy principle			
		Animation with narration		Animation with printed text		Animation with printed text and narration		Animation only with narration	
Session 1		M^{A}	SD	M^{B}	SD	na[E]		na	
Sub 1	Retention	19.04	2.46	19.80	1.00				
	*Transfer	11.92	6.33	11.40	4.90				
	Total	15.42	3.16	15.60	2.53				
Sub 2	Retention	19.23	2.88	18.93	3.15				
	Transfer	5.77	3.37	6.40	3.68				
	Total	11.54	2.62	11.77	2.52				
Sub 3	Retention	7.31	7.77	5.60	7.68				
	Transfer	10.25	5.41	11.99	3.33				
	Total	9.08	3.85	9.44	3.44				
	Mental load	4.58	1.74	4.30	1.89				
Session 2		na		na		M^{C}	SD	M^{D}	SD
Sub 4	Retention					13.14	3.98	14.34	3.93
	Transfer					8.40	5.28	7.05	4.25
	Total					10.95	3.47	10.98	2.38
Sub 5	Retention					4.44	2.31	4.62	2.58
	Transfer					3.86	3.92	5.89	5.10
	Total					4.21	2.09	5.12	1.54
Sub 6	Retention					16.00	1.65	15.77	7.20
	Transfer					7.73	5.86	7.95	9.98
	Total					12.46	8.96	12.42	6.19
	Mental load					5.68	1.65	5.61	2.19

[A] $N = 26$.
[B] $N = 25$.
[C] $N = 25$.
[D] $N = 26$.
[E] Not applicable. No experiments were set up to test this specific hypothesis during this session.
[a] $N = 36$.
[b] $N = 35$.
[c] $N = 25$.
[d] $N = 25$.
[e] $N = 34$.
[f] $N = 35$.
[g] $N = 44$.
[h] $N = 43$.
[i] Not applicable. No experiments were set up to test this specific hypothesis during this session.
[j] Due to a layout error in the package of the students for the condition with external representations, an insufficient number of students replied to the question to estimate their mental load.

tistically significant differences in mean post-test scores, Cohen's *d* is calculated to determine effect size (Thalheimer & Cook, 2002).

6. Results

6.1. General remarks

Table 2 summarizes the descriptive statistical analysis results, based on the test scores of students, in the different conditions. The results of the pre-test scores are not reported since all students obtained a zero-score for both the retention and transfer questions in this test. This clearly indicates that the knowledge content was completely new for the students and of a high difficulty level.

The value of the multimedia principle is tested twice via the experiments set up in both research sessions. The results in Table 2 are clear. With the exception of the post-test scores in relation to the first sub-study, students studying learning materials with no external graphical representations always attain a higher mean post-test score. Analysis of variance (see Table 3) reveals that these differences are significant for the second sub-study. The effect sizes are very large to large: $d = 1.12$ for the transfer test and $d = 0.95$ for the total post-test score in relation to the specific learning styles content.

The analysis of the descriptive results in relation to spatial contiguity shows that the majority of the conditions where illustrations are not spatially integrated result in higher post-test scores than when this is the case. The differences in scores for the transfer question and the total post-test in the first sub-study are significant. In both cases, this results in a medium effect size of $d = 0.72$.

Analysis of the results in relation to the coherence principle suggest that students studying learning materials consisting of summaries with external graphical representations perform better on post-test questions, though none of the differences are significant.

With respect to computer based (multimedia) learning materials the condition where animations are enriched with audio should, according to the modality principle, lead to higher performance than the condition where the animation is enriched with screen text. The descriptive results in the sub-studies do not support this, though none of the differences found are significant.

The post-test scores of students studying non-redundant learning materials, that is animation with narration and without additional text are mostly higher, but here too the differences are not significant.

Finally, since each of the conditions employed build on different applications of CTML-design guidelines it is possible to see whether there are differences in cognitive load in favor of CTML-designs. There were no significant differences, with the exception of conditions presenting alternative designs based on the coherence principle. The cognitive load for students studying the most coherent learning materials was significantly higher with a medium effect size of $d = 0.72$.

Table 3
Overview of ANOVA results

		Multimedia		Spatial contiguity		Coherence		Modalit y		Redundancy	
Session 1		*F*(1,69)	*p*	*F*(1,67)	*p*	na[a]		*F*(1,49)	*p*	na	
Sub 1	Retention	4.09	.05	–[c]	–[c]			2.07	.16		
	Transfer	.69	.42	8.74	**.004**[*]			.11	.74		
	Total	2.73	.10	8.74	**.004**[*]			.02	.88		
Sub 2	Retention	.87	.35	.09	.76			.12	.73		
	Transfer	21.56	**.00**[*]	.13	.72			.41	.53		
	Total	15.49	**.00**[*]	.18	.67			.10	.75		
Sub 3	Retention	.91	.34	.08	.77			.62	.43		
	Transfer	.21	.65	.19	.66			1.90	.17		
	Total	.93	.34	.02	.88			.13	.72		
	Mental load	–[b]	–[b]	1.60	.21			.28	.60		
Session 2		*F*(1,48)	*p*	na		*F*(1,85)	*p*	na		*F*(1,49)	*p*
Sub 4	Retention	.20	.66			6.60	.02			1.17	.28
	Transfer	.93	.34			1.52	.22			1.01	.32
	Total	.73	.40			5.70	.02			.00	.98
Sub 5	Retention	2.92	.09			3.13	.08			.06	.80
	Transfer	.21	.65			.03	.86			2.52	.12
	Total	2.60	.11			2.38	.13			3.17	.08
Sub 6	Retention	3.71	.06			.11	.74			0.2	.90
	Transfer	.14	.71			2.36	.13			.01	.94
	Total	2.18	.15			1.49	.23			.00	.98
	Mental load	1.40	.24			7.99	**.006**[*]			.01	.91

[a] Not applicable. No experiments were set up to test this specific hypothesis during this session.
[b] Due to a layout error in the package of the students for the condition with external representations, an insufficient replied to the question to estimate their mental load.
[c] Since students in both conditions obtain the maximum score for the retention question in relation to this first sub study, no *F*-value can be calculated.
[*] $p < .01$.

7. Discussion

The results of the studies presented here do not present an unequivocal answer to the question of CTML design guidelines are generalizable to different domains. On the one hand, the results raise serious questions (i.e., statistically significant differences in the non-CTML direction) by some of the assumptions of CTML design guidelines, especially those based on the multimedia, spatial contiguity, and modality principles. On the other hand, the lack of significant positive results in line with the CTML-assumptions opens the door to alternative explanations.

One noteworthy result was the significant differences in post-test scores indicating that studying text without external graphical representations sometimes results in higher performance. This is clearly in contrast with the original CTML-hypothesis and suggests that learners have problems when studying from external graphical representations because of inadequate experience with or knowledge of the iconic symbol system used. Support for this can be found in a number of research studies. Cox (1999), for example, states that the impact of graphical versus textual representations might be affected by the degree to which learners' understand the semantics of the representational system. This is also consistent with the findings of Lowe (2003), namely that subjects best extract information from representations where there are clear visual-spatial characteristics, such as structural coherence and distinctive appearance (e.g., closely related to reality). They do not extract information from representations that lack these qualities. He concludes in a study of learning meteorology from weather maps that students do not extract the elements of major meteorological importance from weather maps; knowledge structures (mental models) are "likely to be incomplete, fragmentary and of limited value in building high-quality mental models of weather map dynamics" (p. 174). Support is also found in Schnotz and Bannert (2003) who conclude that adding pictures to text is not generally beneficial, and that it can even have negative effects on learning because they may interfere with the construction of mental models. Finally, Dobson (1999) found that the impact of representations is influenced by the difficulties the students have to interpret the diagrams. He also determined that students actually prefer lexical parts in the learning materials as compared to diagram-representations.

A specific result was the fact that spatially contiguous integration of external graphical representations in printed learning materials does not result in higher post-test scores as compared to learning materials with non-contiguous representations. In both conditions, students apparently experience difficulties with the specific depictive representations. The contiguity of the representations to the text appears to hinder the students whereas in the non-contiguous conditions they can focus on a consistent textual (sentential) representation.

The effect of different aspects of the impact of representations on cognitive load could be tested in five of the experiments. At the descriptive level, there are only small differences in reported cognitive load by the students in the different conditions with a significant difference in only one condition, namely that students studying the more coherent learning materials experience higher cognitive load – a finding that is clearly not in line with CTML-based theory. Tabbers, Martens, and van Merriënboer (in press[b]) also report inconsistent results as to the impact of external graphical representations on cognitive load.

The practical implications of these findings are clear. Instructional designers may not be able to simply "apply" CTML-guidelines to learning materials in a knowledge domain, where no unequivocal iconic sign system is available to students and where representations have a low level of repleteness. This does not imply that the use of the CTML design principles is not recommended, but rather that caution is prescribed in other domains.

8. Methodological issues

A number of methodological questions can be raised in relation to the experiments in this research. A first question focuses on the quality of the external graphical representations: Are the results due to poor external graphical representations? Much time and effort was invested in the design of the representations by a large team and the representations can be considered to be typical for those found in textbooks in the educational sciences. Also, all representations took student task-demands into account. The structure of the six learning style themes were clearly and explicitly depicted or animated in the representations and specific post-test questions also focused on these features. This is important since recent studies (e.g., Schnotz & Bannert, 2003) have proven that non task-appropriate representations do not foster comprehension and mental model construction.

A second methodological point is that CTML-studies of Mayer and his colleagues is almost always of very short duration. Learning processes limited to 180 s are more the norm than the exception. In the present studies, larger chunks of learning content had to be processed by the students, during a longer period of time, so it is possible that the study tasks in the current study were more demanding than in Mayer's studies. Tabbers, Martens, and van Merriënboer (in press[a]) also mention this particular divergence between their studies and Mayer's as a potential source of inconsistency. In the context of a follow-up study, more attention could be paid to monitoring the study time as co-variable.

A critical issue is the fact individual differences were not taken into account. Since the research group was very homogeneous in terms of prior knowledge, it did not seem useful to take this into account. The intention was to make this an issue for future research. Mayer's seventh principle (2001a) refers to the impact of prior knowledge and spatial abilities. Recent research by Cox (1999) (p. 356) reveals that "there are large variations between subjects in the types and modalities of external graphical representations that they use in their solutions". He concludes that external graphical representations might serve different cognitive functions for different subjects. In addition to prior knowledge other variables such as learning styles or spatial abilities can help explain the research results.

Time on task is an important factor in a lot of researches and analyses. This research had, as said in the part materials, no time limit; students could work as long as they wanted on their material. The variable time was not included in this research, but will be taken into account in future researches.

9. Implications for instructional design and future research

The central research hypothesis of this study questions the generic nature of the design guidelines derived from CTML. The results suggest that instructional designers need more carefully consider the nature of the depictive representations they add to their learning materials. In the context of the present study, the focus was upon the educational sciences knowledge domain. This knowledge domain cannot be compared to the natural sciences where it is easier to build up depictive representations with high levels of repleteness. The results of the present study suggest that developers of learning materials pay explicit attention to *repleteness* as a central quality of the representations. Second, they could either design the representations in such a way that it would help learners understand the symbol system used, or they could ask students to develop representations themselves. Van der Pal and Eysink (1999) suggest an additional approach, namely building up a specific formal language that learners have to master in order to build graphical representations.

Considering the methodological remarks and the implications for instructional design, key characteristics of future research can be delineated. Future research should take into account extra co-variables related to individual differences between learners. A number of new research conditions could be included in the studies to contrast students that study learning materials enriched with external graphical representations and receiving or not receiving extra help, with or without prior introduction about/training in the iconic sign system used or in the design of their own representations of the learning content. This last idea could be expanded with groups being supported with the new generations of CSCL-environments in which specific representation tools are available.

In other words, a second generation of CTML-research is needed that considers the unique affordances of graphical representations in relation to their active processing by learners.

10. Uncited reference

Tabbers, Martens, and van Merriënboer (2001).

Acknowledgement

This study was partially funded by the Flemish Government under the BOF program, NR. 011D00103.

References

Anglin, G., Towers, R., & Levie, W. (1996). Visual message design and learning: the role of static and dynamic illustrations. In D. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 755–794). London: Macmillan.

Baddeley, A. D. (1992). Working memory. *Science, 255*, 556–559.

Broadbent, D. (1956). Successive responses to simultaneous stimuli. *Quarterly Journal of Experimental Psychology, 8*, 145–152.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8*, 293–332.

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

Dobson, M. (1995). Predicting learning outcomes through prior quantitative and post qualitative analysis of interactive graphical representation systems. *Paper presented at the conference of the European association for research on learning and instruction*, Nijmegen, The Netherlands.

Dobson, M. (1999). Information enforcement and learning with interactive graphical systems. *Learning and Instruction, 9*, 365–390.

Goldman, S. (2003). Learning in complex domains: when and why do multiple representations help?. *Learning and Instruction, 13*, 239–244.

Goodman, N. (1976). *Languages of art: an approach to a theory of symbols*. Indianapolis, IN: Hacket.

Hegarty, M., & Just, M. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory & Language, 32*(6), 717–742.

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist, 38*(1), 23–31.

Kalyuga, S., Chandler, P., & Sweller, P. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology, 13*, 351–372.

Kozma, R. (2003). The material features of multiple representations and their cognitive and social affordances for science understanding. *Learning and Instruction, 13*, 205–226.

Lewalter, D. (2003). Cognitive strategies for learning from static and dynamic visuals. *Learning and Instruction, 13*, 177–189.

Lowe, R. K. (2003). Animation and learning: selective processing of information in dynamic graphics. *Learning and Instruction, 13*, 157–176.

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

Mayer, R. E. (2001b). Cognitive, metacognitive and motivational aspects of problem solving. In H. Hartman (Ed.), *Metacognition in learning and instruction* (pp. 87–102). Dordrecht, The Netherlands: Kluwer Academic Press.

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., & Anderson, R. (1991). Animations need narrations: an experimental test of a dual-coding hypothesis. *Journal of Educational Psychology, 83*, 484–490.

Mayer, R. E., & Anderson, R. (1992). The instructive animation: helping students build connections between words and pictures in multimedia learning. *Journal of Educational Psychology, 84*, 444–452.

Mayer, R. E., Bove, W., Bryman, A., Mars, R., & Tapangco, L. (1996). When less is more: meaningful learning from visual and verbal summaries of science textbook lessons. *Journal of Educational Psychology, 88*, 64–73.

Mayer, R. E., & Gallini, J. (1990). When is an illustration worth a thousand words?. *Journal of Educational Psychology, 82*, 715–726.

Mayer, R. E., & Sims, V. (1994). For whom is a picture worth a thousand words. Extensions of a dual-coding theory on multimedia learning. *Journal of Educational Psychology, 84*, 389–401.

Moore, D. M., Burton, J. K., & Myers, R. J. (1996). Multiple channel communication: the theoretical foundations of multimedia. In D. H. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 851–878). New York: Macmillan.

Nelson, D. L. (1979). Remembering pictures and words: appearance, significance and name. In L. S. Cernack & F. Craik (Eds.), *Levels of processing in human memory* (pp. 45–76). Hillsdale, NJ: Erlbaum.

Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: a cognitive load approach. *Journal of Educational Psychology, 84*, 429–434.

Paas, F., van Merrienboer, J., & Adam, J. (1994). Measurement of cognitive load in instructional research. *Perceptual and Motor Skills, 79*, 419–430.

Paivio, A. (1978). A dual coding approach to perception and cognition. In J. H. Pick & E. Saltzman (Eds.), *Modes of perceiving and processing information* (pp. 39–51). Hillsdale, NJ: Erlbaum.

Paivio, A. (1991). Dual coding theory: retrospect and current status. *Canadian Journal of Psychology, 45*, 255–287.

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

Roth, W.-M., & Bowen, G. M. (1999). Complexities of graphical representations during ecology lectures: an analysis rooted in semiotics and hermeneutic phenomenology. *Learning and Instruction, 9*, 238–255.

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

Seufert, T. (2003). Supporting coherence formation in learning from multiple representations. *Learning and Instruction, 13*, 227–237.

Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of communication*. Urbana, IL: University of Illinois Press.

Stenning, K. (1999). The cognitive consequences of modality assignment for educational communication: the picture in logic teaching. *Learning and Instruction, 9*, 391–410.

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.

Sweller, J. (1988). Cognitive load during problem solving: effects on learning. *Cognitive Science, 2*, 257–285.

Sweller, J. (1989). Cognitive technology: Some procedures for facilitating learning and problem solving in mathematics and science. *Journal of Educational Psychology, 81*, 463–474.

Sweller, J. (1994). Cognitive load theory, learning difficulty and instructional design. *Learning and Instruction, 4*, 295–312.

Tabbers, H., Martens, R., & van Merriënboer, J. (2001). The modality effect in multimedia instructions. In J. D. Moore & K. Stenning (Eds.), *Proceedings of the 23rd annual conference of the cognitive science society* (pp. 1024–1029). Mahwah, NJ: Erlbaum.

Tabbers, H., Martens, R. & van Merriënboer, J. (in press[a]). The interaction of modality with pacing in multimedia learning. *British Journal of Educational Psychology*.

Tabbers, H., Martens, R. & van Merriënboer, J. (in press[b]). Multimedia instructions and cognitive load theory: Effects of modality and cueing. *British Journal of Educational Psychology*.

Thalheimer, W., Cook, S. (2002, August). How to calculate effect sizes from published research articles: A simplified methodology. Retrieved August 18, 2003 from http://work-learning.com/effect_sizes.htm.

Tufte, E. (1983). *Envisioning information*. Cheshire, CT: Graphics Press.

Tufte, E. (1990). *The visual display of quantitative information*. Cheshire, CT: Graphics Press.

Van der Pal, J., & Eysink, T. (1999). Balancing situativity and formality: the importance of relating a formal language to interactive graphics in logic instruction. *Learning and Instruction, 9*, 327–341.

Wittrock, M. C. (1989). Generative processes of comprehension. *Educational Psychologist, 24*, 345–376.