Multimedia, comprehension and the psychology of learning: A review of four cognitive models

Céline Lemercier and André Tricot

Laboratory Work and Cognition
University of Toulouse
Allées Antonio Machado
31 058 Toulouse cedex – France
celine.lermercier@univ-tlse2.fr

The authors wish to thank John Sweller for providing comments on earlier versions of this chapter.

Introduction

While multimedia techniques have been introduced on a very widespread basis over the last 15 years, numerous empirical studies have helped bring our understanding of the effects of multimedia on comprehension and learning up to date. To account for these results, the cognitive psychology of learning has developed models designed to explain the effect of multimedia on comprehension and even psychological models of the comprehension of multimedia documents. On the basis of these accumulated empirical results, we can ask (a) whether these explanatory models are *ad hoc* in nature or whether they dovetail easily with current-day psychological theories, and (b) whether these models contribute anything to the psychology of learning?

In the field of cognitive psychology, human beings who process and understand a multimedia document in order to learn something are often considered as individuals endowed with a natural information processing system. As such, they are, by definition, beings possessing perceptual, attentional and memory capacities. In this chapter, we shall attempt, on the one hand, to account for the incorporation, or lack of it, of the models of the comprehension of multimedia documents within the theories of attention and memory. On the other, we shall assess the contribution of these models to our knowledge of attentional and memory-related phenomena.

Here, we shall look at the four models so far recognized in the multimedia field as explaining the behavior of understanding, learning subjects: Sweller's cognitive load theory (2003), Mayer's cognitive theory of multimedia learning (2001, Moreno & Mayer, 1999), Schnotz's model of comprehension which integrates text and graphics (1993; Schnotz & Bannert, 2003) and, finally, the model of multidocument representation proposed by Perfetti, Rouet and Britt (1999). All these models have their roots in the psychology of learning and have given rise to articles in the leading journals dealing with this field of study: *Journal of Educational Psychology, Cognition and Instruction, Learning and Instruction.*

Sweller's and Mayer's models take the form of general theories of the processing of multimedia documents in learning situations or, in Sweller's case, even of a general theory of learning. These models have both a theoretical and a practical vocation: they constitute a general framework for the apprehension of learning and comprehension behavior which makes it possible to formulate working hypotheses and, thanks to the accumulated empirical and experimental results, they propose recommendations concerning the design of documents used for teaching purposes. For their part, the models advanced by Schnotz and by Perfetti et al. address more specific aspects of acknowledged models in the field of comprehension and mental representations. The main interest of these authors' works resides in the application of these models in the field of the integrative comprehension of text and graphics in the case of Schnotz (1993) and of documentary sources of differing formats and content for Perfetti et al. (1999). Schnotz thus discusses the hypothesis of a benefit due to dual coding developed by Paivio, whereas Perfetti et al. shed greater light on certain aspects of Kintsch's model.

Sweller (1999) proposes a first explanatory framework in order to account for the effect of multimedia on comprehension and learning. The interest of the cognitive load hypothesis advanced by this author lies in a simple proposition: sometimes, the completion of a learning task and the processing of the material that comprises it are so cognitively costly that if subjects devote all their cognitive resources to accomplishing the task, none are left over for learning and, consequently, they do not learn. Reducing the cost imposed by the accomplishment of the task, either by modifying the task or by modifying the material, frees up cognitive and/or attentional resources and thus permits learning to take place. Though it was developed before the emergence of multimedia, this theory has integrated the arrival of this new technique in a way that has considerably enriched and developed the theory. Recently, Sweller (2003) conferred a new dimension on the theory, making it simultaneously more profound and more general.

For his part, Mayer has developed a model of multimedia learning. Mayer (1997) believes that subjects learning in a multimedia setting mobilize three successive cognitive processes: selection (permitting the construction of the text and graphic base), organization (permitting the creation of models based on the textual and graphic material) and, finally, integration (permitting the establishment of connections between the corresponding events in the models based on the verbal and graphic material). These three stages in the processing of multimedia information in order to understand and learn call on attentional resources (1) for the selection of the relevant information and the inhibition of the irrelevant information and (2) for the sharing of resources between different sources of information, primarily during the stage of the selection of the presented information, as well as short-term and long-term memory resources, in particular during the stages of information organization and integration.

The model of the integrative comprehension of text and graphics proposed by Schnotz and Bannert (2003), which is to a very large extent inspired by Paivio's work on mental representations, has permitted us to gain a better understanding of the processes involved during the comprehension, by a human subject, of mixed text/graphic documents. Among other things, this model revisits the hypothesis that a benefit can be gained from the dual coding of the information, namely a hypothesis that was developed by Paivio during the 1980s in studies of the comprehension of documents which include both text and graphics.

As far as Perfetti, Rouet and Britt's model (1999) is concerned, it represents a response to the inability of Kintsch's model (1998) to account fully for the comprehension of multisource documents. In effect, Kintsch's model assumes that individuals develop a mental representation of the content of a text on the basis of the linguistic and propositional processing of the text, on the one hand, and the mobilization of his or her own knowledge on the other. Kintsch referred to this mental representation as a "situation model" which may or may not be faithful to the intentions of the author of the text. It differs from one reader to the next but remains consistent. Perfetti et al. base their reasoning on the fact that that when required to process complex multimedia documents, individuals are confronted by a particular problem of comprehension: they must process documents which are drawn from different sources and which may be contradictory, opposed etc. Perfetti et al.'s model (1999) accounts for the fact that comprehension may correspond to the elaboration of an inconsistent representation on the basis of sources which are themselves not consistent.

This chapter consists of two parts, one devoted to attention and the other to memory. Each part starts by recalling the key concepts relating to these two aspects of cognitive psychology. It then proceeds to evaluate the way in which each of the four models accounts for the concept in question or contributes additional information concerning its comprehension.

An individual who comprehends and learns is defined as an attentional being

A recall of the attention-related concepts

The perceptual stage refers to the entry of the stimuli into the information processing system. As far as multimedia as used in learning contexts (educational multimedia) is concerned, two main perceptual inputs are involved: visual input (graphics, text) and auditory input (sound, speech). It is at this level that the information to be processed is selected and that which is irrelevant for the activity in question is inhibited. In effect, while the number of sensory stimuli in our environment is infinite, our short-term information processing and storage capacities are limited (Miller, 1956). We must therefore make choices and select from within the environment those items that are the most relevant in the light of the goals we have set ourselves. At the cognitive level, the information is processed as a function of the received instructions (conceptual or spatial processing) and the subject's prior knowledge which is stored in long-term memory. Finally, the motor stage corresponds to the subject's actual response. As far as educational multimedia is concerned, the expected responses may be either manual or vocal. While the information processing theories expect the attentional load to be limited at all levels of stimulus processing (perceptual, cognitive, motor), we shall see below that the models which focus on the effect of multimedia on learning examine the concept of processing load only at the cognitive level. Although the perceptual and motor aspects of multimedia learning are referred to by these models, they do not form the object of direct experimental evaluation. Finally, certain principles, such as the multiple representation or cognitive overload principle which, in the information processing models, refer to perceptual aspects, are incorporated within the cognitive stage of information processing in both Sweller's and Mayer's models. It would be interesting at some future stage to distinguish between different forms of the processing overload effect depending on whether it is situated at the perceptual (visual/sound), cognitive (type of task, difficulty of the verbal and visual tasks) or motor (manual/vocal) level.

The instance responsible both for selecting the relevant information and inhibiting the irrelevant information (from the input of the stimulus into the information processing system through to the output) and which is considered to be the fuel or resource called on by the information processing system is attention (Broadbent, 1958). Conventionally, we distinguish between different forms of attention when referring to different current situations. Here, we shall discuss only the two forms of attention that are relevant to multimedia learning: selective attention and divided attention (Wickens, 1989). First of all, selective attention requires us to select those items which will enable us to succeed in our task from within an information-rich environment. For its part, divided attention refers more to the parallel or concurrent information processing operations mobilized simultaneously in order to accomplish the task.

Sweller's model and attention

For Sweller, split attention is an attentional phenomenon which has a direct influence on comprehension. Sweller (2003) believes that split attention occurs when subjects have to process multiple sources of information which they have to integrate at the mental level in order to infer the meaning from the presented material. For example, in geometry, pupils have to split their attention between the processing of the presented figure and the text relating to this figure (Sweller, Chandler, Tierney & Cooper,1990). The presented material cannot be understood unless the subject integrates the various sources of information at the mental level. The cognitive load theory has been used to suggest that the process of splitting attention as well as that of mentally integrating the material are cognitively costly due to the traditional modes of presenting and structuring information. The physical integration of the sources of information, for example by placing written comments at the corresponding positions on the geometrical figure rather than next to it, eliminates the negative effects of split attention on comprehension. This type of result has been obtained by Wickens and co-workers (Wickens & Andre, 1990 ; Boles & Wickens, 1987 ; Wickens, 1987) who have formulated a proximity compatibility principle.

The modality effect: if eliminating the split attention effect improves learning by reducing the working memory load, then the same positive effect can be obtained by increasing the effective capacity of working memory (Jeung, Chandler & Sweller, 1997; Mousavi, Low & Sweller, 1995; Tindall-Ford, Chandler & Sweller, 1997). To achieve this type of objective, it is necessary to employ dual mode didactic presentations, i.e. presentations in which the various sources of information that have to be integrated are presented in different sensory modalities (auditory and visual). For example, when a geometric figure is presented visually and explained orally then learning is better than when a conventional presentation is used (figure and text presented visually). This effect, which was revealed by Penney (1989), could be explained by the fact that because pictorial and verbal information do not pass along the same information processing channels, they do not need to share the same resources and the processing load in WM is limited. This effect has been amply replicated by Mayer and co-workers (Mayer, 2001; Moreno & Mayer, 1999).

The redundancy effect: the modality effect and the effect of reducing attention splitting are obtained if and only if it is necessary to process disparate information in order to accomplish the task. Thus, learning is facilitated by physically integrating this information and presenting it in a way that addresses different sensory modalities. In contrast, when it is not necessary to process disparate information, when a single information source suffices, then the addition of redundant information (for example, by presenting the same information in two different modalities such as a written text and a spoken text, or in two different formats such as a text and a graphic or, again, two different scopes such as a detailed and a summary text) impairs learning (Cerpa, Chandler & Sweller, 1996; Chandler & Sweller, 1991, 1996; Kalyuga, Chandler & Sweller, 1999). A simple explanation for this effect would be that the excess cognitive cost associated with the processing of the useless information interferes with the processing of the relevant information that is to be understood and learned. This finding has also been replicated by Mayer (Mayer, Heiser & Lonn, 2001 ; Mayer, Bove, Bryman, Mars & Tapangco, 1996). In the same vein, the research conducted by Schnotz (1993; Schnotz & Bannert, 2003) into the comprehension of combined text and graphic shows that an illustration interferes with the comprehension of the textual/graphic document when it takes a form which is incompatible with or unsuited to the task that is to be performed by the subject.

The element interactivity effect: the effects of modality, of reducing attention splitting and of redundancy are observed if and only if the material for processing is cognitively demanding, that is to say if it possesses a large number of interacting elements. Overly simple material, containing only a small number of interacting elements, does not yield the above-mentioned effects (Marcus, Cooper & Sweller, 1996; Sweller & Chandler, 1994). This effect is probably explained by the limited capacity of WM, with sufficient resources being present to process simple situations while being inadequate to handle more complex ones. The simplicity/complexity of a situation is thought to depend on the novelty of the elements that form it as well as on their number and the relations between them.

The isolated interacting elements effect: when a learner has to process a very complex situation, processing can be facilitated by initially isolating the elements of which it consists. Once the learning of each isolated element is well underway, the interactions between the elements can be presented. Formulated in this way, this effect has, unlike the others, been obtained only once (Pollock, Chandler & Sweller, 2002). It is, however, very similar to a frequently observed effect in the field of hypertext and hypermedia comprehension: when a structural overview (or cognitive map) is presented before the learning of the corresponding text (or graphics) then the obtained result is worse than when the same structural overview is presented after learning. This effect has been explained by Hofman and Van Oostendorp (1999): subjects are not able to understand the full set of interactions between the elements until they have understood each element one-by-one and each of the interactions one-by-one.

The expertise reversal effect: when dealing with advanced learners, almost all the effects described above cease to operate. In other words: the various effects due to the enrichment of the material all operate in the same way as the redundancy effect (see Kalyuga, Ayres, Chandler & Sweller, 2003, for a summary; this effect has been replicated in numerous experiments giving rise to six different publications; it has also been obtained by McNamara, Kintsch, Songer & Kintsch, 1996). Thus, this latter effect, which has doubtlessly contributed greatly to the development of the new formulation of the cognitive load theory tells us that to design a learning situation is to design a relation between a certain material, a task and a learner, with each characteristic of one being relative to those of the others. Nothing in this field is absolute.

Mayer's model and attention

The main effects revealed by Mayer are similar to those identified by Sweller.

Since multimedia is, by definition, a medium consisting of multiple sources of information, selective attention is extremely necessary in order to select, within a text or the accompanying illustrations, that information which is conducive to comprehension (Mayer, 2001). The multiple representation principle postulates that an explanation will be understood better when this information is presented using two modes of representation rather than just one. Mayer and Anderson (1991, 1992) were thus able to demonstrate that a lesson relating to the functioning of a bicycle pump was understood better when it consisted of a narration and a corresponding animation rather than just of a narration. This is the multimedia effect.

Studies in the field of educational multimedia also show that learning improves when the narration accompanying the illustrations emphasizes the most important information that is to be remembered and understood, either through the use of a special style (bold, underlined) or by limiting the spoken text to the essential information (summary). The principle illustrating this phenomenon is known as the coherence principle. Mayer thus obtained a 50% improvement in valid solutions in a transfer test requiring the solution of a problem when the learners had read a text purged of non-essential material (i.e. presenting only the information whose memorization was necessary and sufficient with reference to the formation of a bolt of lightning) as an accompaniment to the illustration rather than a more detailed text on the same subject accompanying the same pictures. In other words, emphasizing the information that corresponds to the information preselected for retention facilitates learning. In contrast, adding details interferes with the selection of the information which is relevant for comprehension and consequently impairs learning. If this initial conclusion can be understood as a counter to the more classical conclusions concerning textual comprehension, it also represents an extension to them in that the learning, and therefore also the comprehension, situation here is multimedia in nature. What the authors show is that when the narration accompanies an illustration, it must present the information that is necessary and sufficient for comprehension and avoid presenting useless information, and/or details or information which are provided in the illustration. Otherwise, we observe an impairment in comprehension performance which is known as the redundancy effect and which has already been mentioned in connection with Sweller.

Finally, the term 'divided attention" tends to refer to the parallel or concurrent processing of information which is necessary or desired for the completion of the task. In the multimedia field, divided attention is required, for example, when illustrations are located at the end of a text, thus obliging subjects to divide their attention between these two different sources of information in order to achieve comprehension (Mayer, 2001). Research conducted in the field of educational multimedia has shown that the division of attention has a negative effect on comprehension. Mayer has thus formulated the contiguity principle in order to permit the development of multimedia applications or lessons which can result in the optimization of learning performances. The contiguity principle states that learners understand an explanation better when the text and the corresponding illustrations are presented at the same time or are physically close rather than when they are temporally or spatially separated (Mayer & Anderson, 1991, 1992; Mayer & Sims, 1994). As we have seen above, this corresponds to what Wickens calls the proximity compatibility principle and Sweller refers to as the effect of reducing split attention.

Work into divided attention has made it possible to further develop the concepts of attentional or processing resources. In effect, even though processing resources are limited, Wickens (1989) has shown that they attached to one or other of the processing channels. Thus, reading an explanatory text while looking at the illustrations associated with this text risks overloading the visual channel and thus causing a fall-off in comprehension performance. In contrast, listening to an explanatory talk while looking at the illustrations associated with this talk will make it possible to use the full processing load of the visual channel and of the auditory channel for the processing of a single stimulus (the spoken text in the case of the auditory channel and the illustrations in the case of the visual channel) and will not therefore result in any processing overload at the perceptual level. In Mayer's model, this phenomenon is related to the principle of "split attention" which postulates that learning is facilitated when the verbal information relating to an illustration is presented auditorily rather than visually (1989, Mayer & Gallini, 1990; Moreno & Mayer, 1999). This effect is similar to the modality mentioned in the discussion of Sweller above.

Finally, even though they involve the mechanisms responsible for attentional selection, the division of attention across various sources of information and the limitation of the processing capacities of the subject engaged in the operation of comprehension, the models proposed by Schnotz and Perfetti et al. tell us no more about the operation of attention during the processing of multimedia information than do the works conducted by the above-mentioned authors.

Summary of attention-related concepts and multimedia comprehension

At the conclusion of this section dealing with the involvement of attention in the explanatory models of the effect of multimedia on comprehension and learning, it should be noted that many of the effects and principles that have been formulated are directly or indirectly attributable to the operation of attention within information processing. In effect, to understand something it is necessary to attentively select the relevant information and inhibit information which is less relevant or which is in contradiction with the specified objective, and makes it necessary to avoid overloading the information processing system. It therefore demands that the information is placed within a hierarchy in accordance with each item's level of priority.

While many of our activities are attentional, that is to say intentionally activated, controlled during their execution and costly in terms of processing resources, others, in contrast, are automated. Examples of such processes are reading in literate adults or movement detection (Shiffrin & Schneider, 1977). These processes which are activated automatically, which are not costly in terms of processing resources and which are executed very quickly, interfere with the execution of the more controlled processes and thus impair performance in the associated tasks. It would therefore be interesting to evaluate the attentional interference produced by an animation presented in the margin of a text on the comprehension of the text itself. Furthermore, recent attention-related work which has addressed the issue of negative priming has revealed the joint action of two distinct processes in information processing: an automatic process of "similarity processing" which permits the automatic retrieval of information from long-term memory, and a controlled process for the processing of differences, the role of which is to process new or contradictory information (Sweller, 2003).

Finally, even though the question of cognitive or processing overload has been the object of many studies, it would be interesting, in the light of these studies of divided attention, to study this establishment of an information hierarchy on the basis of the priority or importance of the various items as undertaken by subjects engaged in the activity of understanding and learning. When confronted with a large number of documents with differing formats and contents, how do subjects intentionally organize their information capture? Do they establish a document hierarchy starting with the most and ending with the least relevant? On what criteria is the attentional selection of a specific source of information based? Finally, does the relevance of the information change as its format is modified (narration/visually presented text)? Subsequent work will no doubt make it possible to answer these questions which are reminiscent of the processes involved in the comprehension of multiple sources developed within Perfetti et al.'s model (1999).

While attention has a role in the presented models, its definition is reductionist. It can be summarized in terms of two general functions: the selection of the relevant information (inhibition of the irrelevant information can be inferred but is not developed) and the management of the processing resources. This reductionist vision of attention implies that the authors consider that information is processed only at the cognitive level, with the resolution of the interference between different information sources being handled by higher-level processes. However, if cognitive processing is a fundamental stage of information processing, it would nevertheless amount to nothing in the absence of the sensory and perceptual processing of the information and would yield nothing if not followed by a motor response stage. Even though, for the most part, they involve low-level processes, an investigation of these stages in the processing of information would no doubt help shed new light on the effects and principles developed in these models. By way of example, let us take the modality effect which, according to the work on attention, is directly attributable to the limited processing resources in the perceptual, non-cognitive information channels (Wickens, 1989). In other words, if impaired comprehension performance is observed on a multimedia document that incorporates both text and graphics then this is due, to a large extent, to the fact that these two sources of information (pictorial and verbal) have to share the same perceptual processing resources.

After having explained the involvement of attention in the models chosen in this chapter, we shall next address the question of the role played by memory in these models. As we have just done for attention, we shall start this new section with a brief reminder of the various underlying concepts relating to memory.

An individual who comprehends and learns is defined as a being endowed with memory

A recall of the memory-related concepts

A person learning using a multimedia document, like any other learner, wants to acquire, or memorize, information through his or her learning activity. To learn is therefore also to memorize.

In 1986, Baddeley developed the concept of short-term working memory (WM) and specified that "short-term" refers not only to the storage but also to the processing of the information. The concept of short-term memory (STM) itself initially emerged at the end of the 19th century and was subsequently formulated in terms of limited information processing capacity by Miller (1956). Working memory consists of three systems: a control system, known as the central executive, which has the role of selecting the information to be processed and inhibiting the irrelevant or distracting information: the phonological loop whose role is to process the verbal information and the visuo-spatial scratch pad which is responsible for processing the non-verbal information. These two subsystems have limited capacities. Working in the field of mental representations, Paivio (1971) has shown that information is memorized better when it is presented both verbally and pictorially. On the basis of this finding, he developed the dual-coding theory which stipulates, among other things, the establishment of connections in working memory between the processing system used for the verbal information (logogen) and the processing system devoted to the pictorial information (imagen) when these different types of information are presented simultaneously, a phenomenon referred to as "referential connections". It should be noted that the imagen and the logogen are the *alter egos* of the phonological loop and the visuo-spatial scratch pad in working memory. In the light of the multimodal nature of a number of situations, and rejecting the hypothesis that the working memory subsytsems are integrated or connected, Baddeley (2000) added a new subsystem to working memory: the episodic buffer which has the role of processing multimodal information (visual and auditory). The information taken from this subsystem would then be stored in long-term episodic memory.

To summarize, working memory consists of three systems, one of which is a control system which has the role of performing attentional selection, with the two others being subordinate to this initial system and dedicated to a specific type of processing: the phonological loop or auditory working memory is thought to process verbal information while the visuo-spatial scratch pad or visual working memory processes the pictorial information. These two subsystems are of limited capacity and are interconnected, thus permitting the integration and transfer of information from one subsystem to the other.

The limited capacity of WM does not affect all information, only information which is new for the individual. What we already know, fields of knowledge in which we are already experts, represent categories of information that are not very costly to process and which are not therefore subject to the limitations of WM either in terms of processing times or capacity. It should be noted that this fact has long been known and was even clearly stated by Miller (1956) with reference to the processing of notes of music by musicians. It is, instead, the integration of this fact into the model of short-term memory or limited-capacity working memory that has, until recently, proved to be problematic for psychologists. Ericsson and Kintsch (1995) developed the concept of long-term working memory in order to account for this.

Working memory designates the part of long-term memory (LTM) which is active at a given time in order, for example, to perform a task.

Long-term working memory is considered to be a structure and/or a function. As a structure, it is a sort of organized reserve of unlimited capacity in which knowledge or memories can be retrieved or reactivated largely automatically. This knowledge and these memories are not objective. They are constructed by the individual and result from the interpretation and the action that the individual produces when processing a situation. As a function, LTM obeys an adaptive dynamic that is largely based on inhibition. LTM is a cognitive function that permits the cognitive adaptation of human beings to their environment, in particular by inhibiting knowledge or memories which are useless for the processing of current situations. Let us state once more here that we are aware of the classical considerations concerning memory, which have undoubtedly been known since Ebbinghaus (1885) and which are summarized in Schacter (1996). The central role of inhibition in memory functioning now seems to be largely accepted by LTM specialists (Tipper, Grison & Kessler, 2003), and experts in the field of WM (Conway & Engle, 1994), being observed in recall tasks (Sahakyan & Kelley, 2002) and implicit memory tests (Perfect, Moulin, Conway & Perry, 2002).

Strongly influenced by the explanatory models of working memory, the models of multimedia comprehension have made it possible to update our understanding of a number of classical memory phenomena, while also contributing to the current conception of working memory and long-term memory.

Sweller's model and memory

The question addressed by the cognitive load theory is as follows: what are the advantages, from the point of view of the cognitive development of the human species, of possessing (a) a WM which has a limited capacity when processing new information and an "unlimited" capacity when processing information already present in LTM, and (b) an LTM with unlimited capacity.

According to Sweller, when the processed information is new, WM possess no central executive capable of coordinating the processing of this new information. WM becomes efficient when it processes information that has already been learned since this knowledge can act as the central executive. WM is very large in order to maximize the circumstances in which the central processor function will be available. In other words, since the emergence of Miller's work (1956) and Baddeley's model (1986), specialists, first in the field of STM and then of WM, have used tasks in which subjects have to process information that is new to them. This has resulted in a conception that holds that WM has a limited processing capacity both in terms of capacity and duration. However, the capacity of WM is not limited when subjects process information that is well known to them (Sweller often cites Ericsson & Kintsch, 1995). LTM contains schemas, a sort of basic functional element of knowledge, which make it possible to categorize items of information according to the way in which they are to be used. These schemas are activated and used automatically at no cognitive cost. To be an expert in a field is to possess dozens or even hundreds of schemas that are specific to this field of expertise.

The difficulties encountered in processing information are related to two dimensions: novelty and complexity. Complexity designates the set of relations that exist between the items of information. For example, in "a/b = c/d", there are four elements and three relations. In "My grandmother's son's dog", there are three elements and two relations. Understanding implies the simultaneous processing of the set of elements and relations. It is not possible, even temporarily, to eliminate elements or relations, for example by replacing "a/b = c/d" with "a/b & c/d" where & designates any possible relation. Possessing knowledge, i.e. internalized schemas, makes it possible to reduce the complexity of the information to be processed. For example, "My grandmother's son's dog" can be transformed into "My father's dog" or even into "Rover". When we try to understand a new, complex situation, we come up against the limitations of WM. The processing of well-known situations, based on the mobilization of schemas, is performed increasingly automatically and at little cognitive cost. Processing a complex situation involves more than just WM, i.e. it also mobilizes LTM mechanisms and learning. The two main mechanisms of learning are the acquisition of schemas and automation, i.e. two mechanisms for the construction of knowledge that can be activated in a relevant, automatic and costless way.

For Sweller, the central executive is located in the schemas - it is one of their functions. It is not a function of working memory as such . If it were, we would be able to coordinate the processing of new information, which we cannot. When we have to process a new, complex situation for which no direct learning is possible, we undertake a problem-solving activity which gradually leads to the creation of new schemas. In the same vein, Sweller considers that if the human cognitive system had evolved towards an increase in the capacity for processing new, complex situations directly, that is to say towards an increase in WM capacity, then it should be able to process a very large number of possible combinations of the relations between the elements in a situation. In the case of four elements, 4! combinations are possible and this would appear to be a reasonable value. However, as soon as we progress to 5! or 6! possible combinations, the cost of processing becomes exorbitant. The role of the central executive in the schemas is to choose which of the 5! or 6! possible combinations is the right one. This choice is learned. According to Sweller, this argument has already been proposed within a formal perspective by authors such as Dirlam (1972) and MacGregor (1987) and by Elman (1993) and Newport (1990) in the field of grammar learning. Kareev (1995) has shown that in the processing of correlations, a small sample increases the probability that the sample will have a higher correlation than the population. Thus, if a correlation exists, a limited-capacity working memory will be better able to increase the probability that the correlation will be detected. Kareev, Liberman and Lev (1997) have shown empirically that individuals with a smaller WM detect more correlations than individuals with larger WMs.

Sweller (2003) represents this set of considerations in a cognitive matrix of continua (Figure 1).

Insert Figure 1.

The cognitive matrix of continua accounts for a number of phenomena in a coherent way. The information with which we are confronted can be placed on the learning continuum (line 1), with the extreme left corresponding to situations in which we possess no schema and the right to situations in which we do possess schemas . The consequence of this (line 2) is the availability (on the right) or lack of it (on the left) of a central executive function. Thus (line 3), as schemas allow us to select and organize the processed information (on the right), we can understand and react directly when dealing with familiar information. In contrast, (on the left), we have to mobilize a problem-solving activity in cases where we possess no schema. Consequently, the combination of information elements is positioned on a continuum going from order (on the right) to random arrangement (on the left), thus bringing the WM limitations into play (on the left) or not (on the right).

Why has the human cognitive system not evolved in a way that permits WM to process new, complex situations? The answer is that it has evolved towards learning. According to Sweller, the goals of learning are therefore probably largely determined by the adaptive relation between the cognitive system and its environment and not by the subject's conscious will.

To summarize Sweller's theory, we can say that comprehension is the representational mechanism that underpins the interaction between an individual and a situation. Memory and learning are two mutually complementary, adaptive functions which assist comprehension in cases where the interaction between the individual and the situation is taking place for the first time as well as in cases where it tends to be repeated. Teaching is an artificial intervention that has the aim of mobilizing these two functions (memory and learning) in cases where the interaction between the individual and the situation is lacking or is insufficient in itself to mobilize these functions.

The cognitive load theory thus clearly derives from these three major cognitive psychological theories: the theory of schemas (Piaget, 1924; Bartlett, 1932; Gick & Holyoak, 1980, 1983; Chi, Glaser & Rees, 1982), the theory of cognitive architecture in terms of WM/STM (Atkinson & Shiffrin, 1968) and the theory of WM (Baddeley, 1992). However, given the way it has developed, Sweller's theory now goes beyond its precursors to integrate the theory of LTWM (Ericsson & Kintsch, 1995) and provide an evolutionary theory of WM which situates the central executive within the schemas and not within WM and interprets our ability to learn as the counterpart of WM functioning

Mayer's model and memory

According to Mayer (2001; Moreno & Mayer, 1999), visually presented information is initially processed in visual working memory whereas auditorily presented information is first processed in auditory working memory. For example, when a text is read, the words may first pass through visual working memory and then, once translated into sounds, pass into auditory working memory. In a multimedia learning situation in which a narration was presented together with a corresponding animation concerning lightning formation, Mayer showed that performance was better than in situations in which these two types of information (verbal/pictorial) are presented sequentially. Mayer holds that subjects in a multimedia presentation situation are simultaneously able to take account of the corresponding verbal and pictorial representations in working memory. They are thus better able to construct referential connections between these representations. This phenomenon is related to the principle of the multiple representation of information.

These results nevertheless need to be seen in the light of the work conducted by Shnotz into the comprehension of documents containing combined text and graphics in which he shows that the format in which the pictures are presented can sometimes impair the learner's comprehension of the document. Here, Schnotz has revealed a variation in comprehension performance that is associated with a modification in the format in which the graphic information is presented. In a similar vein, Mayer has also revealed a modulation in the contiguity, "split attention" and multimedia effects as a function of individuals' specific characteristics. For example, beginner learners in a field of knowledge exhibit significantly greater multimedia and contiguity effects than expert learners in the same field (Mayer & Gallini, 1991, Mayer, Steinhoff, Bower & Mars, 1995).

Schnotz's model and memory

Schnotz and Bannert (2003) distinguish between two forms of representation: symbolic representations (which relate to the analysis of the text) and analogue representations (which relate to the visual processing of the graphic). To understand a text which is illustrated by graphics, the reader must both construct a propositional representation of the text and a mental model of the graphic which illustrates it and then compare these two types of representation. Among other things, the construction of a coherent conceptual organization of an illustrated text document requires the operation of selective attentional processes (the role of which is to select the information relevant to the task), the organization of the selected information and interactions with long-term memory in order to construct a mental model which integrates the information drawn from the text and the graphic. On the basis of his work, the author demonstrates that the advantage associated with the dual-coding of information is not systematic and even goes on to question Paivio's (1986) dual-coding theory. In effect, according to Paivio, the dual-coding of the text and the graphics should lead to the construction of more elaborate representations or cognitive structures than are generated by the coding of the text alone, thus promoting a better memorization of the text and better knowledge acquisition performances. Since Paivio's dual-coding theory does not advance any hypothesis concerning the quality of the presented graphic or illustration on comprehension performance, the presence of a graphic or illustration should facilitate comprehension whatever the task to be performed by the subject and irrespective of the form taken by the graphic material. Schnotz (1993), however, maintains that the type of presented graphic or illustration influences the subject's comprehension performance as a function of the type of task to be performed. He thus demonstrates that if the form taken by the illustration (carpet diagram versus circular diagram) is not compatible with the proposed task then it interferes with the learner's comprehension performance. The incompatibility of the form of the illustration with the task to be performed, on the one hand, and the text on the other, makes it more difficult for the subject to establish a link between the data drawn from the text and the graphic and thus hinders the integration of the two types of information (textual and graphic) within a coherent conceptual organization.

Perfetti et al.'s model and memory

Classically, when dealing with the comprehension of a simple document which comes from a single source, it is accepted that the reader's representation is coherent. However, the increasingly widespread distribution of multisource documents poses a problem for psychologists since the principle of representational coherence no longer obtains. To understand a document constructed from multiple contradictory sources it is necessary to elaborate a contradictory, incoherent representation. Studies of the comprehension of multisource documents have revealed the mobilization of specific types of processing which result in the elaboration of a multidocument representation (Perfetti, Rouet & Britt, 1999). According to Wineburg (1991, 1994) and Rouet (2000) expert readers perform document indexing when processing multisource documents. What is the type of text? Who is the author? What is the date of publication? etc. They also compare the texts. What are the contradictions? What corroborations are there? Finally, they proceed to contextualize the content. What places, times, conditions are referred to? This processing is believed to result in the construction of a multidocument representation which, according to Perfetti et al. (1999), consists of two elements: on the one hand, an intertext model in which the different sources of information are represented together with certain content elements and intersource relations and, on the other, an intersituation model in which the different situations proposed by the documents are represented. Furthermore, this model may relate to one and the same situation in which certain aspects are specific and others variable depending on the sources in question. Depending on the case in question, the user may construct separate representations of the different contents, a merged representation of the different contents, a full indexing of each content to its source or a selective indexing of certain contents to certain sources (Britt, Perfetti, Sandak & Rouet, 1999). This latter representation would be the most frequent in good readers of documents, corresponding to what Perfetti et al. (1999) refer to as a multidocument representation. The construction of such a representation would therefore involve the mobilization of expertise which is specific to the processing of multiple documents and the mobilization of conceptual knowledge relating to the field to which the content relates. The role of LTM is therefore much the same as in Kintsch's model, with two exceptions. In LTM, some knowledge is represented both in terms of content and source and not just in terms of content (Rouet, Britt, Mason & Perfetti, 1996). This is a novel suggestion which clearly does not simply map the classic semantic/pragmatic duality and for which the authors provide convincing evidence. In LTM, some knowledge is represented in a non-coherent way or, alternatively, some knowledge which is coherent thanks to the intersource model, establishes relations between contents which are not coherent with one another. The intersituation model is not coherent whereas the intersource model takes account of this incoherence. Here, we can see a development from the 1996 version of the model (Rouet et al., 1996) and its 1999 formulation (Perfetti et al., 1999). In 1996, the authors had not yet coined the term "multidocument representation". Instead, they spoke of "rhetorical model" in order to designate the representation constructed by an individual reading multiple documents and seemed to consider the intersource model to be coherent. In 1999, not only were the authors more precise but they also seemed to adopt the idea that non-coherent representations might exist both in WM and LTM. This idea of non-coherence has different implications depending on whether we are considering WM or LTM. We believe that the question of identifying the consequences represents a challenge in the field of cognitive psychology. When discussing comprehension, it will also be necessary to try to be clear as to the meaning attributed to the word "coherence". In effect, the psychology of reasoning has taught us that reasoning can be incorrect at the logical level but perfectly correct at the normative level (in the sense of social norms). If we study cognitive biases solely at the logical level then we come to the conclusion that coherence may possibly not be a central value in the organization of knowledge in WM. If we consider matters at the normative level, or indeed the adaptive level, this lack of coherence in reasoning and of the organization of knowledge in WM is much less evident. To return very briefly to a classical idea proposed by Johnson-Laird (1983), if humans often process implications as reciprocal implications then it is because this is often true in their environment.

When individuals process new documents simultaneously or even when they sequentially process a large number of different documents, for example when performing research work on the Web, do they construct a coherent representation? Does what they construct bear the slightest resemblance to a situation model? Do they learn anything coherent? As far as comprehension is concerned, it would appear that addressing is more important than coherence, whether we are speaking of the processed situations or knowledge in LTM.

Summary of the concepts relating to memory and multimedia comprehension

The models of multimedia comprehension are clearly closely associated with the theories of memory. The models produced by Baddeley (1986), as well as Atkinson and Shiffrin (1968) constitute explicit or implicit reference models. More recent developments in terms of long-term working memory (Ericsson & Kintsch, 1995) or the adaptive theory of memory whether within an overtly Darwinian perspective (Klein, Cosmides, Tooby & Chance, 2002) or not (Schacter, 1996) are also incorporated in a more or less explicit way, at least by Sweller. In contrast, inhibition, which is now conceived of as one of the fundamental mechanisms of memory functioning does not seem to be clearly integrated in the models of multimedia comprehension that we have described. Is it possible that comprehension could be interpreted as an activity that inhibits irrelevant sources or representations? This is clearly the case in Kintsch's model which describes the suppression inferences. Might not inhibition be one of the mechanisms mobilized by individuals when they have to confront redundant documents or interfering sources?

Furthermore, the contribution made by the models of multimedia comprehension to our knowledge of memory phenomena is far from negligible. Sweller's and Mayer's models both constitute an original, general proposition concerning the functioning of memory and learning. The model developed by Perfetti et al. raises an important question, namely that of the non-coherence of representations, to which research in the field of cognitive psychology must find an answer, both with reference to WM and LTM. Finally, Schnotz's model represents an important review of the model proposed by Paivio.

Conclusion

In the introduction, we asked whether the theoretical literature relating to multimedia comprehension was an *ad hoc* phenomenon or whether it entered clearly within the framework of today's psychological theories. We also asked what contribution this literature has made to the psychology of learning. In this conclusion, we shall attempt to show how the models presented here fit within the framework of current psychological theories and how this research into multimedia comprehension represents a source of development for the more classical models of cognitive psychology.

Insert Figure 2

Figure 2 presents an overview of multimedia comprehension as it is conceived of in the models presented above, thus allowing us to reply to the two questions we have just posed. Multimedia comprehension can be envisaged either on the basis of the document, by relating each aspect of the document (document type and format - contents - expected response) to the various levels of information processing (sensory, semantic, technical level; Schnotz & Lowe, 2003), or in terms of the processes involved - selection, organization, integration and action (Mayer, 2001; Sweller, 1999) - which mobilize, on the one hand, attentional or processing resources and, on the other, short-term and long-term memory structures.

As Schnotz and Lowe (2003) point out, a multimedia document can be considered at three levels: the sensory level, the semantic level and the technical level. Our examination of Perfetti et al.'s work has shown us that, in the case of multiple documents, the semantic level may not be coherent. The cognitive processing performed by an individual can be represented with reference to these levels of the document.

At the sensory level, the document primarily prompts the process for the selection of relevant information developed by Mayer. This process includes the early attentional selection processes. We have listed the effects revealed by Sweller's and Mayer's models at this level of processing. However, we have also emphasized that these authors have devoted little effort to the study of this level *per se*, that is to say as a psychological process. We have, in this regard, posed a number of questions which, as far as know, are still awaiting an answer within the field of multimedia comprehension.

The processes which Mayer refers to as organization and integration correspond to the semantic level of the document. The works which we have surveyed emphasize the special nature of organization in the case of multimedia and/or multisource documents. In particular, they stress the fundamental role of media and source integration as a constituent element of comprehension. Mayer and Sweller have indicated a significant number of ways of positively influencing this integrative activity through the configuration of different subsections of the document. Perfetti et al. consider that this integration may not necessarily be coherent or, at least, that it may be more addressed than coherent, and distinguish between the intersource and intersituation model levels.

Finally, the individual's or user's action corresponds to the technical level of the document. We have pointed out that this level is ignored in the models of comprehension.

It is precisely the models of multimedia comprehension that address the way in which memory is involved in these types of processing. Working memory as presented in these models is based on the classic definition of this memory structure (Baddeley, 1986; Paivio, 1986) which the authors believe to be (a) composed of two different systems for short-term information processing (one for graphic information and one for verbal information) and an attentional component which selects the information to be processed, (b) endowed with a limited capacity to process new, complex information, (c) endowed with an unlimited capacity to process information which is already known (d) capable of having parts of its capacity freed up by certain specific configurations of the multimedia document, a fact which, logically, echoes what we have just seen in connection with attentional processing. WM is therefore directly involved in the selection and organization operations. We have emphasized the fact that Sweller's idea that the central executives are the schemas, i.e. located in LTM and not in WM as such, is an original and persuasive proposal. Finally, we have indicated how knowledge stored in LTM is involved in organizational and integrative processing.

We shall now discuss a critical aspect of the relations between these four models. While all the models used in this chapter tend to explain the influence of multimedia on learning, it should be noted that the concepts of "multimedia" or "multiple information sources" do not mean the same thing for all authors. For Sweller and Mayer, multiple information sources are often linguistic and pictorial representations, the content of which refers in a homogeneous way to one and the same phenomenon or entity (a wiring diagram, an electrical drawing, the heart and the circulation of the blood, storm formation). Typically, Sweller and Mayer focus on the interaction between descriptive (the text) and depictive (the graphics) representations as defined by Schnotz (2001), even though they do not cite this author. In contrast to this initial definition of multimedia, in Perfetti et al. (1999; see also Rouet, Britt, Mason & Perfetti, 1996; Rouet, Favart, Britt & Perfetti, 1997) the term "multiple sources" refers to documents created by different authors and/or using different media (historical controversy concerning the Panama canal; newspaper debate concerning the demonstrations in Noumea). For Sweller, Schnotz and Mayer, the aim is to develop a coherent representation (from the author's perspective) of the multimedia document by the subject who is attempting to understand it. Given this approach, these authors are particularly interested in the internal cognitive factors such as expertise in the field of knowledge in question which may exert an influence on comprehension. In particular, Sweller's work on the influence of electrical know-how on the comprehension of a multimedia document presenting an electrical assembly has made it possible, among other things, to develop the concept of redundancy.

For Perfetti et al., however, the aim is the development by the recipient (the subject who is seeking to understand) of a multidocument representation (which is not necessarily coherent from the point of view of the authors) which incorporates both an intersituation and an intertext model. Working within this perspective, the authors have developed studies designed to gain a better understanding of the influence of external factors such as the type of task asked of the subject on the subject's comprehension (Rouet, Vidal-Abarca, Bert-Erboul & Millogo, 2001) without, however, ignoring the internal factors (Rouet et al., 1997). The study of cognitive factors, both internal (age, sex, expertise, motivation) and external (aim of the task, type of material, environmental factors), therefore represents an interesting avenue of research in our search for a better understanding of multimedia effects on comprehension.

Throughout this chapter, we have recalled the importance of the concepts of attentional division and selection, on the one hand, and, on the other, the concepts of working memory and long-term memory as explanatory factors for the effects of multimedia on comprehension. The value of this approach lies in the fact that, by drawing together various theoretical strands, it contributes new experimental paradigms to the study of multimedia comprehension. As we have emphasized here, we believe that it is important to be able to consider the perceptual and attentional processes in themselves within work devoted to the question of multimedia comprehension. It would, for example, be interesting, within the framework of the study of the effect of "split attention" on comprehension, to use eye-movement recording equipment, to evaluate subjects' eye movements over the various presented documents as a function of whether or not the subjects consider these documents to constitute a reference or an illustration in connection with the topic for comprehension, in other words as a priority or secondary source of information. Within the same perspective, the study of eye movements and fixation times on the presented documents as a function of the task to be performed (solve a problem, answer precise questions, make inferences) or the general context of the task would make it possible to assess the relative importance of each document for comprehension. In the same vein, it would be interesting to study the redundancy effect in connection with the work devoted to interference in the field of selective attention in order to gain a better understanding of why two identical items of information, i.e. items that are congruent in every respect, impair rather than facilitate comprehension. The most recent work undertaken in the field of the cognitive load theory is starting to address the question of eye movements thanks to the development of tools for the recording and analysis of such movements during the processing of a multimedia document

(Tabbers, 2002). This will make it possible to use existing protocols and paradigms for the differentiated analysis of the multimedia effect on learning.

References

Atkinson, R. & Shiffrin, R. (1968). Human memory: A proposed system and its control processes. In K. Spence & J. Spence (Eds.), The psychology of learning and motivation (Vol. 2, pp. 89-195). New York: Academic Press.

Baddeley, A. (1986). Working memory. Oxford: Clarendon Press.

Baddeley, A. (1992). Working memory. Science, 255, 556-559.

Baddeley, A. (2000). The episodic buffer: a new component of working memory? Trends in Cognitive Sciences, 4(11), 417-423.

Bartlett, F. (1932). Remembering: A study in experimental and social psychology. London: Cambridge University Press.

Boles, D.B., & Wickens, C.D. (1987). Display formatting in information integration and non integration tasks. Human Factors, 29, 395-406.

Britt, M. A., Perfetti, C. A., Sandak, R., & Rouet, J. F. (1999). Content integration and source separation in learning from multiple texts. In S. R. Goldman, A. C. Graesser, & P. van den Broek (Eds.). Narrative comprehension, causality, and coherence: Essays in honor of Tom Trabasso (pp. 209-233). Mahwah, NJ: Lawrence Erlbaum Associates.

Broadbent, D. E. - (1958). Perception and communication. London: Pergamon Press.

Cerpa, N., Chandler, P., & Sweller, J. (1996). Some conditions under which integrated computer-based training software can facilitate learning. Journal of Educational Computing Research, 15, 345-367.

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

Chandler, P., & Sweller, J. (1996). Cognitive load while learning to use a computer program. Applied Cognitive Psychology, 10, 151-170.

Chi, M., Glaser, R. & Rees, E. (1982). Expertise in problem solving. In R. Sternberg (Ed) Advances in the psychology of human intelligence (pp. 7-75). Hillsdale, NJ: Erlbaum.

Conway, A.R.A. & Engle, R.W. (1994). Working memory and retrieval: A resource-dependent inhibition model. Journal of Experimental Psychology: General, 123, 354-373.

Dirlam, D. (1972). Most efficient chunk sizes. Cognitive Psychology, 3, 355-359.

Ebbinghaus, H. (1885). Über das Gedächtnis. Untersuchungen zur experimentellen Psychologie. Leipzig: Duncker & Humblot.

Elman, J. (1993). Learning and development in neural networks: the importance of starting small. Cognition, 48, 71-99.

Ericsson, K. A. & Kintsch, W. (1995). Long-term working memory. Psychological Review, 102, 211-245.

Gick, M. L. & Holyoak, K. J. (1980). Analogical problem solving. Cognitive Psychology, 12, 306-355.

Gick, M. L. & Holyoak, K. J. (1983). Schema induction and analogical transfer. Cognitive Psychology, 15, 1-38.

Hofman, R. & Van Oostendorp, H. (1999). Cognitive effects of a structural overview in a hypertext. British Journal of Educational Technology, 30(2), 129-141.

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

Johnson-Laird, P.N. (1983). Mental models. Cambridge, MA: Harvard University Press.

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. Educational Psychologist, 38, 23-32.

Kalyuga, S., Chandler, P. & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. Applied Cognitive Psychology, 13, 351-371.

Kareev, Y. (1995). Through a narrow window: Working memory capacity and the detection of covariation. Cognition, 56, 263-269.

Kareev, Y., Lieberman, I. & Lev, M. (1997). Through a narrow window: Sample size and the perception of correlation. Journal of Experimental Psychology: General, 126, 278-287.

Kintsch, W. (1998). Comprehension: A paradigm for cognition. New York: Cambridge University Press.

Klein, S.B., Cosmides, L., Tooby, J., & Chance, S. (2002). Decisions and the evolution of memory: Multiple systems, multiple functions. Psychological Review, 109, 306-329.

MacGregor, J. (1987). Short-term memory capacity: Limitation or optimization? Psychological Review, 94, 107-108.

Marcus, N., Cooper, M., & Sweller, J. (1996). Understanding instructions. Journal of Educational Psychology, 88, 49-63.

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

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

Mayer, R.E. (1997). Multimedia learning: Are we asking the right questions? Educational Psychologist, 32, 1-19.

Mayer, R.E. (1989). Models for understanding. Review of Educational Research, 59, 43-64.

Mayer, R. (2001). Multimedia learning. New York: Cambridge University Press.

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

Mayer, R., & 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., 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., Heiser, J. & Lonn, S. (2001). Cognitive contraints on multimedia learning: When presenting more material results in less understanding. Journal of Educational Psychology, 93, 187-198.

Mayer, R.E. & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual-processing systems in working memory. Journal of Educational Psychology, 90(2), 312 - 320.

Mayer, R., Steinhoff, K., Bower, G. & Mars, R. (1995). A generative theory of textbook design: Using annotated illustrations to foster meaningful learning of science text. Educational Technology Reasearch and Development, 43 (1), 31-43.

McNamara, D., Kintsch, E., Songer, N.B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, Background knowledge, and levels of understanding in learning from text. Cognition and Instruction, 14, 1-43.

Miller, G.A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review, 63, 81-97.

Moreno, R., & Mayer, R. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. Journal of Educational Psychology, 91, 358-368.

Mousavi, S.Y., Low, R. & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. Journal of Educational Psychology, 87, 319-334.

Newport, E. (1990). Maturational constraints on language learning. Cognitive Science, 14, 11-28.

Paivio, A. (1971). Imagery and verbal processes. New York: Holt, Rinehart, & Winston.

Paivio, A. (1986). Mental representations. New York: Oxford University Press.

Penney, C.G. (1989). Modality effects and the structure of short term verbal memory. Memory and Cognition, 17, 398-422.

Perfect, T.J., Moulin, C.J.A., Conway, M.A. & Perry, E. (2002). Assessing the inhibitory account of retrieval induced forgetting with implicit memory tests. Journal of Experimental Psychology: Learning, Memory and Cognition, 28, 1111-1119.

Perfetti, C.A., Rouet, J.-F., & Britt, M.A. (1999). Towards a theory of documents representation. in H. van Oostendorp & S.R. Goldman (Eds.), The construction of mental representations during reading (pp. 99-122). Mahwah, NJ: Lawrence Erlbaum Associates.

Piaget, J. (1924). Le jugement et le raisonnement chez l'enfant. Neuchâtel : Delachaux et Nieztlé.

Pollock, E., Chandler, P., & Sweller, J. (2002). Assimilating complex information. Learning and Instruction, 12, 61-86.

Reder, L. & Anderson, J. R. (1980). A comparison of texts and their summaries: Memorial consequences. Journal of Verbal Learning and Verbal Behavior, 19, 121-134.

Reder, L. & Anderson, J. R. (1982). Effects of spacing and embellishment on memory for main points of a text. Memory & Cognition, 10, 97-102.

Rouet, J.-F. (2000). Les activités documentaires complexes. Aspects cognitifs et développementaux. Mémoire pour l'habilitation à diriger des recherches. Université de Poitiers ?

Rouet, J.-F., Britt, M.A., Mason, R.A. & Perfetti, C.A. (1996). Using multiple sources of evidence to reason about history. Journal of Educational Psychology, 88(3), 478-493.

Rouet, J.-F., Favart, M., Britt, M.A. & Perfetti, C.A. (1997). Representation and use of multiple documents by novice and expert history students. Cognition and Instruction, 15(1), 95-106.

Rouet, J.-F., Vidal-Abarca, E., Bert-Erboul, A. & Millogo, V.E. (2001). Effects of information search tasks on the comprehension of instructional text. Discourse Processes, 31(2), 163-186

Sahakyan, L., & Kelley, C.M. (2002) A contextual change account of the directed forgetting effect. Journal of Experimental Psychology: Learning, Memory, and Cognition, 28, 1064-1072

Schacter, D.L. (1996). Searching for memory: The brain, the mind, and the past. New York: Basic Books

Shiffrin, R.M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. Psychological Review, 84, 127-190.

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

Schnotz, W. (2001). Sign systems, technologies, and the acquisition of knowledge. In J.F. Rouet, J. Levonen & A. Biardeau (Eds.), Multimedia learning. Cognitive and instructional issues (pp. 9-29). Amsterdam: Elsevier.

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

Schnotz, W. & Lowe, R. (2003). External and internal representations in multimedia learning, Learning and Instruction, 13, 117-123.

Sweller J. (1999). Instructional design. Melbourne : ACER Press.

Sweller, J. (2003). Evolution of human cognitive architecture. In B.H. Ross (Ed.), The psychology of learning and motivation (vol. 43, pp. 215-266). New York: Academic Press.

Sweller, J., & Chandler, P. (1994) Why some material is difficult to learn. Cognition & Instruction, 12, 185-233.

Sweller, J., Chandler, P., Tierney, P. & Cooper, M. (1990). Cognitive load and selective attention as factors in the structuring of technical material. Journal of Experimental Psychology: General, 119, 176-192.

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

Tindall-Ford, S., Chandler, P., & Sweller, J. (1997). When two sensory modes are better than one. Journal of Experimental Psychology: Applied, 3, 257-287.

Tipper, S.P., Grison, S., & Kessler, K. (2003) Long-term inhibition of return of attention. Psychological Science 14 (1), 19-25

Tuovinen, J., & Sweller, J. (1999). A comparison of cognitive load associated with discovery learning and worked examples. Journal of Educational Psychology, 91, 334-341.

Wickens, C.D. (1989). Resource management and time sharing. In, J.I. Elkind, S.K. Card, J. Hochberg and B.M. Huey (Eds.), Human performance models for computer-aided engineering (pp.180-202). Washington, DC: National Academy Press..

Wickens, C.D. & Andre, A.D. (1990). Proximity compatibility and information display: effects of color, space, and onjectness on information integration. Human Factors, 32, 61-77.

Wickens, C.D. (1987). Attention. In P. Hancock (Ed.), Human factors in psychology. Amsterdam : Noth Holland.

Wineburg, S. (1991). Historical problem solving: A study of the cognitive processes used in the evaluation of documentary and pictorial evidence. Journal of Educational Psychology, 83, 73-87.

Wineburg, S. (1994). The cognitive representation of historical texts. In G. Leinhardt, I. Beck, & C. Stainton, (Eds.), Teaching and learning in history (pp. 85-135). Hillsdale, NJ: Erlbaum.

List of figures

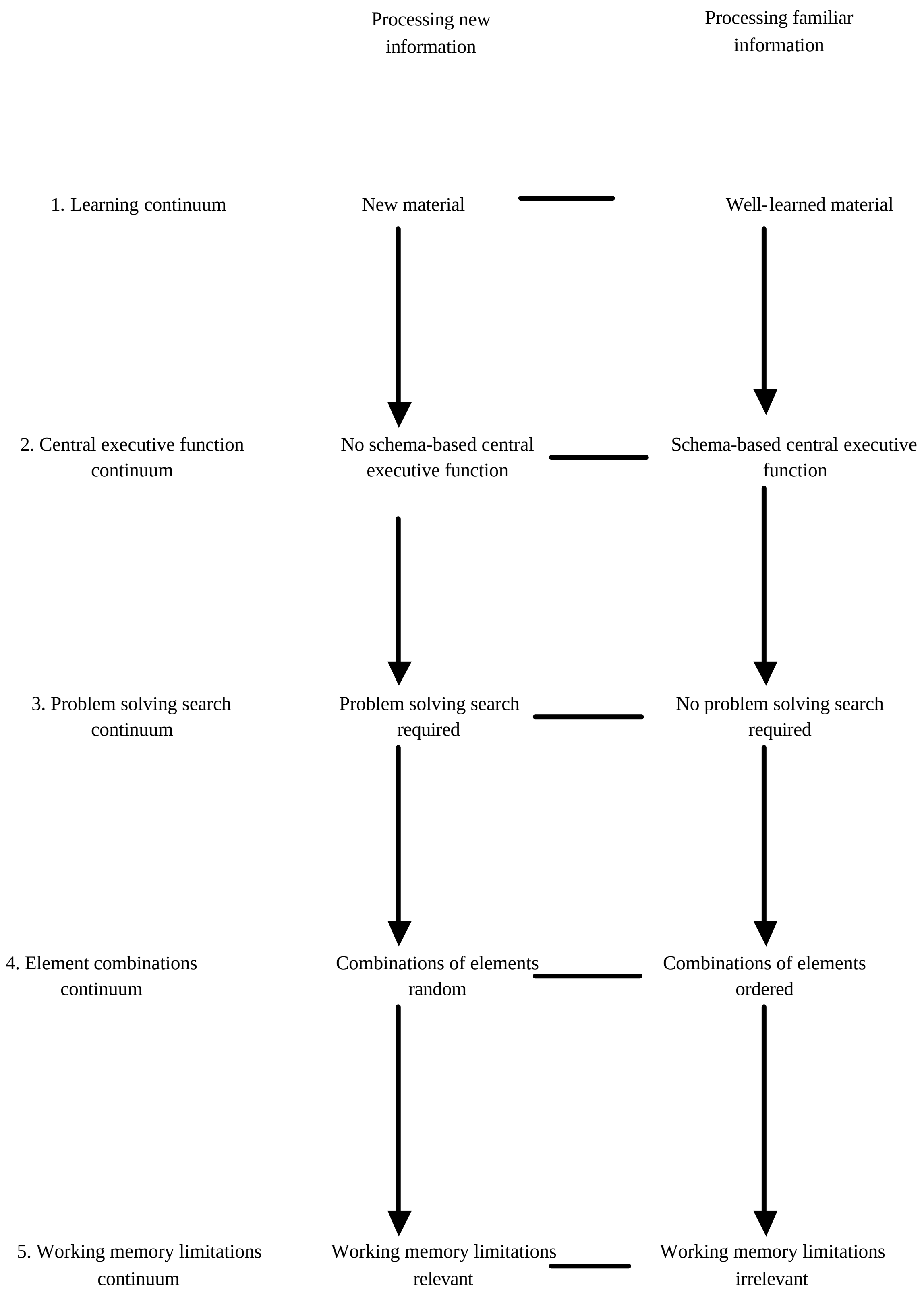

Figure 1. A cognitive matrix of continua (Sweller, 2003)

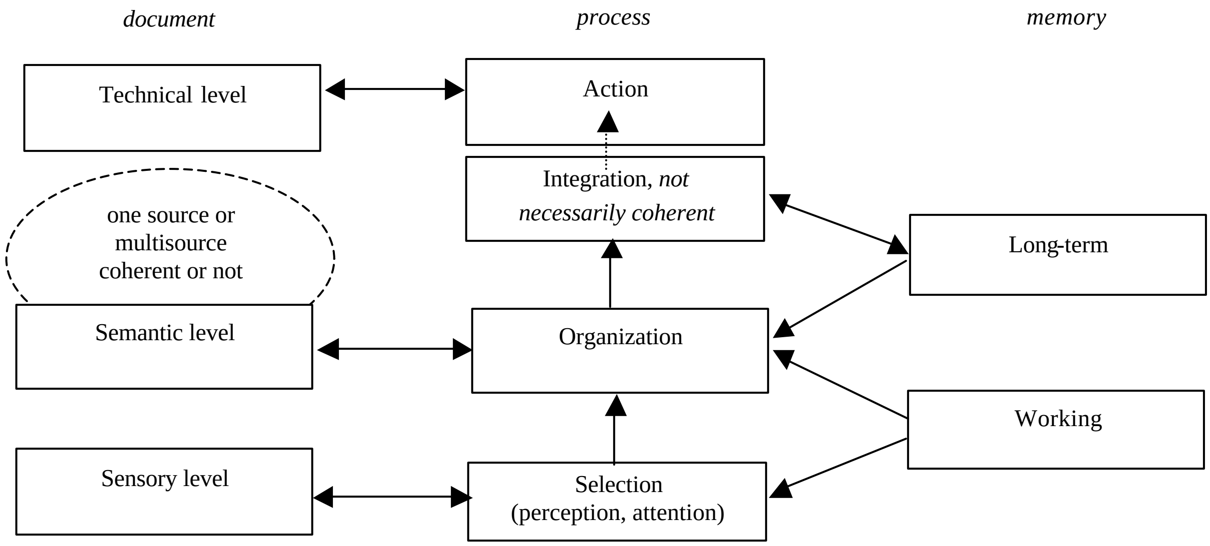

Figure 2. Document levels, processing and memory types involved in multimedia comprehension