

The Design of Multimedia Learning Based on Working Memory

Liying Zhao

College of Educational Science and Technology

Zhejiang University of Technology

Hangzhou, China

zly@zjut.edu.cn

Abstract-The limited working memory is one of the significant factors that constrain the effectiveness of multimedia learning. In the framework of working memory and its related model, this paper provides a tutorial overview of four fundamental theories on the design of multimedia learning, and investigates some guidelines for managing cognitive load and enhancing instructional efficiency of multimedia learning environments.

Keywords-working memory; multimedia learning; cognitive load

I. INTRODUCTION

Multimedia allows the combination of different presentation formats, such as texts, pictures, animations and music in flexible ways via different sensory modalities. So there is a strong temptation to simply assume that multimedia can motivate learners, increase their invested cognitive effort, as well as allow for better adaption of instruction to the learners' needs and preferences, and ultimately results in better learning. However, mixed findings in research on effectiveness of multimedia learning have been found. One of the main reasons is that there is a potential problem in the acquisition of information via multiple representations is that the cognitive demands may overwhelm working memory capacity. Based on the research on working memory, this paper analyzes the bases of the main theories of multimedia comprehension and describes these models, then explores some issues that should be paid attention in the process of multimedia learning design.

II. WORKING MEMORY AND WORKING MEMORY MODEL

A. Working Memory

Working memory is a system for temporarily storing and managing the information required to carry out complex cognitive tasks such as learning, reasoning, and comprehension [1]. However, working memory is a limited system that can store about seven elements but operates on just two to four elements. It is able to deal with information for no more than a few seconds with almost all information lost after about 20 s unless it is refreshed by rehearsal and into long term memory.

B. Baddeley's Working Memory Model

Baddeley and Hitch originally proposed working memory model, which has become a dominant conception of WM in cognitive psychology. In the model, WM is thought of as a multicomponent temporary storage and processing system including one central executive system and two storage subsystem: the phonological loop and visuospatial scratchpad. The phonological loop or VWM keeps phonological entries active under the control of an articulated process. The visuospatial scratchpad or VSWM is responsible for retaining visual and spatial information, and possibly for the formation and manipulation of mental images [2]. According to this model, if the multimedia information involved with multiple channel, then the learners can process information in VWM and VSWM simultaneously, this is actually a kind of extension

of working memory capacity and is beneficial for learning novel information. In a new version of working memory model, Baddeley proposed a new component-episodic buffer, which is conceived to be able to integrate information from various source under the control of the central executive, and as a temporary store that retains complex information [3]. It is in the episodic buffer multimedia information such as graphics and text could be temporarily stored and integrated.

Based on the characteristics of working memory and Baddeley's working memory model, many psychologists have developed theoretical models of multimedia learning to investigate into the issue that how multimedia instruction interacts with the human cognitive architecture.

III. MULTIMEDIA LEARNING THEORIES BASED ON WORKING MEMORY

A. *Dual Coding Theory*

The dual coding theory proposed by Paivio attempts to give equal weight to non-verbal and verbal processing. The theory assumes that there are two cognitive subsystems, one specialized for the representation and processing of nonverbal objects/events (i.e., imagery), and the other specialized for dealing with language. Paivio also postulated two different types of representational units: "imagens" for mental images and "logogens" for verbal entities which he describes as being similar to "chunks" as described by Miller. Logogens are organized in terms of associations and hierarchies while imagens are organized in terms of part-whole relationships [4]. The dual coding theory assume that having two memory codes to represent an item provides a better chance of remembering that item than having only one single code, which is an important theoretical explanation of multimedia advantages.

B. *Cognitive Load Theory*

Sweller has developed cognitive load theory, which attempts to interpret the relation between human information processing system and the framework of instructional procedures and learning. This theory distinguishes three sources of cognitive load: intrinsic, extraneous, and germane cognitive load [5]. Sweller and his collaborators have conducted many experiments that have

shown better learning of instructional materials when presented in an integrated, rather than a split format, and they have developed several instructional techniques, such as split-attention effect, modality effect, completion effect, worked-example effect, to aid the design of multimedia learning [6]. However, Kalyuga has found that designs and techniques mentioned above that are effective with low-knowledge individuals can lose their effectiveness or even have negative consequences for more proficient learners, this is called expertise reversal effect [7]. All these research in CLT have important implications for multimedia learning design.

C. *Mayer's Multimedia Learning Theory*

Based on Paivio dual coding theory, Baddeley's concept of working memory and Sweller's cognitive load theory, Mayer developed a model specifically for multimedia learning. In his model, he assumes that the sounds are organized into a verbal model and the visual images into a pictorial model. Working memory is used to integrate the verbal model, pictorial model, and prior knowledge stored in LTM [8]. Mayer preferred mode of presentation is to present auditory words (not text) so they do not conflict with the visual code that is needed for pictures. Mayer and his students have done dozens of experiments and developed seven principles for multimedia learning design: multimedia principle, spatial contiguity principle, temporal contiguity principle, coherence principle, modality principle, redundancy principle, and individual differences principle [9].

D. *Schnotz's Integrated Model For Multimedia Learning*

Schnotz and Bannert's model of multimedia learning is aimed at explaining how information from different external representations is integrated. It consists of a descriptive and a depictive branch of processes. The descriptive branch comprises processes of symbol analysis that first construct a surface representation and then a propositional representation of the externally represented text. The depictive branch comprises analog structure mapping processes that first construct a visual image and then a mental model of the externally represented picture or diagram [10]. The mental model is the representation that integrates attributions from the text base, pictorial elements from the visual image, and

general world knowledge into a new, coherent structure representing the entities that text and picture are jointly referring to, this integration process requires abundant working memory resources. Schnotz and Bannert's model is especially suitable to derive hypotheses on working memory demands required for learning based on verbal and pictorial materials [11].

IV. SOME GUIDELINES ON MULTIMEDIA LEARNING DESIGN

A. *The Integration of Screen Text and Graphic*

One of the important principles about the integration of text and diagrams or pictures is mentioned in cognitive load theory called split-attention effect, which occurs when multiple sources of information refer to each other but are presented in isolation, the mental coordination among these information will require additional cognitive resources that are not directly related to learning, resulting the waste of substantial working memory resources [12]. One way to solve this effect is that each textual statement should be located just near its matching parts on the picture. However, an important condition for the split-attention effect to occur is the interdependency between different information sources, when neither resource can be understood on its own without attending to the related ones. So if one form of instruction is adequate, providing the same information in a different form will produce an extraneous cognitive load, resulting in redundancy effect. Therefore, simultaneous presentation of text and picture messages, the two complement should compensate with each other, and not just repeat the same content, otherwise it will increase the cognitive load and is also a waste of time and effort.

Another thing should also be concerned when screen-text and picture are displayed simultaneously. Research has shown that learner often incorrectly identified multimedia presentations as picture-based only. At some level this suggests that individuals may intuitively feel that pictures are more important or effective than text for understanding learning tasks. This does not mean that participants actively ignore text, but rather that in retrospect, they find the pictures more readily identifiable as useful for completing procedural tasks [13]. One implication of such findings is that

presentation may need to readily focus users' attention to text when it is critically important, or when it conveys information that the pictures could not include.

B. *The Combination of Visual and Auditory Information*

Many multimedia presentations include visual and auditory information at the same time. Just as the modality effect in cognitive load theory asserts that effective working memory capacity can be increased by using auditory and visual working memory together rather than using one or the other alone. We can explain this effect within the framework of working memory. There are two areas for storage of limited information in working memory: one for auditory data and one for visual data. So when we explain visuals with words in text, we overload the visual working memory; when we explain visuals with words in audio, we balance our data between the two working memory subsystems, thereby maximizing the limited capacity of working memory. However, in a recent meta-analysis of the modality effect made by Ginns (2005), it is found that this effect will depend on the complexity of the material as well as the learner's ability to control the pace of presentation of the information [14], that is, the modality effect only exists when studying complex learning material. And if the learner can determine his or her learning steps, then this effect cannot be found. The modality principle is also unsuitable for learners who are studying in a second language. These learners benefit more from taking the time they need to read onscreen text rather than listening to audio narration.

One disadvantage of audio should be mentioned here, which is its transiency. Once it plays, there is no record of the words, at this point, on-screen text seems to support memory better. So in the situation where learners need to refer to words over time, it's better to use text. Take the directions to exercises for example, when learners are doing an exercise, it is better to put exercise directions and feedback in the text-not the audio format so that the learner can easily refer back to the directions anytime needed [15]. Meanwhile, whenever use audio to describe the image information, it is better to have a replay button so the learner can replay as they want.

C. *Animation*

It is usually believed that animation can display the

changing of matter dynamically as well as attract the attention of learners, therefore, in multimedia learning, it should be an effective information presentation. However, the effectiveness of the animation is not such a simple definition. Kalyuga (2007) has found that continuous animations could be too cognitively demanding for novice learners because of high levels of transitivity. So if static and dynamic visualizations are equivalent in terms of provided supportive information, novice learners would benefit more from studying a set of static diagrams [16]. For more knowledgeable learners, available knowledge structures may help them to handle the transitivity of animated instructions. On the other hand, details displayed in static graphics may need to be integrated and reconciled with knowledge base of these learners imposing additional working memory demands, so it is better for these proficient learners to use animation.

D. Authenticity of Multimedia Learning Environment

Constructivism learning theory assumes that authentic learning environment can promote the learning. However, we should remember that authentic learning environment often contains too much information, some are detailed and unrelated to learning task and engender high cognitive load. Mayer and Moreno(2002) found that students learn more deeply from learning materials when extraneous words, sounds, and video are excluded rather than included, this is the coherence principle [17]. The theoretical rationale is that the learner may attend to the irrelevant material and therefore have less cognitive resource available for building mental connections between relevant portions of the information.

The authenticity in multimedia learning environment is embodied not only in physical environment, but also in social experience. It is found that at an unconscious level, people tend to process more deeply when they are in a social-like setting. One way to engage the learners socially is to use onscreen characters, called learning agents. Clark and Mayer found that the appearance of learning agents does not make much difference, we can achieve the same learning effectiveness with simple images, what matters is the agent's voice. Agents that use conversational audio narration in a familiar human accent rather than onscreen text or machine-generated language have better learning results [18].

E. Individualization in Multimedia Learning

The principle of individual differences aims at adapting elements in multimedia environment to different ability levels and preferences of learners. In the framework of working memory, learners' prior knowledge and their cognitive spatial abilities are two fundamental factors influencing the design of multimedia learning. The effects of prior knowledge are related to the higher level of processing. As Sweller and Mayer conclude from their research, learners with higher prior knowledge are able to compensate for weaknesses of learning environment because they do not need so many cognitive resources for mental model construction as learners with lower level prior knowledge [19]. With regard to spatial ability, it is found that learners with high spatial ability performed better when words and graphics were presented simultaneously rather than successively, whereas learners with low spatial abilities showed no differences in both kinds of presentation [20]. This finding suggests that learners with high spatial abilities can integrate verbal and pictorial information better than learners with low spatial abilities.

V. CONCLUSION

The findings reviewed in this paper and the corresponding guidelines indicate that hypermedia may be effective only if used in a sensible way. This paper is only a tentative introduction, further researches on hypermedia learning are demanding urgently.

REFERENCES

- [1] Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 2, pp. 89–195). New York: Academic Press
- [2] Baddeley, A., & Logie, R. (1999). Working memory: The multiple-component model. In A. Miyake, & P. Shah (Eds.) *In A Models of working memory: Mechanisms of active maintenance and executive control* (pp. 28–61). New York: Cambridge University Press.
- [3] Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Science* (Vol. 4, No. 11)
- [4] Paivio, A. (1986). *Mental representations: A dual coding approach*. New York: Oxford University Press.
- [5] Sweller, J. (1994). Cognitive load theory, learning difficulty and instructional design. *Learning and Instruction*, 4, 295–312.
- [6] Sweller, J. (1999) *Instructional Design in Technical Areas*. Melbourne, Australia: ACER Press.
- [7] Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). Expertise reversal effect. *Educational Psychologist*, 38, 23–31.
- [8] Mayer, R. E. (2001). *Multimedia learning*. New York: Cambridge University press.
- [9] Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31–48).

New York: Cambridge University.

- [10] Schnotz, W. Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction*, 13, 141–156.
- [11] Stephan Dutke, Mike Rinck (2006) Multimedia learning: Working memory and the learning of word and picture diagrams *Learning and Instruction* 16 (2006) 526-537
- [12] Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, 8, 293-332.
- [13] Brunye', T. T., Taylor, H. A., Rapp, D. N., & Spiro, A. B. (2006). Learning procedures: The role of working memory in multimedia learning experiences. *Applied Cognitive Psychology*, 20, 917–940.
- [14] Ginns, P. (2005). Meta-analysis of the modality effect. *Learning and Instruction*, 4, 313–331.
- [15] Clark, R. C. Mayer, R. E.(2007). *E-Learning and the Science of Instruction: Proven Guidelines for Consumers and Designers of Multimedia*

Learning, Business & Economics.

- [16] Kalyuga (2007). Expertise Reversal Effect and Its Implications for Learner-Tailored Instruction. *Educational Psychology Review*, 19:509–539
- [17] Mayer, R. E, Moreno, R (2002). Animation as an Aid to Multimedia Learning, *Educational Psychology Review*, Vol. 14, No. 1, March 2002
- [18] Clark, R. C. Mayer, R. E.(2007). *E-Learning and the Science of Instruction: Proven Guidelines for Consumers and Designers of Multimedia Learning, Business & Economics*.
- [19] Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 19–30). New York: Cambridge University Press
- [20] Gyselinck, V., Jamet, E., & Dubois, V. (2008). The Role of Working Memory Components in Multimedia Comprehension. *Applied Cognitive Psychology*, 22, 353–374